



UNIVERSITAT DE BARCELONA

Exploring molecular mechanisms of key targets in metabolic and infectious diseases

Elnaz Aledavood

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

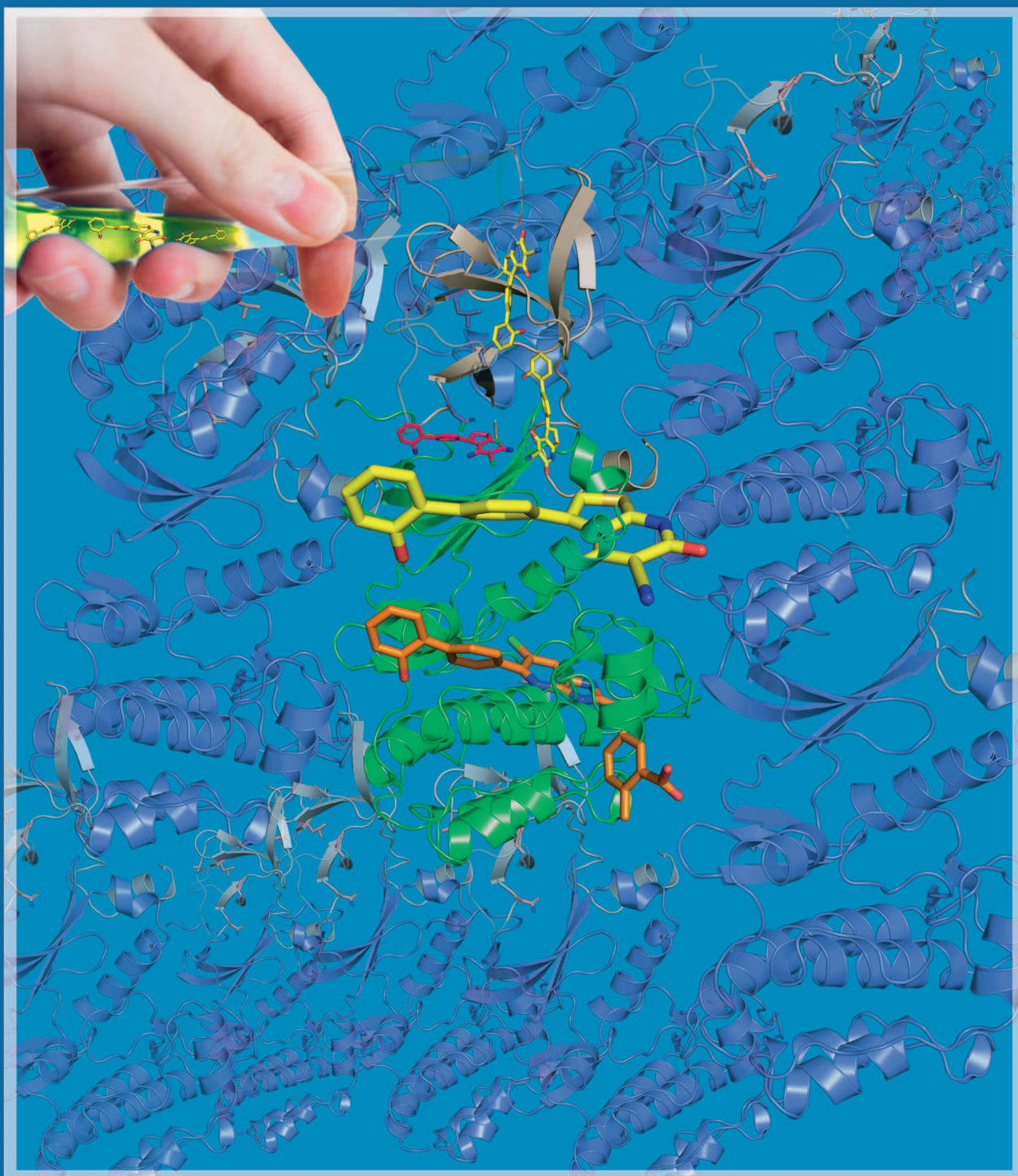
ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



UNIVERSITAT DE
BARCELONA

**Exploring molecular mechanisms of key targets in
metabolic and infectious diseases**



Elnaz Aledavood
2021



UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA i CIÈNCIES DE L'ALIMENTACIÓ
(Programa de Doctorat en Biotecnologia)

EXPLORING MOLECULAR MECHANISMS OF KEY TARGETS IN
METABOLIC AND INFECTIOUS DISEASES

ELNAZ ALEDAVOOD

2021

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA i CIÈNCIES DE L'ALIMENTACIÓ

PROGRAMA DE DOCTORAT EN BIOTECNOLOGIA

*Exploring molecular mechanisms of key targets in metabolic and
infectious diseases*

Memoria presentada por Elnaz Aledavood para optar al título de doctor por la
Universidad de Barcelona

Prof. F. Javier Luque Garriga
Director
Tutor

Prof. Carolina Estarellas
Director

Elnaz Aledavood
Doctorando

ACKNOWLEDGEMENT

Thank God for gifting me the life and understanding, by virtue of which could discriminate the good from the bad and prescribed how to promote my wisdom through learning, and dedicated me wise and faithful teachers to teach me the knowledge, and lead me to the enlightenment, under its glow, I started my journey toward perfection which is nothing unless serving human being.

May God give me the ability to be good and to remain so, to continue my journey on finding the keys of secrets in the universe to ease the human pains and agonies.

I would like to express my gratitude to my great instructor and supervisor, Professor Luque. He is a valuable and knowledgeable person, from whom I learnt great lessons; the art of a teacher is not just teaching science whereas during unforgettable working period with him, humanity in all its aspects became meaningful to me. He replaced all my worries, about being in a distant country, with kindness and support. Today, I wholeheartedly pay tribute to his scientific and human status and wish him a long life with health and happiness. I could have not imagined a better supervisor and mentor for my Ph.D. study and preparation of this thesis would have never been possible without his constructive suggestions, continual encouragement and assistance. Muchas gracias por todo. Nunca olvidaré su apoyo durante estos 5 años.

Many thanks to Prof. Carolina Estarellas. Dear Carolina, I sincerely express my gratitude for sharing your valuable scientific and managerial experiences with me. You have shown me that no goal will be achieved without consensus and cooperation. An important part of the progress of this research is indebted to your perfect training, efficient workshops, conferences, discipline and follow-up. You will be always a symbol of effort for me to know that fatigue has no meaning in the dictionary of a successful woman.

These acknowledgements could not be complete without mentioning my colleagues: Dr. Antonio Viayna, Dr. Constantí Seira and Dr. William Zamora (office 6, Gaudi building). It was a great pleasure working with them and I appreciate their ideas, help and good humour. Thanks for all the fun that we had in the last five years. I hope we will continue discussing for years. I will always remember the first classes teaching with Toni and the first molecular dynamics simulations that I performed by the help of Constantí.

Many thanks to Dr. Javier Vázquez, Dr. Federico Issoglio and David Vílchez for sharing valuable data and being so helpful.

Thanks to Prof. Axel Bidon-Chanal for his computing and scripting recommendations, his computer skills and for making our cluster works properly.

I want to thank Dr. Tiziana Ginex for her computational tips and also for all the delicious Sicilian sweets.

I would like to thank Dr. Salomé Llabrés for organizing interesting and useful seminars.

I am grateful to my Italian friends, especially my bff Eleonora and Alessia for all the great moments that we have shared together and for the special recipes of Tiramisu and pesto that they have taught me.

Many thanks to Prof. Ivan Rivalta for giving me the opportunity to stay in his research group at the University of Bologna and teaching me the Perturbation network analysis. I sincerely thank him for his hospitality.

Finally, I would like to thank my Iranian friends, Milad, Mahsa, Pouya, Sepideh, Hamidreza, Rezvan, Hamid and Neda. You have always been there for me through the good and the bad times. You are not just my friends, you are my family. Thank you for making so many ordinary moments, extraordinary.

I believe this is not the end; it's only the beginning of something great.

This thesis is dedicated

to my kind mother, who caressed my body with her hands and polished my soul with her way of thinking, and flying towards today's radiant thoughts and successes is achieved over the dream wings she has made for me.

And to my dear father, whose presence has always been a great support in the path of life. Faith and belief in being able, has always been defined to me because of him.

to my dear husband Shahin, the great gift of God with the meaning of love and a reason for everything in me. A journey has begun with him that the beauty of all its moments is due to his patience, perseverance, and kindness, and every moment I learnt from him how to be who I am.

to my beautiful and loving sister, Anita, who is my best friend at every moment, who filled my childhood's world with lots of joy and happiness and my heart today is full of our memories.

to the unique grandparents whose shadow is always a safe and comfortable image of life for me and their kindness became the blessing and tenderness of our lives.

to my husband's parents whose presence in all difficulties of life made a salve of love and kindness.

to my dear uncles, Maziar, Mehran and Payam, whose creativity and arts have always inspired me.

تقریب به هر ستمی که با فلک شمس استیم هم در نوع تقدیرش بود
لذت که در آنجا بود و در وقت برآمدن صبح با هوا و با کسب آن که اولیای شمس است
و تقریب به هر عزیزی که در جوشش است و در هر حال که در اینجاست و اعتقاد به بدست گرفتن و تقسیم
با او مغان شده است.

تقریب به هر عزیزی که در جوشش است، همین نزلت فلک شمس است و در طبع هر عزیزی که با او ستمی است
که زین می نامند و شمس در اینجاست و در هر حال که در اینجاست و تقسیم به هر عزیزی که در جوشش است
تقریب به هر عزیزی که در جوشش است و در هر حال که در اینجاست و تقسیم به هر عزیزی که در جوشش است
خط و کتابت.

تقریب به هر عزیزی که در جوشش است و در هر حال که در اینجاست و تقسیم به هر عزیزی که در جوشش است
و در هر حال که در اینجاست و تقسیم به هر عزیزی که در جوشش است
تقریب به هر عزیزی که در جوشش است و در هر حال که در اینجاست و تقسیم به هر عزیزی که در جوشش است
تقریب به هر عزیزی که در جوشش است و در هر حال که در اینجاست و تقسیم به هر عزیزی که در جوشش است

Summary

This doctoral thesis is focused on disclosing molecular factors that regulate the functional mechanism of key targets in i) metabolic disorders, such as diabetes and obesity, and ii) tuberculosis (TB). These diseases have a high impact on the public health, and TB is one of the diseases that causes highest mortality rates based on the World Health Organization reports. Accordingly, gaining insight on functionally relevant protein structures with key roles in these diseases and understanding their regulatory mechanisms can pave a way to identify novel therapeutic targets.

Regarding the metabolic disorders, I have specifically studied the direct activation mechanism of adenosine-monophosphate activated protein kinase (AMPK) by different activators, such as A-769662, SC4 and PF-739, since AMPK plays an important role in cellular energy homeostasis and it is directly related to diabetes type 2 and obesity. These studies were carried out to explore the molecular basis of the selective isoform activation of AMPK, especially focusing on the role played by the two different isoforms of β subunit. The first part of this research was focused on the direct activation by A-769662 compound in $\alpha 2\beta 1$ isoform, and the second part addressed the important issue of isoform selectivity, it being dedicated to examine the structural and dynamical properties of $\beta 1$ - and $\beta 2$ -containing AMPK complexes formed with A-769662, SC4 and PF-739. The results revealed the mechanical sensitivity of the $\alpha 2\beta 1$ complex, in contrast with a larger resilience of the $\alpha 2\beta 2$ species. Moreover, binding of activators to $\alpha 2\beta 1$ promotes the pre-organization of the ATP-binding site, favoring the adoption of activated states of the enzyme.

In another part of the research related to AMPK, in collaboration with Prof. Ana Castro from Institute of Medicinal Chemistry and Prof. María S. Fernández-Alonso from Complutense University of Madrid, I studied a novel indolic compound synthesized and experimentally analysed by their group as modulator of endothelial AMPK that acts as a mixed-type inhibitor. It means that this modulator, IND6, may bind the ATP-binding site leading to competitive inhibition of the enzyme but also suggest that IND6 may regulate the AMPK activity through binding to an additional pocket.

Regarding TB, I have studied the truncated hemoglobin N (trHbN) of *Mycobacterium tuberculosis* (Mtb). These studies aimed to disclose innovative therapeutic approaches by identifying the molecular basis of the nitrosative stress resistance of Mtb. Our hypothesis was that a decrease in the nitric oxide (NO) resistance should affect the survival of the bacillus, increasing the efficacy of current therapeutic treatments. Nitric oxide detoxification (NOD) is carried out by trHbN, which converts NO into the harmless nitrate anion, becoming essential for the defense mechanism of Mtb. However, two points prevent the use of this system as therapeutic target. On the one side, trHbN is not druggable and there is the risk that a drug

targeting trHbN may also affect other hemoglobins. On the other side, the reductase protein that helps trHbN to restore the ferrous state required to initiate the NOD is still unknown. Our studies have led to the identification of three reductase candidates for NOD activity. We have built the 3D models for the three candidates and performed the protein-protein docking in order to obtain a reductase-trHbN complex. The important role played by one of these candidates, ndh, has been highlighted in several experimental studies. This led us to focus our attention in the trHbN-ndh complex and establish the steps that could help to advance in the definition of new therapeutic strategies for TB.

Resumen

Esta tesis doctoral se centra en el análisis de los factores moleculares que regulan el mecanismo funcional de dianas clave en: i) trastornos metabólicos, como diabetes y obesidad, y ii) tuberculosis (TB). Estas enfermedades tienen un elevado impacto en la salud pública, y en particular la TB se halla entre las enfermedades con una mayor tasa de mortalidad según los informes de la Organización Mundial de la Salud. Por lo tanto, conocer las estructuras proteicas que tienen una función clave en estas enfermedades y comprender sus mecanismos reguladores puede allanar el camino para identificar nuevos objetivos terapéuticos.

Respecto a los trastornos metabólicos, he estudiado el mecanismo de activación directa de la proteína quinasa activada por adenosina-monofosfato (AMPK) por diferentes activadores, como A-769662, SC4 y PF-739, ya que AMPK juega un papel importante en la homeostasis energética celular y está directamente relacionada con diabetes tipo 2 y obesidad. Estos estudios se llevaron a cabo con el objetivo de explorar los factores moleculares que regulan la activación selectiva de isoformas de AMPK, centrándome especialmente en el papel que desempeñan las dos isoformas de la subunidad β . La primera parte de esta investigación se centró en la activación directa por el compuesto A-769662 en la isoforma $\alpha 2\beta 1$, y la segunda parte se dedicó a examinar la base molecular de la selectividad entre isoformas, considerando específicamente las propiedades estructurales y dinámicas de los complejos de AMPK que contienen isoformas $\beta 1$ y $\beta 2$ formados con A-769662, SC4 and PF-739. Los resultados revelaron la sensibilidad mecánica del complejo $\alpha 2\beta 1$, en contraste con una mayor resiliencia de la especie $\alpha 2\beta 2$. Además, la unión de los activadores en la isoforma $\alpha 2\beta 1$ promueve la organización previa del sitio de unión de ATP, favoreciendo la adopción de estados activados de la enzima.

Asimismo, en colaboración con la Prof. Ana Castro del Instituto de Química Médica y la Prof. María S. Fernández-Alonso de la Universidad Complutense de Madrid, estudiamos un novedoso compuesto indólico sintetizado y analizado experimentalmente por sus grupos, que actúa como inhibidor de tipo mixto y es capaz de modular el AMPK endotelial. Nuestros cálculos sugieren que este modulador, IND6, puede unirse al sitio de unión de ATP que conduciría a una inhibición competitiva del enzima, pero también apuntan que IND6 puede regular la actividad de AMPK mediante la unión a un bolsillo adicional.

En cuanto a TB, he estudiado la hemoglobina truncada N (trHbN) de *Mycobacterium tuberculosis* (Mtb). El objetivo ha sido obtener un nuevo enfoque terapéutico mediante la identificación de la base molecular de la resistencia al estrés nitrosativo de Mtb. Nuestra hipótesis es que una disminución de la resistencia al óxido nítrico (NO) debería afectar la supervivencia del bacilo, aumentando la eficacia de los tratamientos terapéuticos actuales. La desintoxicación del óxido nítrico (NOD, de sus siglas en inglés) se realiza mediante trHbN, que transforma NO en el inocuo anión nitrato, convirtiéndose en un mecanismo de defensa esencial

de Mtb. Sin embargo, dos puntos impiden el uso de este sistema como diana terapéutica. Por un lado, la trHbN no puede ver modificada su acción mediante ningún fármaco ya que carece de bolsillos que permitan enlazar el fármaco, y segundo, porque sería muy difícil intentar modificar únicamente esta hemoglobina en el paciente sin provocar efectos secundarios. Por otro lado, aún se desconoce la proteína reductasa que ayuda a restaurar el estado ferroso de trHbN requerido para iniciar el NOD. Mi objetivo ha sido identificar esta reductasa. Por esta razón, los estudios se han centrado en la identificación de tres posibles candidatos capaces de permitir la actividad NOD de Mtb. Una vez identificados, se han construido modelos 3D para cada uno de ellos y realizado el acoplamiento proteína-proteína para obtener un complejo reductasa-trHbN de partida. Durante esta investigación, varios trabajos experimentales apuntaron la importancia de uno de los candidatos identificados, la reductasa ndh. Este hecho nos permitió centrar nuestra atención en el complejo trHbN-ndh y establecer los pasos que podrían ayudar a avanzar en la definición de una nueva estrategia terapéutica para TB.

GENERAL INDEX

Index

1. Introduction	25
1.1. Disease is a (<i>harmful</i>) relationship of molecules	25
1.2. Adenosine Monophosphate-activated Protein Kinase	29
1.2.1. AMPK structure	29
1.2.2. AMPK activation mechanisms	32
1.2.2.1. AMPK direct activators	34
1.2.2.2. AMPK modulator (IND6)	35
1.3. Truncated hemoglobin N of the Mycobacterium Tuberculosis	36
1.3.1. TrHbN structure	36
1.3.2. Steps of the NOD function inside the trHbN	38
1.4. How does computational chemistry help drug design?	40
1.4.1. Quantum Calculations	42
1.4.2. Molecular Dynamics	43
1.4.2.1. Restrained Electrostatic Potential (RESP)	45
1.4.2.2. Ensembles	46
1.4.3. Monte Carlo	46
1.4.4. Homology Modeling	46
1.4.4.1. SWISS-MODEL	47
1.4.4.2. I-TASSER	48
1.4.5. Protein-protein interaction	48
1.4.5.1 ClusPro	49
1.4.6. Structural analysis	49
1.4.6.1. Essential dynamics	50
1.4.6.2. Dynamical cross-correlation matrix	50
1.4.6.3. Interaction energy network	50
2. Objectives	55
2.1. AMPK project	55
2.2. TrHbN project	55
3. Publications	59
3.1. AMPK	59
3.1.1. Paper 1: “Understanding the Mechanism of Direct Activation of AMP-Kinase: Toward a Fine Allosteric Tuning of the Kinase Activity”	60
3.1.2. Paper 2: “Structural Basis of the Selective Activation of Enzyme Isoform: Allosteric Response to Activators of β 1- and β 2-Containing AMPK Complexes”	77
3.1.3. Paper 3: “Elucidating the activation mechanism of AMPK by direct pan-activator PF-739”	109
3.1.4. Paper 4: “Novel Indolic AMPK Modulators Increase Nitric Oxide Release in Human Endothelial Cells”	129
3.2. TrHbN	161
3.2.1. Paper 5: “Searching for Novel Targets in Tuberculosis: Potential Reductase Partner in the Nitric Oxide Detoxification Activity by Mycobacterium tuberculosis Truncated Hemoglobin N”	162
4. Results summary	181
4.1. Study of the direct activation mechanism of AMPK	181
4.1.1. Structural Analysis	183
4.1.2. Dynamics of AMPK complexes	185
4.1.3. Comparing the network of interaction in β 1 and β 2 containing complexes	188
4.1.4. Identifying a new AMPK modulator	192
4.2. Exploring the probable reductase partner for trHbN	194
4.2.1. Searching for the reductase candidates	194

4.2.2. Building the 3D structures	195
4.2.3. Protein-protein interaction	196
5. Discussion	203
5.1. AMPK	203
5.1.1. Study of the direct activation mechanism of AMPK	203
5.1.2. AMPK modulator	206
5.2. Exploring the reductase partner for trHbN	207
6. Conclusion	211
6.1. AMPK	211
6.2. TrHbN	212
Bibliography	215

ABBREVIATIONS

<i>Abbreviation</i>	<i>Definition</i>
ADP	Adenosine di-phosphate
AMP	Adenosine monophosphate
AMPK	Adenosine monophosphate-activated protein kinase
ATP	Adenosine tri-phosphate
ADaM	Allosteric drug and metabolic
AMPK	Adenosine monophosphate-activated protein kinase
CBM	Carbohydrate binding module
CADD	Computer-aided drug design
CASP	Critical assessment of protein structure prediction
CTD	C-terminal domain
CBS	Cystathionine β -synthase domain
DFT	Density Functional Theory
DCCM	Dynamical cross-correlation matrix
ET	Electron transfer
E.COLI	Escherichia coli
ED	Essential dynamics
FF	Force field
GB	Generalized Born
HA	Hemagglutinin
Hb	Hemoglobin
KD	Kinase domain
LBDD	Ligand-based drug design
LKB1	Liver kinase B1
MD	Molecular dynamics
MC	Monte Carlo
Mtb	Mycobacterium tuberculosis
NA	Neuraminidase
NAI	Neuroaminidase inhibitors
NO₃⁻	Nitrate anion
NO	Nitric oxide
NOD	Nitric oxide dioxygenase
PDB	Protein data bank
QM	Quantum mechanics
RMSF	Residue mean square fluctuation
RMSD	Root mean square deviation
SBDD	Structure-based drug design
3D	Three-dimensional
TMH	Transmembrane helix
trHbN	Truncated hemoglobin N
TB	Tuberculosis
TSC	Tuberous sclerosis complex 2
T2D	Type 2 diabetes
WISP	Weighted Implementation of Suboptimal Paths



INTRODUCTION

1. Introduction

'Life is a relationship between molecules, not a property of any one molecule.

So is therefore disease, which endangers life'

(Zuckerkandle and Pauling 1962)

Every year millions of people suffer from any kind of illnesses. Generally speaking, there are four main types of diseases, namely deficiency, infection, hereditary and physiological diseases. According to the World Health Organization's (WHO) reports, the four diseases with highest mortality rate in the world include cancer, diabetes, influenza and tuberculosis.¹ Besides these known diseases, every few decades some new illnesses such as COVID-19 appear and threaten human's life.

Determining how many people die each year and the reason of their death is very important in order to catch the attention of the Public Health and the scientists to take an action to face these diseases. Despite of the discovery of novel medications, failure of the drug treatment is still an important challenge for biomedicine scientists and pharmaceutical companies. This failure can happen due to different reasons such as wrong diagnosis, selection of low efficacy drugs and drug resistance.² Amongst these reasons, drug resistance is a growing issue that reduces and even precludes the suitability of the therapeutic treatment selected to deal with the disease.³ Indeed, drug resistance is a life threatening phenomena; so there is urgent need for finding new therapeutic strategies.⁴ In the near future, the duty of scientists and pharmaceutical companies is to deal with this key problem in order to find new efficient strategies that can tackle this issue.

1.1. Disease is a (*harmful*) relationship of molecules

In the life bodies there exist four main classes of biomacromolecules, which include proteins, lipids, carbohydrates and nucleic acids. The proper functioning of these four kinds of macromolecules is of outstanding importance for the normal development of life. In this thesis, we have centered our efforts in the study of selected proteins, which were chosen due to their key functional roles.

Proteins are described based on their large range of functions in the body. They can be listed in five categories, including antibody, enzyme, messenger, structural component and transporter or storages.⁵ Protein function directly relies on its three-dimensional (3D) fold. The primary structure is formed by the sequence of amino acid residues, which contain apolar alicyclic and aromatic chains, and various polar groups, such as alcohols, thiols, carboxamides, imidazole, carboxylic acid, amine and guanidine. These groups are chemically reactive and, when the amino acid residues are properly arranged in the 3D fold, they provide series of

Introduction

physicochemical properties that are important for the protein functions. Normally, proteins interact with one or more molecules in order to exert their function (like enzymes, which catalyze chemical reactions) and mediate metabolic and signaling pathways and cellular processes.⁶ Because of this role in biological functions, the study of protein interactions provides a better understanding of basic cellular biochemistry and physiology,⁷ since they control the mechanisms related to healthy and diseased states in the body.⁸ The dynamics of protein binding pockets are important for their interaction specificity. Proteins can be adapted due to the interaction with small organic molecules, namely ligands, since they are structurally flexible. A cavity on the surface or in the interior of a protein that has suitable properties for binding a ligand is usually referred to as a binding pocket (**Figure 1**). The physicochemical characteristics of a binding pocket is determined by the set of amino acid residues around it and together with its shape and location in protein, it defines the protein functionality.⁹

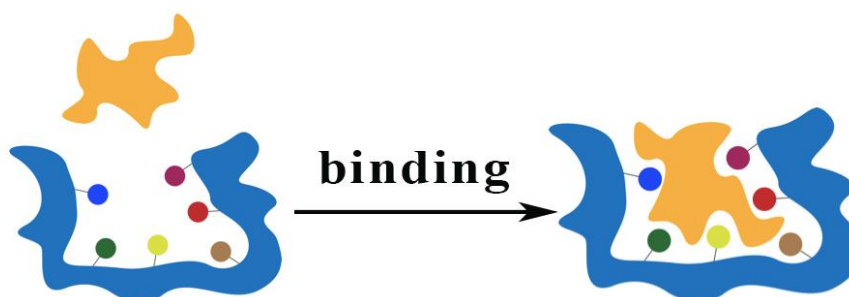


Figure 1. Schematic representation of the formation of the binding site demonstrating the effect of the aminoacids (represented by different colour spheres) in the physicochemical properties of the binding cavity. The ligand is represented by yellow form.

Moreover, protein folding and misfolding that occur in different kinds of diseases, including Alzheimer and Parkinson, are associated with the ability of the protein to form/break specific noncovalent interactions, such as hydrogen bonding, ionic pairs and van der Waals contacts. On the other hand, the ability of a protein to bind to a ligand also depends on the formation of these noncovalent interactions (**Figure 2**). Therefore, they play a fundamental contribution in dictating the binding affinity between the interacting partners, such as the interaction between proteins with small molecules, peptides, proteins, nucleic acids, etc.

For this reason, the understanding of protein non-covalent interaction network and how this network affects the binding affinity can help us to elucidate the molecular basis of diseases. These interactions can be altered by mutations affecting the binding site of small molecules, or the interface that mediates recognition with other macromolecules. This may result in biochemically dysfunctional changes, affecting the binding mode/affinity that may cause different diseases.^{10,11} Altogether, it is clear that unraveling the molecular mechanism and the role of these interactions is an outstanding ingredient to understand the infection strategies of

pathogens,¹² and necessary to prevent, diagnose and treat the main diseases which are nowadays affecting to the world population.¹³

Type of Interaction	Model	Example
(a) Charge–charge Longest-range force; nondirectional		—NH_3^+ —COO^-
(b) Charge–dipole Depends on orientation of dipole		—NH_3^+ H_2O
(c) Dipole–dipole Depends on mutual orientation of dipoles		H_2O H_2O
(d) Charge–induced dipole Depends on polarizability of molecule in which dipole is induced		—NH_3^+ C_6H_6
(e) Dipole–induced dipole Depends on polarizability of molecule in which dipole is induced		H_2O C_6H_6
(f) Dispersion Involves mutual synchronization of fluctuating charges		C_6H_6 C_6H_6
(g) van der Waals repulsion Occurs when outer electron orbitals overlap		C_6H_6 C_6H_6
(h) Hydrogen bond Charge attraction + partial covalent bond		$\text{N—H} \cdots \text{O}=\text{C}$ Hydrogen bond length

Figure 2. Different types of non-covalent interactions. Adapted from Bioinformatics III notes from the Institute of Bioinformatics, Johannes Kepler University Linz written by Noura Chelbat.¹⁴

Evaluating the affinity of ligands to the primary target is a key aspect in drug discovery to find and optimize the ligands. Binding kinetics of ligands (drugs) is calculated with a rate constant of ligand association (k_{on}) and ligand dissociation (k_{off}). The equilibrium dissociation constant (K_d) is the ratio of the dissociation to the association rate constant, determining the fraction of the receptor occupancy at a given ligand concentration.¹⁵

There are two common situations with respect to target-ligand equilibria. In the first both association and dissociation occur in single steps, so that receptor and ligand combine to shape a binary complex with association and dissociation rate constant k_1 (k_{on}) and k_2 (k_{off}) (eq1).



Introduction

The second situation consists of receptor isomerization after initial ligand binding, which results in a higher binding affinity (eq 2). In this mechanism, the ligand encounters the receptor in a conformational state that is suboptimally complementary to the ligand for binding. After the initial encounter (RL), the receptor undergoes conformational change to a more committed state (R*) where the new binary complex (R*L) has a more competent binding affinity than RL.



In this case, the values of k_{on} and k_{off} are not specified by a single microscopic rate constant. In particular, the k_{off} is composed of both reverse and forward rate constants in receptor isomerization such that (eq 3):

$$k_{off} = \frac{k_2 k_4}{k_2 + k_3 + k_4} \quad (3)$$

Two scenarios can be considered for the analysis of the kinetic rate constant that define the (un)binding of a ligand to its target receptor.¹⁶ In the first one, which is called closed system, the total concentration of receptor and ligand are constant over time under strict control of the variables of the system. The only change that happens over time is the concentration of free (R + L) and bound (R*L) species as the system reaches the equilibrium. The equilibrium dissociation constant is used to approximate the target-drug complex lifetime. In the second one, which is named open system that would better represent the situation in the living context, the receptor is kept at a fixed concentration while the ligand concentration can vary, so the equilibrium measurements are not feasible because the concentration of the ligand continuously changes as a result of factors such as absorption and metabolic clearance.¹⁷ A drug is efficacious only as long as it is bound to, and modulating the action of its physiological target. Therefore, the binding of drugs to their macromolecule targets is considered as paramount for pharmacological activity.

Taking all these elements together, in this thesis, we gained exposure to research in exploring the molecular mechanisms of key targets in metabolic (studying the adenosine mono-phosphate activated protein kinase) and infectious diseases (studying the reductase partner of the truncated hemoglobin N) by means of *in-silico* techniques and computational chemistry. In the next sections, the two main targets of our research will be explained followed by the description of the computational methods that assists us to complete this thesis.

1.2. Adenosine Monophosphate-activated Protein Kinase

Protein kinases are very important in nearly every aspect of cell biology since they can modify the function of a protein.¹⁸ Mammalian adenosine monophosphate-activated protein kinase (AMPK) is a Ser/Thr protein kinase that has an important role in cellular energy homeostasis and its activity is tightly regulated by the cellular ratios between AMP, ADP and ATP.¹⁹ When AMPK is activated, it reduces the rate of the anabolic pathway and up-regulate the catabolic pathways leading to the raise of the ATP levels (**Figure 3**). In the physiological condition, the ATP concentration is about 100-fold higher than AMP and 10-fold higher than ADP, that is, AMPK binds to AMP/ADP more strongly than it binds to ATP,²⁰ and AMP/ATP ratio becomes a sensitive indicator of the energy equilibrium inside the cell.²¹

Because of its critical function in metabolism, AMPK is a central regulator of different metabolic pathways and it is implicated in various metabolic disorders such as obesity, insulin resistance, type 2 diabetes (T2D) and cardiovascular diseases. The activation of AMPK plays a crucial role in increasing glucose uptake and fatty acid oxidation, while reducing the synthesis of fatty acids, protein and cholesterol.²²

Several studies have reported different results about the role of AMPK in cancer. It has been shown recently that AMPK is phosphorylated and activated by the upstream kinase, liver kinase B1 (LKB1), which is a tumor suppressor. Additionally, another tumor suppressor, tuberous sclerosis complex 2 (TSC) is activated and phosphorylated by AMPK.²³ So, AMPK can act as a tumor suppressor or promoter depending on the stage and type of the cancer.²⁴

Due to the importance of AMPK in different metabolic pathways and disorders, this protein has become a potential therapeutic target in the past decades²⁵ and for this reason was one of the selected targets to be studied during this thesis.

1.2.1. AMPK structure

AMPK is a 145 kDa heterotrimeric complex which consists of a catalytic α -subunit and two regulatory subunits β and γ . All three subunits are encoded by multiple genes in mammals ($\alpha 1$, $\alpha 2$; $\beta 1$, $\beta 2$; $\gamma 1$, $\gamma 2$, $\gamma 3$). So a diverse collection of twelve $\alpha\beta\gamma$ heterotrimers can exist.²⁵ The activities of each complex are not fully understood but there are different expression patterns across tissues. For instance, $\gamma 1$ and $\gamma 2$ are widely expressed in the body, but expression of $\gamma 3$ is restricted to skeletal muscle.²⁶

The α catalytic subunit has a canonical Ser/Thr kinase domain (KD) at the N-terminus, an autoinhibitory domain (AID), an adenine nucleotide sensor segment termed as α -linker, and a α subunit-interacting C-terminal domain (α -CTD) which is necessary for the formation of the complex with the other two subunits (**Figure 4. A**).

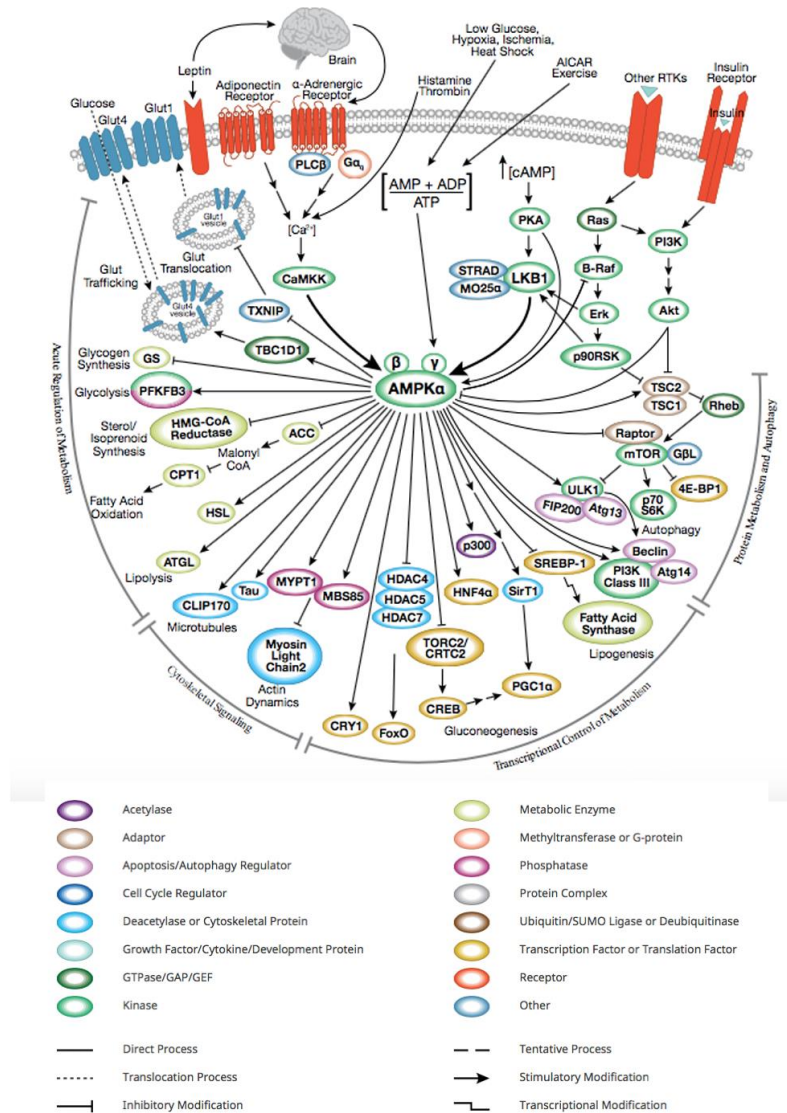


Figure 3. AMPK signaling pathway. Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com).

KD has a highly conserved structure consisting of a smaller N-terminal lobe composed of a β -sheet and the αB and αC helices, and a larger α -helical C-terminal lobe. Key regulatory elements of the KD are the activation loop (aa 160-175) at the entrance of the catalytic cleft, the αC -helix (aa 55-71) in the N-lobe, and the peptide substrate-binding catalytic loop, commonly known as P-loop (aa 15-33). The activation loop of the α -subunit has the amino acid residue Thr-172, which its phosphorylation has been considered as an essential step in AMPK activation. The activation loop locates at the interface between the catalytic and nucleotide-binding modules, which is also close to the C-terminal of β -subunit. This arrangement makes the phosphorylation and dephosphorylation of Thr172 sensitive to conformational rearrangements induced by nucleotide binding.²⁷ The P-loop and the αC -helix shape an ATP-binding site in the protein structure. Also, it is suggested that the induced formation of C-

interacting helix in the β -subunit through the binding of an activator in the ADaM site, stabilizes the α C-helix and thus results in the activation of AMPK.²⁷

The β subunit has a central carbohydrate binding module (CBM, **Figure 4**) which causes AMPK to associate with glycogen particles.²¹ It is also known as glycogen-binding domain. The physiological function of this subunit is still unclear but some scientists suggest that it allows AMPK to sense the availability of glycogen reserves in the cell.²⁵ The CBM contacts the N-lobe of the KD and the interface between these two modules shapes a distinct pocket named allosteric drug and metabolic (ADaM) binding site that was reported as the binding site for many direct AMPK-activating compounds.²⁸ The C-terminal domain (CTD) of the β -subunit acts as a scaffold for the interaction of the complex and interacts with the α - and γ -subunits.

The γ -subunit (**Figure 4**, colored in orange) has four tandem repeats of cystathionine β -synthase domain (CBS) forming two Bateman motifs, which have head-to-head arrangement. The CBS domains form four adenine nucleotide binding sites.²⁹ Cocrystal of the mammalian AMPK reveals that site two is always empty, site four is tightly bound to AMP and site one and three can bind to nucleotides depending on their concentration. Mutagenesis studies show that site 3 and 4 are important for the AMPK allosteric activation.³⁰ AMPK is a highly dynamic complex with a stable core formed by the γ -subunit and the α - and β -CTD. Attached to the core there are moveable domains whose position is determined by ligand binding and post-translational modifications.

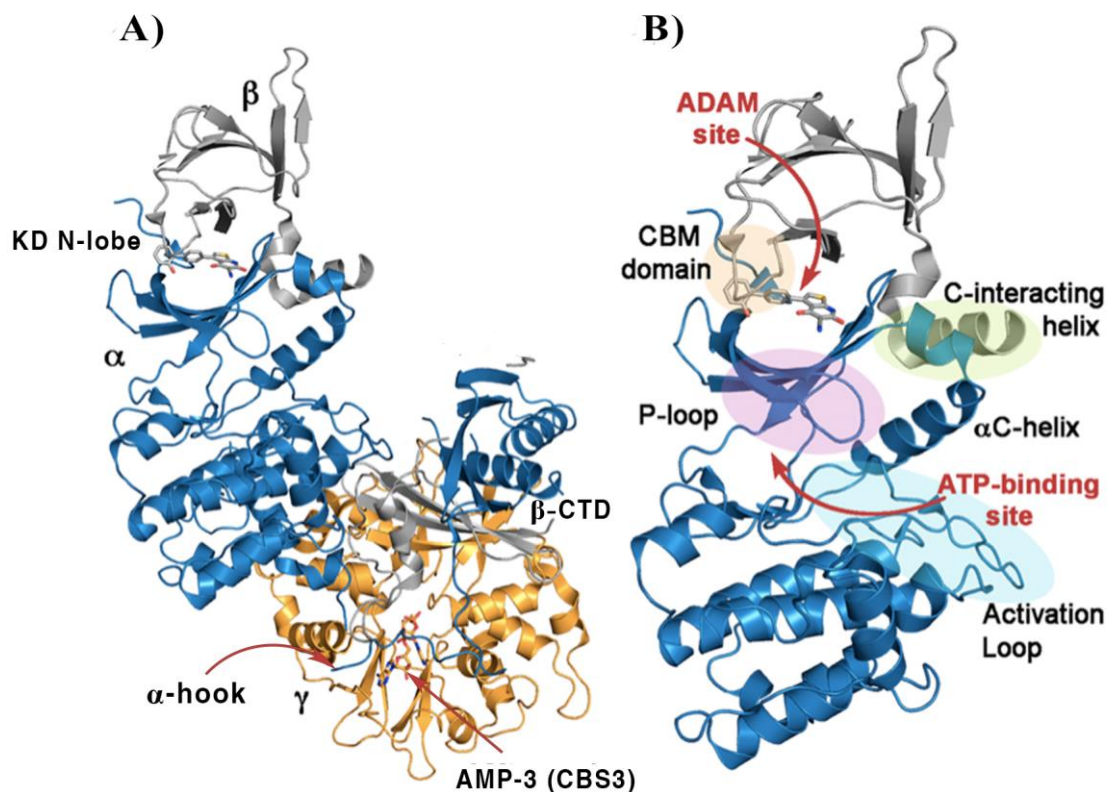


Figure 4. A) AMPK heterotrimeric complex composed of subunits $\alpha 2$ (blue), $\beta 1$ (gray), and $\gamma 1$ (orange) bound to the A-769662 activator (PDB code 4CFF). (B) Representation of the simulated system selected for the study of the direct activation mechanism, which is composed of subunits $\alpha 2$ and $\beta 1$. The most important regions of these subunits are highlighted; P-loop (purple), activation loop (cyan), CBM domain (orange) and C-interacting helix (green).

1.2.2. AMPK activation mechanisms

Because of its crucial role in cellular energy homeostasis,³¹ the AMPK activity is regulated by different mechanisms.³²

a) Allosteric activation: AMPK activity is regulated at three different levels: (i) activation loop phosphorylation by upstream kinases, (ii) protection against activation loop dephosphorylation by protein phosphatase, and (iii) at the level of phosphorylation-independent, allosteric kinase activation.³³ The main form of regulation is done by phosphorylation of Thr172 in the activation loop segment of the kinase domain of the α subunit (see **Figure 4B**). First, AMP binds to the CBS domain in the regulatory γ subunit and then phosphorylation of Thr172 by upstream kinases such as LKB1 and calcium/calmodulin-dependent protein kinase converts the inactive form of AMPK to a form that is several thousand-fold more active. When AMP or ADP binds to site 3 of the γ subunit, the α -hook (**Figure 4A**), which pertains to the α -subunit, makes a conformational change leading to the interaction between the

activation loop and the C-terminal regions of the α and β subunits. Therefore, access to the phosphorylated Thr172 in the activation loop is blocked and phosphorylated AMPK will be maintained in an active state.³⁴ Binding of AMP to the site 1 increase the activity of AMPK more than 2-5 fold but the mechanism of this activation is still unknown.

b) Indirect activation: AMPK is activated by different drugs like metformin, phenformin and oligomycin in intestines and liver. The activation occurs through increases the intracellular levels of AMP. In this process, disruption of complex I of the mitochondrial respiratory chain occurs which result in reducing the ATP synthesis and an increase in the intracellular AMP:ATP ratio.³⁵

c) Direct activation: In 2006 Barbara Cool et al. reported a novel mechanism of action that showed the first direct activation of AMPK by the thienopyridone drug A-769662.³⁶ This compound binds to a binding site called allosteric drug and metabolite-binding site (ADaM site (**Figure 4. B**) located at the interface between α and β subunits. The activation mechanism of A-769662 is independent from phosphorylation of Thr172, but it requires autophosphorylation of Ser108 in the CBM domain of the β -subunit.³⁷ Binding of A-769662 can increase the activity of AMPK more than 90-fold where the Ser108 phosphorylated is present. The phosphorylated Ser108 does not directly participate in the drug binding but it forms hydrogen bonds with α -Thr21, α -Lys29 and α -Lys31, which probably contribute to the stability of the binding pocket. Although neither A-769662 nor AMP, they cannot individually activate the dephosphorylated AMPK (**Figure 5A**); together they increase the activity of AMPK more than 1000-fold (**Figure 5B**).³⁸

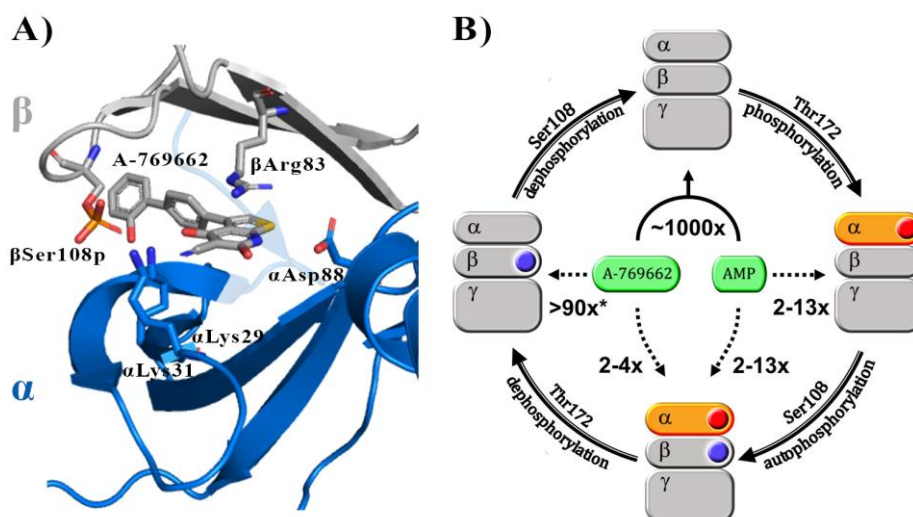


Figure 5. A) Allosteric Drug and Metabolite (ADaM) binding site of the direct activator A-769662. Relevant residues of the α and β subunits are shown as blue and grey sticks, respectively. B) Hypothetical model of allosteric regulation of AMPK by AMP and A-769662 described by Scott and coworkers.³⁹

1.2.2.1. AMPK direct activators

In the recent years, different kinds of direct activators of AMPK were reported (**Figure 6. A**), which their activation mechanism remains only partially understood. Moreover, some of them are selective for specific AMPK complex isoforms while others can activate various isoforms. Among all the direct activators obtained in the last years, we have centred our efforts in the study of A-769662, SC4 and PF-739 to gain insight into the allosteric mechanism that modulates the direct activation of AMPK by small compounds. Our aims are i) to identify the molecular factors that govern the ligand binding, ii) to comprehend the relationship between activator binding and structural changes in the protein, iii) to assess the impact of these changes in the enzyme activity and finally, and iv) to study the different sensitivity of AMPK isoforms specially regarding the $\beta 1$ and $\beta 2$ subunits.

A769662: Different hypotheses have been suggested in order to explain the activation mechanism induced by binding of A-769662 to AMPK. It was assumed that A-769662 could induce conformational changes in AMPK structure, which result in the rearrangement and stabilization of the activation loop. Binding of this compound facilitates the interaction between the αC -helix, from the α -subunit, and the C-interacting helix, from the β -subunit, which leads to the allosteric activation of AMPK.⁴⁰ Furthermore, phosphorylated Thr172 can be protected against dephosphorylation upon binding of A-769662 because it enhances the interaction between CBM and kinase domain. It was also reported that A-769662 is only active in $\alpha 2\beta 1\gamma 1$ isoform of AMPK and it is not active in the $\alpha 2\beta 2\gamma 1$, at least up to concentrations of $10\mu\text{M}$. Despite all this information, the details of the conformational rearrangement triggered by binding of A-769662 have not been demonstrated until we studied this system during this thesis.

SC4 is a potent activator of AMPK complexes including the $\beta 2$ isoforms and it stimulates the glucose uptake by skeleton muscles. SC4 can activate the complexes containing either isoforms of β but it shows selectivity preference for isoform $\alpha 2$ with respect to the $\alpha 1$ irrespectively from the β isoform. SC4 synergistically activates dephosphorylated $\alpha 1\beta 1\gamma 1$ in the presence of AMP; however, AMP/drug synergistic activation did not translate to the $\alpha 2\beta 2\gamma 1$ isoform complex.⁴¹

PF-739 is a pan-activator for AMPK. Pan activators are expected to activate AMPK in all cells of body and they are not dependent on the α - and β -subunits isoform selection.⁴² This activator is a benzimidazole derivate that can activate both $\alpha 2\beta 1\gamma 1$ and $\alpha 2\beta 2\gamma 1$ complexes,⁴³ though it still exhibits a larger affinity for the $\beta 1$ -containing isoforms. PF-739 seems also to have a larger effect on the $\alpha 2$ -containing AMPK complexes.

Despite the mentioned information, the details about the activation mechanism that is induced by the binding of these ligands to the ADaM site, as well as the different dynamical behaviour that their binding could promote between isoforms is still unknown. Understanding the

molecular mechanisms that determine the regulatory effect of direct activators, especially targeting selectively a given isoform, is of utmost relevance for gaining insight into the tissue-dependent modulation of AMPK. Since the ADaM site is shaped only by α and β isoforms and the lack of precise structural information about C-terminal stretches of both α and β subunits, which are located close to the γ subunit, we did not consider the γ subunit in our simulation.

1.2.2.2. AMPK modulator (IND6)

In another part of our studies, which were performed in collaboration with the group of Prof. Ana Castro from Institute of Medicinal Chemistry and Prof. María S. Fernández-Alonso from Complutense University of Madrid, we studied a novel indolic compound synthesized and experimentally analysed by their group as modulator of endothelial AMPK that shows affinity for $\alpha 1\beta 1\gamma 1$ isoform, acting as a mixed-type inhibitor (**Figure 6. B**). IND6 might bind the ATP-binding site results in competitive inhibition of the enzyme, but may also regulate the AMPK activity through binding to an additional binding pocket.⁴⁴ So, the binding modes of IND6 have been investigated by molecular dynamics simulations by allocating this modulator in both the ATP-binding site and the ADaM site of the $\alpha 1\beta 1\gamma 1$ AMPK.

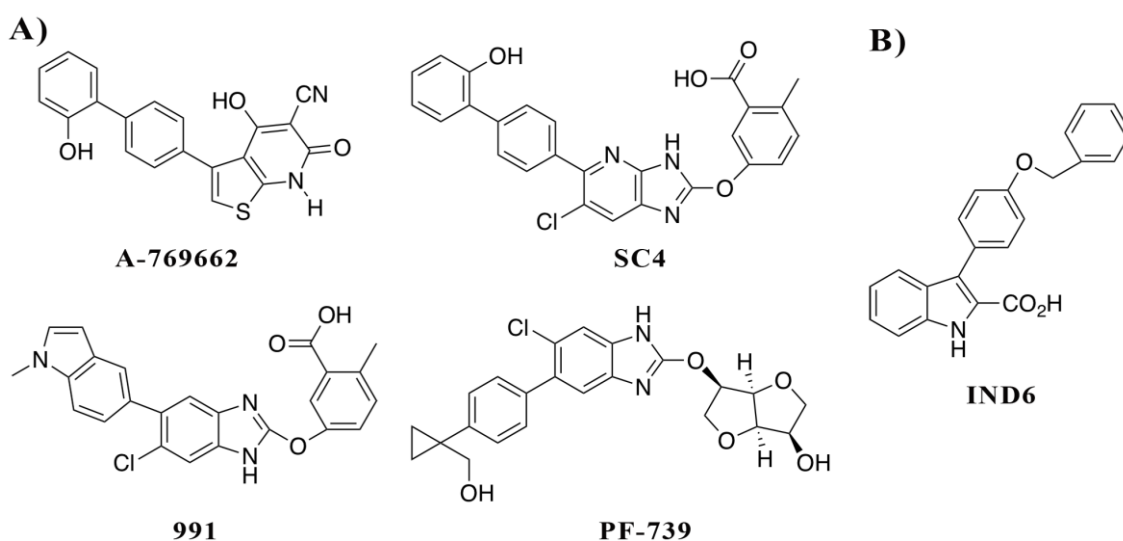


Figure 6. A) Chemical structures of different AMPK activators. B) Chemical structure of AMPK modulator, IND6.

1.3. Truncated hemoglobin N of the *Mycobacterium Tuberculosis*

Tuberculosis (TB) is amongst the world wide health threats. Around 10 million people are affected each year and it is the ninth leading cause of death worldwide.⁴⁵ It is caused by *Mycobacterium tuberculosis* (Mtb).⁴⁶ Current treatments are still faced with multi-drug resistant restraints and also drug side effects, especially in countries lacking the long and costly treatment adapted to patients.⁴⁷ Although different researches lead to discovery of novel compounds with inhibitory properties against different targets, TB is still one of the challenging diseases around the world.⁴⁸ Nowadays around 20 drugs are used in clinical trials.⁴⁹ The multi-drug resistance of *Mycobacterium* is the reason that these treatments are become required for the patients for almost two years with all the toxicity and expensive cost included.

The pathogenicity of Mtb is related to the remarkable adaptability to face with hazardous level of reactive nitrogen/oxygen species within the intracellular environment.⁵⁰ During the growth infection age, level of nitric oxide (NO) and nitrogen species, which are produced by activated macrophages, limit the intracellular survival of Mtb and it may involves in the restricting the bacteria into latency. But, Mtb has an efficient defence mechanism by which it can escape from the toxicity of NO. The reaction of NO with the oxygenated truncated hemoglobin N (trHbN) has been suggested as the defence mechanism that the organism can avoid the toxic effect of NO.⁵¹ It converts NO to harmless nitrate anion (NO_3^-). This process which is called nitric oxide dioxygenase (NOD) reaction, is crucial in relieving nitrosative stress.⁵²

1.3.1. TrHbN structure

Truncated hemoglobins are a distinct group in the hemoglobin (Hb) superfamily. The trHbN is a small heme protein that is encoded by the *glbN* gene.⁵³ It has a tertiary structure which is based on a two-over-two α helical fold consisting of 128 residues⁵⁴ (α -helical building the structure of trHbs are labeled as A-F according to their sequential order; **Figure 7**)⁵⁵. HisF8 (81) is bound to the heme group and it is the only residue that is conserved between the Hb and trHb families. The heme stabilization is achieved through hydrogen bonds to ThrCD4(49) and Ala(75), as well as by salt bridges to ArgE6(53) and ArgFG3(84).⁵⁶

According to the previous computational studies conducted in our research group, trHbN showed a new dual-path mechanism to drive migration of O_2 and NO to the heme cavity via the short and long branches (Figure 7).^{12,51} The O_2 reaches to the heme cavity through the tunnel short branch, which is approximately 10 Å and is shaped by residues in helices G and H (see **Figure 7**). Binding to the heme is conducted by opening the long branch with 20 Å length that is formed of helices B and E where a ligand-induced conformational change regulates the opening of the tunnel branch via PheE15(62) residue that acts as a gate for ligand entry. The ability of PheE15 gate to switch between the open and closed states depends on the dynamics and conformational changes induced by Pre-A motif.⁵⁷

As the NOD function of trHbN relies on the dispersion of NO to the O₂-bound heme via the tunnel long branch, the PheE15 plays an important role as a fundamental residue in specifying the efficiency of the NO scavenging. The closing and opening of the gate result in the NOD function of the trHbN.^{58,59}

Due to the binding of O₂, the dynamical motion of the protein backbone is changed. The important motions of the deoxygenated trHbN influence C, G and H helices. On the other hand, the contribution of the flexibility in the oxygenated protein comes from helices B and E, which form the walls of the tunnel long branch. Thus, the transition between open and closed state of the gate should be facilitated by the increased motion of these helices.⁶⁰

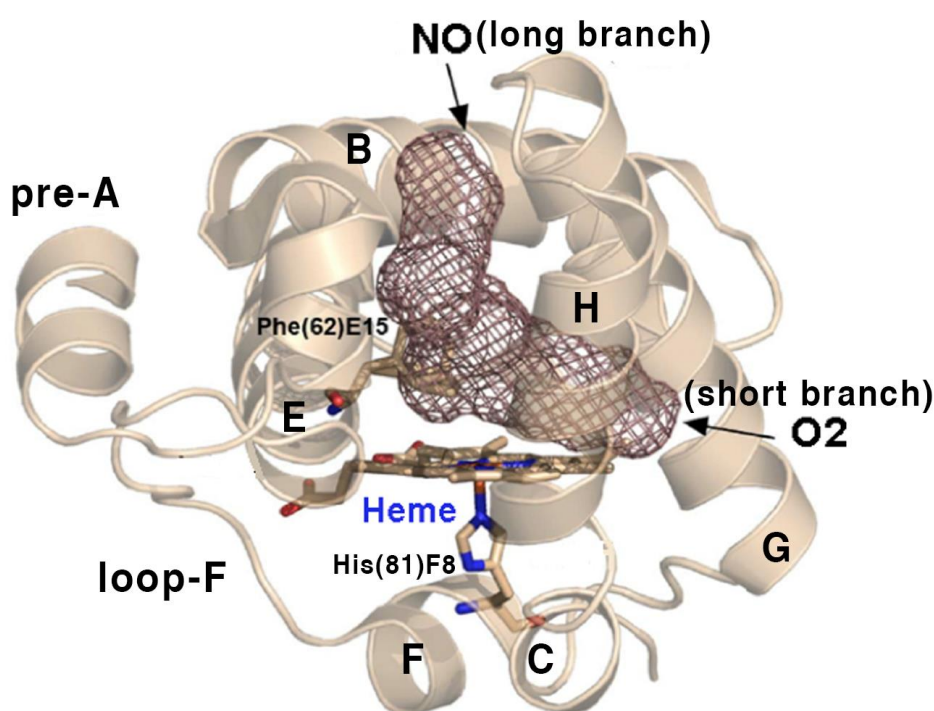


Figure 7. X-ray structure of trHbN (PDB entry 1IDR) showing the labels of the helices, the location of the heme group and the location of the two branches of the tunnel system (brown mesh). The heme, PheE15 and HisF8 are shown as sticks.

The tunnel system of the protein simplifies the ligand access to the heme site⁵⁶ and the flexible N-terminal pre-A motif composed of 12-residue is needed for the optimal NOD activity.⁵⁷ In 2009, the computational studies reveals that if the pre-A motif is removed from the trHbN structure, the movement of the protein backbone changes significantly. The difference is detected around helices B and E and also the EF loop region.⁵¹ These helices, which define the long branch, showed smaller fluctuation in the trHbN, deleted pre-A compared with the wild type trHbN, which may affect the function of the long tunnel branch and the PheE15 gate. The analysis of the trajectories revealed that the major motions in trHbN deleted N-terminus are

related to the helices G and H and loop CD. Therefore, removal of the Pre-A motif change the protein backbone motion and the PheE15 gate is forced to remain in a closed conformation, leading to the blockage of the easy access of NO toward the active site.^{57,61} Although deleting the pre-A motif of trHbN can affect the function of this protein; this is not a feasible way to stop the NOD activity. So, it is necessary to find an alternative solution.

1.3.2. Steps of the NOD function inside the trHbN

The conversion process of NO catalyzed by trHbN is considered to involve the following steps:

- 1.- O₂ enters to Hb via the short tunnel branch and oxyferrous heme is generated.
- 2.- Binding of O₂ triggers dynamical alterations in the protein backbone that regulate the opening of the long branch facilitating access of NO to the O₂-bound heme. The oxyferrous heme interacts with NO and nitrate anion and ferric heme are produced.
- 3.- An electron will be shuffled from NADH group of the flavin domain (or a reductase partner) to the heme to covert the heme to the ferrous state and recycle the process. Both oxidative and reductive processes affect the efficiency of the NOD function.⁶²⁻⁶⁴

The NOD function depends on the stabilization of O₂ into the deoxygenated heme, a rapid migration of NO to the heme cavity, and a fast electron transfer from an electron donor in order to recover the ferrous state to ferric and restart a new cycle.^{65,66} So, other members of the Hb superfamily can perform NO deoxygenation when they pair with an appropriate electron donor.⁶⁷ Despite lacking of a reductase domain, the NO dioxygenation efficiency of trHbN is high ($K_{\text{NOD}} \sim 745 \mu\text{M}^{-1} \text{S}^{-1}$) and it is comparable with the flavoHbs that already contain a reductase domain.^{68,69} The question is *how trHbN can carry out so efficiently the NOD function that relies on the structural and functional compatibility with a reductase partner?*

To the best on our knowledge, there is no information about the reductase partner of trHbN that can catalyze the reduction of trHbN. Furthermore, the effect of interaction between trHbN and the reductase partner on the migration of the ligands (NO and O₂) through the protein and how this interaction can affect the conversion to the nitrate anion are still unknown.

To see if the survival of Mtb requires the expression of a specific reductase partner, Singh *et al.* carried out a research to confirm this hypothesis. Since trHbN is functionally active in *Escherichia Coli* (*E. coli*) and it helps its host to metabolize NO and convert it to nitrate anion, it should be able to use reductases to carry out NO dioxygenation, although this reductase has not been identified yet in its native host. For an efficient electron transfer, the reductase should have a correct coordination and structural alignment with trHbN globin domain and the FAD cofactor of reductase and heme of trHbN binding regions should place at an acceptable distance for an efficient electron transfer. So, they examined the efficiency of different reductases in assisting the electron transfer process. The results pointed out a marked dependence of the

efficiency of this process on the nature of presence of a reductase. Specifically, this study showed that NADH-ferredoxin reductase (FdR) could perform efficiently the reduction of ferric to ferrous trHbN. This interaction is mediated by flexible CD loop and the dynamical motion of the protein backbone is crucial for facilitating the interaction which is modulated by the Pre-A region of trHbN.⁶²

Based on these evidences and the point that trHbN i) is not a druggable target, due to the lack of druggability hallmarks in its structure which restricts its suitability for drug discovery purposes, and ii) it belongs to the Hb family that is found in the human body and can result in unwanted side effects, we decided to start a project in order to find the reductase partner for NOD activity of trHbN.

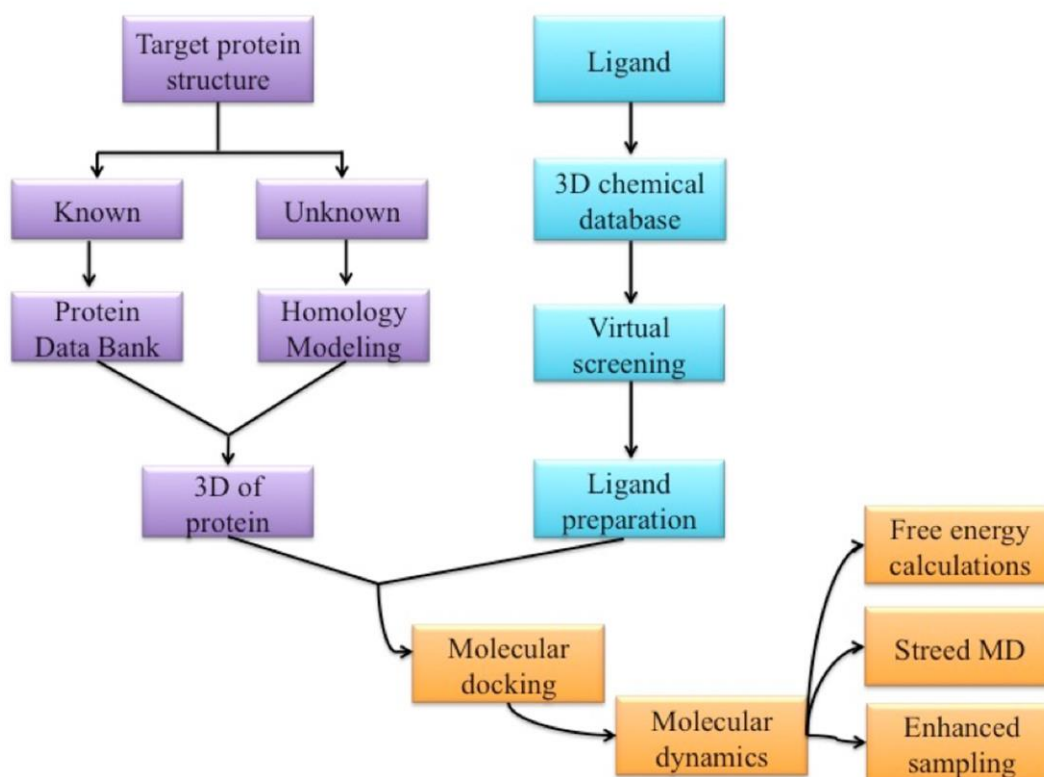
Therefore, identifying and characterizing the reductase partner is of utmost importance for either inhibiting its ability to act as redox partner or by preventing its ability form a functional complex with trHbN.

1.4. How does computational chemistry help drug design?

Knowledge of the structure-function relationships of proteins can be gained resorting to a combination of experimental biophysical techniques. However, sometimes it is not possible to monitor structure of molecules *in vivo* or *in vitro* environments and, the experimental studies are costly and time-consuming. Because of these reasons, computational chemistry emerged as an efficient tool to solve these issues in more efficient and feasible way.

Computer-aided drug design (CADD) uses computational approaches to discover, develop, and analyze drugs and similar biologically active molecules. CADD comprises a diverse range of theoretical and computational approaches that are part of modern drug discovery. Based on the structural information that we have, it is possible to follow one of the CADD areas: structure-based drug design (SBDD) and ligand-based drug design (LBDD). SBDD is related to the design and optimization of a structure with the goal of identifying a compound suitable for clinical testing like a drug candidate. LBDD approach involves the analysis of ligands (a library of ligands) known to interact with a target of interest.⁷⁰

Scheme 1 shows a typical protocol used in the SBDD. Since during the progression of this thesis I have mainly used this protocol with some differences, below I will explain the tools used for each of the two developed research projects.



Scheme 1. Structure-based drug design work flow.

The Nobel Prize in Chemistry awarded to Arieh Warshel, Martin Karplus and Michael Levitt for the development of multiscale models for complex chemical systems in 2013 demonstrates the relevant contribution of this field in the ligand-target interaction, and particularly the potential impact in the drug discovery process. Currently the combination of computational and experimental studies is a powerful approach to understand the molecular basis of many processes that encode the processes of the cell.⁷¹

Computational chemistry is a branch of Chemistry that uses the results of theoretical chemistry integrated into efficient computer programs to calculate the structures and characteristic of molecules. These characteristics include structure, interaction energy, charges, vibrational frequencies, etc. This term is also used to cover any science areas that connect computer science and chemistry and it is utilized computer simulations to assist in solving chemical problems.⁷² Computational chemistry improves our understanding of how the biomolecules interact with each other, providing a molecular basis to decipher the cell life, but is also fundamental to facilitate the design of novel molecular entities, particularly in pharmaceutical companies.

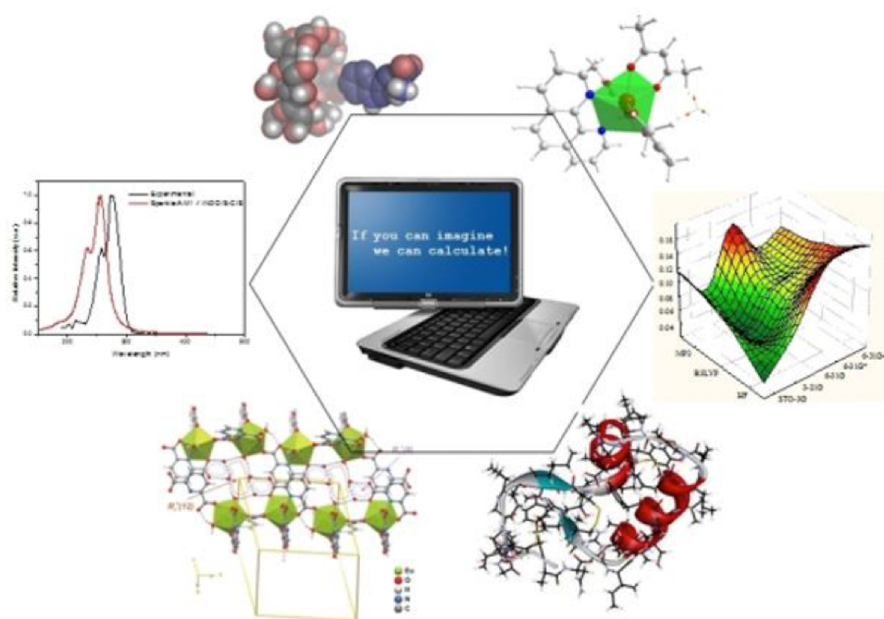
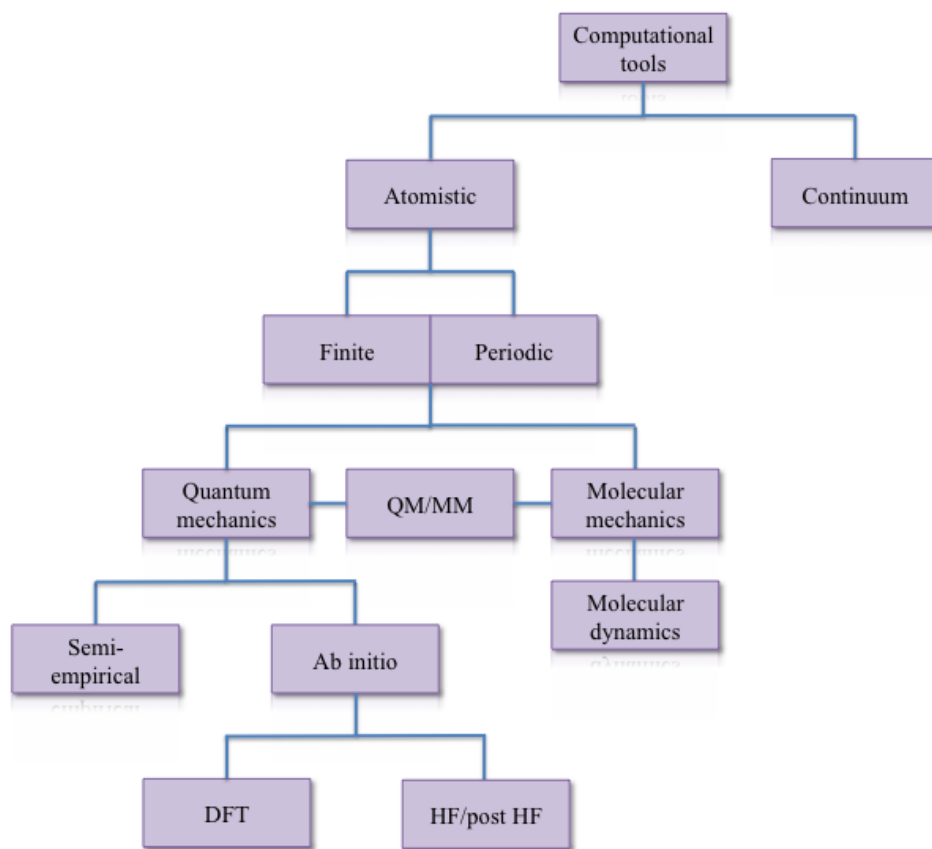


Figure 8. Schematic example of computational chemistry applications. Adopted from <https://www.solutions.bocsci.com/computational-chemistry.htm>⁷³

The computational researches can be done using high-performance computing.⁷⁴ Computational chemists can use a variety of methods depend on the size and the characterization of the systems, which can describe the system of interest at the classical and quantum mechanical levels. Quantum mechanics (QM) can explain the features of small-scale and discrete quantities and interactions, which cannot be explained by classical methods. QM is often the only theory

that can reveal the electronic structure of molecules consisting of few hundreds of atoms, including chemical reactivity. In contrast, classical methods can be utilized to study the conformational fluctuations of macroscopic systems including up to 10^6 atoms.



Scheme 2. Different types of computational chemistry tools.^{75,76}

Below, some features and the main differences of the quantum and classical methods will be briefly described.

1.4.1. Quantum Calculations

Based on the different level of first-principle physics and physical models used, various computational approaches have been developed. The most rigorous approach is based on Quantum Mechanics (QM). QM is a science of small, enabling us to explain the behaviour of matter and its interaction with energy on the atomic and sub-atomic scales resorting to both *ab initio* and semi-empirical methods. QM methods are based on the Schrödinger equation (eq. 4). This equation describes the behaviour of the electrons in a molecule and gives us the energy and a wavefunction (Ψ). The electron distribution can be calculated via the wavefunction,⁷⁷ thus providing the electronic properties of the molecule through the application of suitable operators. Due to the nature of wavefunction, it does not possess any physical meaning by itself, however, the scalar product of the wavefunction with its complex conjugate bears the notion of the

probability of finding a particle at a particular position q and time t which is known as Born interpretation.⁷⁸

$$i\hbar \frac{\partial}{\partial t} \Psi(q, t) = \mathbf{H}\Psi(q, t) \quad (4)$$

\mathbf{H} represents the Hamiltonian operator, which explains the total energy of the system at a given time and \hbar is Planck's constant divided by 2π .

Ab initio techniques attempt to compute electronic state energies and other physical properties, as function of the position of nuclei and electrons, from first principles without the use of experimental input. Although semi-empirical calculations are also based on the Schrödinger equation, this method use experimental parameters or the results of *ab initio* calculations to determine some of the matrix elements, which let to obtain faster calculations.⁷⁹ Semi-empirical methods treat at least the valence electrons explicitly, so they are able to answer questions such as electronic transitions, dipole moments, polarizability and bond breaking and forming. The computational requirements of semi-empirical methods locate between those of the *ab initio* methods and the force field techniques.

Ab initio methods include intensive computation, so tend to be limited to smaller systems (up to hundred atoms) due to practical reasons of computer time, but they are essential because semi-empirical methods are of little use on systems whose electronic properties have not been included in the data base used to construct the parameters of such models.

Density Functional Theory (DFT) calculations determine the electron density function using the Kohn-Sham formalism.⁷⁹ DFT follows a formally exact approach to solve static many-electron problem by replacing the formalism of the many-electron wave function with the electron density. The main advantage of DFT compared to the traditional wave function based methods is the fact that it substitutes the many-body problems with a problem of discovery of electron density, which relies on three spatial coordinates.

1.4.2. Molecular Dynamics

To study the large systems such as proteins, the methodologies based on Molecular Mechanics (MM) describe the molecules as a collection of classical particles, while bonds are represented as springs, thus neglecting the explicit treatment of electrons.⁸⁰ Accordingly, the energy of a molecule can be determined using a force field, which encompasses classical expressions for the different bonded (stretching, bending and torsions of bonds) and non-bonded (electrostatic and van der Waals, often supplemented with polarization) contributions.⁷⁴

In a simulated system, the potential energy depends on the interaction between atoms. The interactions between them are controlled by a parameterized force field (FF). In the force field

Introduction

the molecule is presented as an ensemble of spheres with mass (representing the atoms), which are linked by springs that indicate the interactions (eq 5).

$$\begin{aligned}
 U = & \sum_{\text{bonds}} \frac{1}{2} k_b (r - r_0)^2 + \sum_{\text{angles}} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{\text{torsions}} \frac{V_n}{2} [1 + \cos(n\phi) - \delta] \\
 & + \sum_{\text{improper}} V_{imp} + \sum_{LJ} 4\epsilon_{ij} \left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right) + \sum_{\text{elec}} \frac{q_i q_j}{r_{ij}}
 \end{aligned} \tag{5}$$

where the first four terms refer to intramolecular or local contributions to the total energy (bond stretching, angle bending, and dihedral and improper torsions), and the last two terms serve to describe the Van der Waals and the Coulombic interactions.

Molecular dynamics (MD) simulations rely on the use of classical physics law. MD simulations explore the conformational space sampling an ensemble of structures accessible under the simulation conditions. In a system composed by fixed particles, the potential energy depends on the interactions between the atoms thanks to their reciprocal positions and the potential energy varies at the variation of the geometry of the system. The sampling is performed by solving Newton's laws of motion for the system, which enables to obtain a time evolution of the starting snapshot of the system (i.e., a trajectory).

$$F = ma \tag{6}$$

where F is force on an atom, m is mass of the atom, and a is the atom acceleration.

In MD simulations, the physics of the system is simplified but it provides close approximation to treat complex molecular systems such as proteins with affordable cost and high level of reliability. The force field gives rise to the Hamiltonian of the simulated system and the bonded and non-bonded interactions happening between the particles in the system are described by analytical expressions. The modern force fields consist of at least four terms that are used to explain both the intra- and intermolecular interactions. One of the most frequently used force fields is the AMBER force field (eq. 7) in which the function of the potential energy is defined as:

$$\begin{aligned}
 V(r) = & \sum_{\text{bonds}} k_b (b - b_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 \\
 & + \sum_{\text{torsions}} k_\phi [\cos(n\phi + \delta) + 1] + \sum_{\text{nonbonded pairs}} \left[\frac{q_i q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^6} \right]
 \end{aligned} \tag{7}$$

Similarly to equation 5, in the AMBER force field the first three functions show the interactions between bonded atoms including the bonds, the angles and the torsional angles. The last functional term describes all of the non-bonded interactions, where the Coulomb law describes the electrostatic interactions and the van der Waals forces are approximated by the Lennard-Jones potential.⁸¹⁻⁸³

The ff99SB force field in Amber has been mostly used with the Amber suite of programs.⁸⁴ It implicates a set of parameters for all-atom simulations, which are suitable for protein simulations. Explicit use of all hydrogen atoms, no specific functional form for hydrogen bonding and fixed partial charges on atom centers are some characteristic features of this force field. Also, there is a better balance of secondary structure elements by improved distribution of backbone dihedrals for glycine and alanine.⁸⁵ In the modified version of this force field, which has termed ff99SB-ILDN, new torsion potentials described for four residues (isoleucine, leucine, aspartate and asparagine that appear to be the most problematic in ff99SB).

Apart AMBER, other programs for modeling various systems are available including GROMACS,⁸⁶ NAMD,⁸⁷ OpenMM⁸⁸ or Desmond⁸⁹. The AMBER package have a general so-called AMBER force field named gaff⁹⁰ which is used to parameterized small organic molecule. In this thesis, ff99SB⁹¹ and gaff were utilized in conjugation with RESP charges calculated at the B3LYP/6-31(G)⁹² levels⁹³.

1.4.2.1. Restrained Electrostatic Potential (RESP)

Force field parameters can be achieved by fitting to the experimental measurement or data calculated using high level of theory. One of the most important terms is the electrostatic one, as small errors in the atomic charges can prompt large changes in the conformational behaviour of a molecule. So, the partial charges play a crucial role in the description of our system.

Electrostatic potential (ESP) method can be used to calculate the partial atomic charges for molecular dynamics simulations. This method is based on the assignment of charges to the atomic centers using the least square fitting method. In this method, the obtained charges reproduce the electrostatic potential using quantum mechanics methods.

The drawback of this method is the weak prediction of the charges that exist at the large distance from the points on the QM-obtained ESP grid. To solve this problem, the RESP method was introduced. In this method the fitting process is done in two stages. In the first step, the partial charges of all atoms are free to change and in the second stage, the fit partial charges of all atoms are restrained except for degenerate hydrogen atoms. The advantages of this method in comparison to the ESP charges are two things; firstly the calculated values are more closer to the expected values and secondly, this method lowers the polarity of the polarized C-H bonds which result in a more correct parameterization of organic molecules.⁹⁴

1.4.2.2. Ensembles

The ensemble is described as all the possible microstates that a simulated system can occupy under specific constraints that define the system macroscopically including constant volume and constant temperature. Different ensembles are produced based on choosing the fixed variables:

1.- Microcanonical ensembles (NVE): The number of atoms (N), the volume of the system (V) and the total energy (E) are constant. In this ensemble, there is no exchange of the energy with surroundings.

2.- Canonical ensemble (NVT): The number of atoms (N), the volume of the system (V) and the temperature (T) are remained fixed. So, the exchange of the energy with environment can happen.

3.- Isobaric-isothermal ensemble (NPT): The number of atoms (N), the volume of the system (V) and the temperature (T) are kept fixed. From the chemistry point of view, this ensemble is the most natural ensemble as it represents the usual condition the same as the experimental ones.⁸³

1.4.3. Monte Carlo

Monte Carlo (MC) methods use probabilistic rules to generate a new configuration from the previous one. With the MC algorithm, a sequence of states will be generated that can be used to estimate structural and thermodynamic properties. MC simulations, except the Kinetic Monte Carlo techniques, are time-independent and dynamical properties based on the evolution over time cannot be estimated. In MC simulations, with randomly altering specific degrees of freedom, a new configuration of a system is generated.

Despite MD simulations, MC simulations are free from restrictions of solving Newton's equations of motion; this can lead to huge speedups up to 10^{10} in the sampling of equilibrium properties. Moreover, MC methods are easily parallelizable and with some techniques being ideal for use with large CPU cluster.

Regarding the disadvantages of MC simulations, because it does not solve the Newton's equation of motion, a traditional MC simulation cannot generate any dynamical information. The difficulty of conducting large-scale moves is one of the main difficulties of MC simulations of proteins in an explicit solvent. Any move that significantly changes the internal coordinates of the protein without moving the solvent particles will lead to a large overlap of atoms and the rejection of the trial configuration. Coarse-grained protein models are the most popular systems that use MC methods, because a implicit solvent do not suffer from the mentioned drawback.⁹⁵

1.4.4. Homology Modeling

All the SBDD studies relies on knowledge of the three-dimensional structures of the biological targets obtained through methods such as X-ray crystallography and NMR spectroscopy for the known structures or homology modeling for those which their 3D-structure is not available. In

the first two projects (AMPK and M2 channel), the 3D structures of the proteins were adopted from Protein Data Bank (PDB)⁹⁶, but respect to the trHbN project due to the lack of information of 3D structure of this protein, we tried to model it by homology modeling software such as I-TASSER⁹⁷ and SWISS-MODEL⁹⁸.

Protein structure predictions use two different approaches: template-based structure prediction techniques that achieve a three-dimensional and atomistic model by applying the structure-sequence relationship. De-novo structure prediction, which generate different sets of alternative structures and chooses the best scoring model using complex scoring functions.

Homology modeling predicts the three-dimensional structure of a query protein via the sequence alignment of template proteins. Most procedures include four consecutive steps:

- a) Specifying the protein structure related to the target sequence with an alignment of target/template.
- b) Building of the target sequence based on the template's information.
- c) Refinement of the model
- d) Evaluation of model quality and ranking of generated models.

All these steps might be repeated in order to obtain a satisfactory model. In traditional methods, local alignment such as BLAST was used for the alignment. These methods are only accurate for the sequence alignment above 40-50% sequence identity but when no close homologs are identified, other algorithms are needed. The new approaches are based on sequence-profile, sequence-HMM and HMM-HMM alignments. Also, there are other programs that are classified into fold recognition methods.

In this thesis, two homology modeling servers were used for build the structure of ndh reductase of Mtb.

1.4.4.1. SWISS-MODEL

In this server, the user can model a sequence with a minimum input. In the simplest case, it works only with the amino acid sequence of a target protein. Additional user input such as selecting a different template or adjusting the target-template alignment might be necessary for some complex modeling researches. The template library named ExPDB is extracted from Protein Data Bank (PDB) and the PDB coordinate files are divided into individual protein chains and unreliable entries, so the low-quality structures are removed. The software searches for the template in the template structure library. The modeling process will be split into separate independent batches if the selected templates cover distinct regions of the target sequence. Five template structures will be superposed per batch by an iterative least squares algorithm. After removing incompatible templates, a structural alignment is created. The models with high C_{α} root mean square deviations to the first template will be omitted. Finally, a local

Introduction

pair-wise alignment is done and the placement of insertions and deletions is optimized. The backbone atom positions of the template are averaged in order to build the core of the model. To model the insertions and deletions in the target-template alignment, an ensemble of fragments is constructed by constraint space programming (CSP). To build the side chains, weighted positions of corresponding residues in the template structures will be used and a scoring function checks favorable interactions such as hydrogen bonds and disulfide bridges and unfavorably close contacts to select the best conformation. At the end, an energy minimization using the GROMOS96 force field is applied to the final structure. (<https://swissmodel.expasy.org>)⁹⁸

1.4.4.2. I-TASSER

To build the model, the first step is to perform the threading. Threading means identifying template proteins from solved structure databases that have a similar structure or similar structural motifs as the target sequence. A non-redundant sequence database is used to match against the target sequence by PSI-BLAST. Then a sequence profile is created which is used to predict the secondary structure using PSIPRED. The target sequence is threaded through a PDB structure library using different servers and the best template from each threading server is chosen. In the next stage, the continuous fragments in threading alignment are taken from the template structure and these parts are used to assemble structural conformations of the sections that aligned well with the unaligned regions built by *ab initio* modeling. In the following step, the fragment assembly simulation is carried out and the lowest energy structures are selected as an input for REMO server that build the final models. In the final stage, the function of the target protein is concluded by structurally matching the built 3D model against the proteins with the known structure and function in the PDB.

The quality of the predicted model is assessed with a confidence score named C-score. C-score is defined based on the quality of the threading alignments and the coverage of the I-TASSER. (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>)⁹⁷

1.4.5. Protein-protein interaction

SBDD begins with the identification of a potential ligand-binding site on the target, which is normally a pocket or protuberance with a variety of potential hydrogen bond donors and acceptors, hydrophobic characteristics and sizes of molecular surfaces. The active site in an enzyme or an assembly site with another macromolecule can be the active site.⁷⁰ Respect to the AMPK, its binding sites were specified in the 3D structures obtained from PDB and we used the same site to locate the activators in AMPK. In the case of trHbN, we needed the protein-protein complex of this protein with its probable reductase partner. To this end, we used ClusPro⁹⁹ webserver.

Protein-protein interactions are important to understand the function and organization in the cell. Docking of two proteins generates different detailed models that specify the position of all atoms. There are two different docking methods: direct and template-based methods. Direct methods try to find the structure of the target complex that is located at the minimum of Gibbs free energy in the conformational space, so it needs an effective minimization algorithm. The template-based docking uses the rule in which interacting pairs sharing higher than 30% sequence identity interact in the same way, so the structure of a complex can be obtained by homology modeling tools when there is a correct template complex of a known structure.

1.4.5.1 ClusPro

This web-based server is used the direct docking to define the interactions between two protein. It follows three steps: i) rigid-body docking by sampling of billions of conformations. The docking uses PIPER program, which is based on the Fast Fourier Transform (FFT) correlation approach. ii) Root mean square deviation-based clustering of the 1000 structures with the lowest energy to find the largest cluster that will show the best models of the complex. iii) Refinement of the selected structures by energy minimization (**Figure 9**).⁹⁹

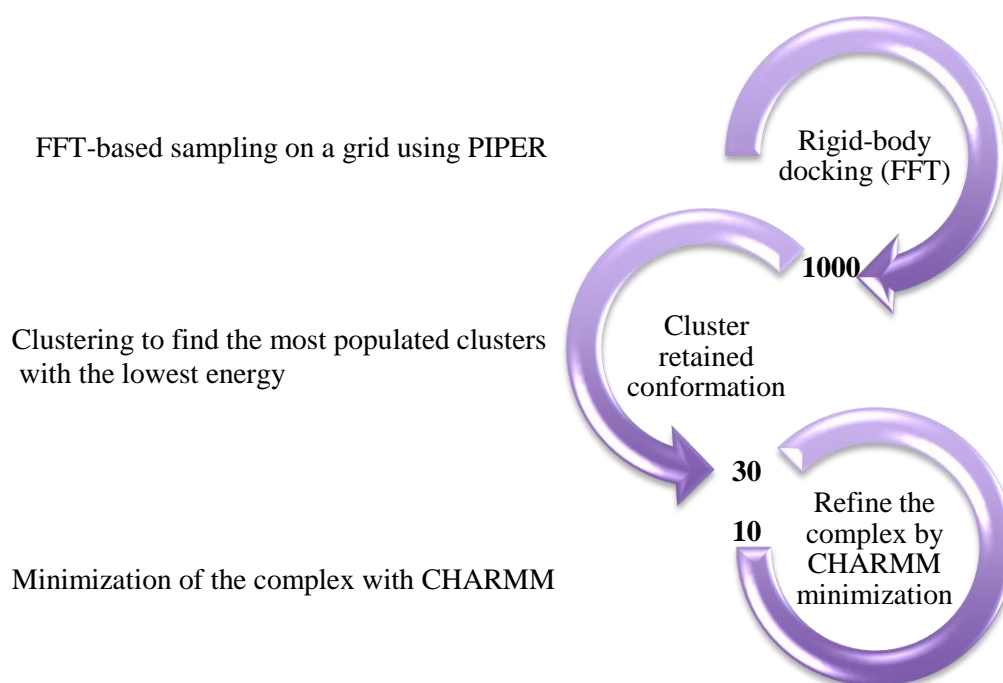


Figure 9. ClusPro algorithm. The number of structures obtained after each step were indicated in the plot.

1.4.6. Structural analysis

Once the structure and the target site were identified, we carried out different MD simulations. Finally, the produced trajectories were analysed by various methods; the general analyses including root-mean square deviation (RMSD) and residue-mean square fluctuation (RMSF) to

see the conformational changes in the protein structure and the binding site and the more specific analyses to check how binding of the ligand can affect the motions of the important structural elements and find the critical residues that play important roles in the activation of proteins. Amongst these methods, we can mention essential dynamics (ED), dynamical cross-correlation matrix (DCCM), and interaction energy network.

1.4.6.1. Essential dynamics

This method was utilized to specify the most important motions from the structural variance sampled in MD simulations. In ED,¹⁰⁰ the dynamics along the individual modes can be studied and visualized separately, so we can filter the main collective motions during our simulations. Therefore, the positional covariance matrix is created and diagonalized in order to achieve the collective deformation modes, i.e. the eigenvectors, while the eigenvalues account for the contribution of each motion to the structural variance of the protein. The calculations were performed with PCAsuite program, which is integrated in the pyPCczip program, a suite of tools for compression and analysis of molecular simulations.¹⁰¹

1.4.6.2. Dynamical cross-correlation matrix

To complement the information gained from the ED analysis, DCCM was used to examine the correlation motion of residues along a given trajectory. To this end, all the snapshots were aligned by means of least-square fitting of C α atoms of the whole protein to the equilibrated starting configuration. Then, the DCC matrix was determined as noted in Eq. 8.

$$C_{ij} = \frac{c_{ij}}{c_{ii}^{1/2} c_{jj}^{1/2}} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\left[(\langle r_i^2 \rangle - \langle r_i \rangle^2) (\langle r_j^2 \rangle - \langle r_j \rangle^2) \right]^{1/2}} \quad (8)$$

where the position vectors of two C α atoms i and j fitted in the structure at time t are denoted as $r_i(t)$ and $r_j(t)$, respectively.

The cross-correlation coefficients range from -1 to +1, which represent anticorrelated and correlated motions, respectively, whereas values close to zero indicate the absence of correlated motions.¹⁰² This analysis was performed using the module available in AMBER package.

1.4.6.3. Interaction energy network

Networks of local interactions are intrinsically linked to the structural response of proteins to external factors.¹⁰³ For our purposes, Weighted Implementation of Suboptimal Path (WISP)¹⁰⁴ was utilized to analyse the allosteric network. This method enabled us to perform a dynamic network analysis to understand how the binding of a ligand in an allosteric cavity could affect another binding site. In particular, WISP relies on the dynamical interdependence among the protein residues. To this end, each amino acid is treated as a node, which was located at the

residue center-of-mass, and the interdependence among nodes is represented as a connecting edge with an associated numeric value that reflects its strength. The interdependence is determined from a $N \times N$ matrix C (N is the number of nodes) with values corresponding to the weights of each edge, reflecting the correlated motion among node-node pairs. Finally, the weight between the edge that connects nodes i and j is expressed as $w_{ij} = -\log(|C_{ij}|)$, so that highly correlated or anticorrelated motions are characterized by small values of w_{ij} .

CHAPTER



OBJECTIVES

2. Objectives

Understanding the factors that regulate the activity of key targets in diseases is important to identify potential hits and novel therapeutic strategies against these diseases. Moreover, despite of the discovery of novel medications, treatment failure still remains a big challenge for scientists and pharmaceutical companies. Understanding the reasons of this failure, disclosing molecular mechanisms and attempting to find new targets can pave a way to tackle this challenge. In our studies we focused our attention on the study the proteins implicated in metabolic and infectious diseases, which are amongst the highest mortality rate diseases in the world according to the WHO reports.

In this context, this thesis uses the computational chemistry methods to gain insight in two different proteins structures; AMPK and trHbN, to explore the following objectives.

2.1. AMPK project

1. To gain insight on the molecular mechanisms that underlies the regulatory effect of direct activator A-769662.
2. To unveil the changes in the network of interactions between key structural elements of AMPK.
3. To explore the molecular basis of the selective isoform activation of AMPK, specially focusing on the role played by $\beta 1$ and $\beta 2$ subunits upon binding of direct activators.
4. To describe a novel indolic compound as modulator of endothelial AMPK.
5. To propose new possibilities for regulating the modulation of endothelial cells.

2.2. TrHbN project

1. To explore putative membrane proteins able to interact with trHbN and assist recycling to the ferrous state.
2. To examine the interaction interface of trHbN with its reductase partner to:
 - a) identify and characterize a binding pocket, and
 - b) assess the electron transfer rate conducted by the trHbN-reductase complex, with the aim to design a drug able to disrupt this complex, and therefore the efficiency of the NOD activity.

CHAPTER



PUBLICATIONS

3. Publications

Four papers have been compiled for the defence of this thesis, two of them have been published and the other two are submitted.

3.1. AMPK

3.1.1: In the first paper, entitled “*Understanding the Mechanism of Direct Activation of AMP-Kinase: Toward a Fine Allosteric Tuning of the Kinase Activity*”, MD simulations have been used to gain insight into the allosteric mechanism that modulates the direct activation by small compound A-769662, suggesting that the activator may act as a glue, coupling the dynamical motion of the β subunit and N-terminal domain of the α subunit and assisting the preorganization of the ATP-binding site (sections 4.1.1 and 4.1.2).

3.1.2: In the second paper, entitled “*Structural Basis of the Selective Activation of Enzyme Isoform: Allosteric Response to Activators of β 1- and β 2-Containing AMPK Complexes*”, we examined the molecular basis of the selective isoform activation of AMPK and the structural and dynamical properties of β 1 and β 2 containing AMPK complexes formed with A-769662 and SC4. The findings were discussed in light of the changes in the residue content of β subunit isoforms, particularly regarding the β 1Asn111 \rightarrow β 2Asp111 mutation as a key factor (sections 4.1.2 and 4.1.3).

3.1.3. In the third paper, entitled “*Elucidating the activation mechanism of AMPK by direct pan-activator PF-739*”, we have assessed the structural and mechanical properties of β 1- and β 2-containing AMPK complexes towards the binding of a pan-activator PF-739. The analysis let us to understand how the PF-739 can activate the AMPK through both β -isoforms and help us to discern the slightly higher affinity for β 1-isoform, due to the β 1Asn111 to β 2Asp111 substitution as a critical factor in modulating the dynamical sensitivity of the β -isoforms. (sections 4.1.2 and 4.1.3)

3.1.4: The fourth paper, entitled “*Novel Indolic AMPK Modulators Increase Nitric Oxide Release in Human Endothelial Cells*”, report the identification of a novel AMPK modulator, IND6, which shows affinity for AMPK α 1 β 1 γ 1. The findings explained the ability of IND6 to act as a mixed-type inhibitor that can also promote the enzyme activation by adopting two distinct binding modes at the ADaM site (section 4.1.4)

3.1.1. Paper 1: “Understanding the Mechanism of Direct Activation of AMP-Kinase: Toward a Fine Allosteric Tuning of the Kinase Activity”

Elnaz Aledavood,^{†,||} Gleiciane Moraes,^{‡,||} Jeronimo Lameira,[‡] Ana Castro,[§] F, Javier Luque,^{*,†}
and Carolina Estarellas^{*,†}

[†]Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicina (IBUB) and Institute of Theoretical and Computational Chemistry (IQTCUB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain

[‡]Faculdade de Ciências Naturais, Campus Universitário do Marajó-Breves, Universidade Federal do Pará(CUMB-UFPA), Breves, Brasil

[§]Instituto de Química Médica, Consejo Superior de Investigaciones Científicas (IQM-CSIC), 28006 Madrid, Spain

Published in Journal of Chemical Information and Modeling, **2019**, 59 (6), 2859-2870

Understanding the Mechanism of Direct Activation of AMP-Kinase: Toward a Fine Allosteric Tuning of the Kinase Activity

Elnaz Aledavood,^{†,||} Gleiciane Moraes,^{‡,||} Jeronimo Lameira,^{‡,||} Ana Castro,[§] F. Javier Luque,^{*,†,||} and Carolina Estarellas^{*,†,||}

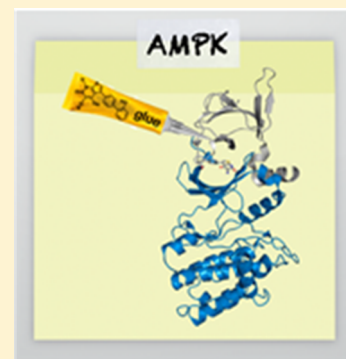
[†]Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicina (IBUB) and Institute of Theoretical and Computational Chemistry (IQTUCUB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain

[‡]Faculdade de Ciências Naturais, Campus Universitário do Marajó-Breves, Universidade Federal do Pará (CUMB-UFPA), Breves, Brasil

[§]Instituto de Química Médica, Consejo Superior de Investigaciones Científicas (IQM-CSIC), 28006 Madrid, Spain

S Supporting Information

ABSTRACT: Mammalian AMP-activated protein kinase (AMPK) is a Ser/Thr protein kinase with a key role as a sensor in cellular energy homeostasis. It has a major role in numerous metabolic disorders, such as type 2 diabetes, obesity, and cancer, and hence it has gained progressive interest as a potential therapeutic target. AMPK is a heterotrimeric enzyme composed by an α -catalytic subunit and two regulatory subunits, β and γ . It is regulated by several mechanisms, including indirect activators such as metformin and direct activators such as compound A-769662. The crystal structure of AMPK bound to A-769662 has been recently reported, suggesting a hypothetical allosteric mechanism of AMPK activation assisted by phosphorylated Ser108 at the β -subunit. Here, we have studied the direct activation mechanism of A-769662 by means of molecular dynamics simulations, suggesting that the activator may act as a glue, coupling the dynamical motion of the β -subunit and the N-terminal domain of the α -subunit, and assisting the preorganization of the ATP-binding site. This is achieved through the formation of an allosteric network that connects the activator and ATP-binding sites, particularly through key interactions formed between α Asp88 and β Arg83 and between β pSer108 and α Lys29. Overall, these studies shed light into key mechanistic determinants of the allosteric regulation of this cellular energy sensor, and pave the way for the fine-tuning of the rational design of direct activators of this cellular energy sensor.



■ INTRODUCTION

Mammalian adenosine monophosphate-activated protein kinase (AMPK) is a Ser/Thr protein kinase with a key role as a sensor in cellular energy homeostasis.¹ Once activated, AMPK promotes a reduction in the rate of anabolic pathways and up-regulation of catabolic pathways, the net balance being an increase of the ATP levels. Both ADP and AMP may allosterically activate the enzyme,¹ although physiologically the ATP concentration is about 10-fold higher than ADP and 100-fold higher than AMP, making the AMP/ATP ratio a sensitive indicator of the energy equilibrium inside the cell. Due to its critical function in cell metabolism, AMPK is implicated in numerous metabolic disorders such as type 2 diabetes, cardiovascular diseases, and obesity.² Moreover, several studies have reported seemingly paradoxical results about the role of AMPK in cancer, as it could function as a tumor promoter or suppressor depending on the stage and type of cancer.^{3,4} Overall, AMPK has emerged as a potential therapeutic target in the past decade.⁵

AMPK is a heterotrimeric complex composed of a catalytic α -subunit together with the regulatory β and γ subunits

(Figure 1A).^{6,7} The α -subunit contains the kinase domain in the N-terminal region. The β -subunit encompasses a central carbohydrate-binding module (CBM), which is related to those found in proteins that are capable of binding to glycogen particles (also known as the glycogen-binding domain).⁸ Finally, the γ -subunit comprises four tandem repeats of cystathionine β -synthase domains (CBS), arranged in two Bateman motifs, which are in almost symmetrical “head-to-head” arrangement, and contains four sites able to bind AMP, ADP, and ATP. Among them, site 2 is found to be always empty in crystal structures of AMPK/adenine nucleotide complexes, and site 4 is tightly bound to AMP, whereas sites 1 and 3 can bind nucleotides reversibly in solution depending on their relative concentrations.^{9,10}

Due to its central role in energy homeostasis, the AMPK activity is susceptible to being regulated by several mechanisms, including AMP binding to the CBS domains in the γ -regulatory subunit, a process that promotes phosphorylation of

Received: December 1, 2018

Published: March 29, 2019

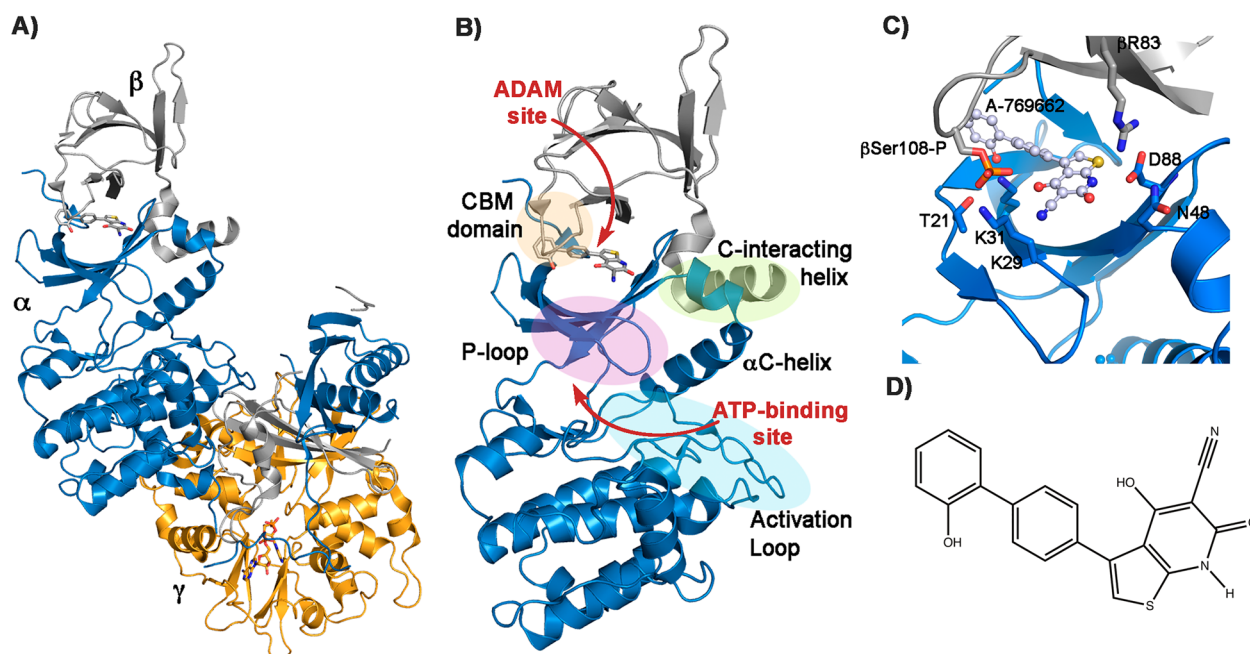


Figure 1. (A) AMPK heterotrimeric complex composed of subunits α_2 (blue), β_1 (gray), and γ_1 (orange) bound to the A-769662 activator (PDB code 4CFF). (B) Representation of the simulation system selected for the study of the direct activation mechanism, which is composed of subunits α_2 and β_1 . The most important regions of these subunits are highlighted; P-loop (purple), activation loop (cyan), CBM domain (orange) and C-interacting helix (green). (C) Binding site of the direct activator A-769662. Relevant residues of the activator binding site are shown as sticks, and the activator as ball and sticks. (D) Chemical structure of A-769662.

Thr172 (pThr172) in the activation loop at the kinase domain of the α -subunit by upstream kinases, such as liver kinase B1 (LKB1) and calcium/calmodulin-dependent protein kinase (CaMKK β), and causes allosteric activation.¹¹ There are also indirect AMPK activators that act by increasing the cellular AMP concentration, such as metformin, phenformin, or oligomycin.¹² Furthermore, in 2006, Abbott Laboratories described a novel mechanism of action that involves the first direct activation of AMPK by the thienopyridone drug A-769662 (Figure 1D).¹³ In contrast to adenine nucleotides, A-769662 does not bind to the CBS motifs in the γ -subunit but to a binding site located at the interface between α and β subunits (Figure 1C),^{14,15} as revealed by the X-ray structure of AMPK bound to the direct activator A-769662 (PDB entry: 4CFF).¹⁶

The activation mechanism of A-769662 is of particular interest, because it activates AMPK independently of α -Thr172 phosphorylation, provided Ser108 in the CBM of the β -subunit is phosphorylated, via a *cis*-autophosphorylation event, so that binding of A-769662 enhances the AMPK activity up to more than 90-fold when pSer108 is present.^{17,18} β_1 -Ser108 is also a substrate for ULK1 as well as other kinases like NimA related kinases (NEKs).¹⁹ Although not directly involved in drug binding, inspection of the X-ray structure of the A-769662-AMPK complex shows that pSer108 forms hydrogen bonds with α -Thr21, α -Lys29, and α -Lys31, which suggests that it contributes to the stability of the binding pocket (see Figure 1C). Nevertheless, phosphorylation of α -Thr172 is not necessary for *in vitro* AMPK activation by direct activators; phosphorylation of this residue is crucial for significant activation of the enzyme *in vivo*.²⁰ Finally, although neither A-769662 nor AMP individually stimulate the activity of dephosphorylated AMPK, together they potentiate the enzymatic activity by >1000-fold, revealing that AMPK

activation can be achieved without the requirement for α -Thr172 phosphorylation.²⁰

These findings reflect the complex interplay of factors that regulate the activity of AMPK, which in turn raises a question about the underlying molecular mechanisms implicated in the enzyme activation. At this point, it has been suggested that A-769662 may induce a conformational change leading to the rearrangement and stabilization of the activation loop.¹⁹ The structural studies have shown that A-769662 facilitates the interaction between the α C helix in the kinase domain and the C-interacting helix of the β -subunit, effectively contributing to the allosteric activation of AMPK.¹⁶ Furthermore, the interaction of the CBM with the kinase domain might be enhanced upon binding of A-769662, which could contribute to protecting pThr172 against dephosphorylation. However, the details of the conformational rearrangements triggered upon binding of A-769662 have not been elucidated yet. In this context, A-769662 deserves interest as a starting point for the design of novel direct activators, which could in turn be useful for gaining insight into AMPK pharmacology.²¹

Here, Molecular Dynamics (MD) simulations have been used to gain insight into the allosteric mechanism that modulates the direct activation of AMPK by small compounds, such as A-769662. Our aim is to take advantage of this knowledge to provide a molecular basis that enables the study of other activators, gaining insight into the different sensitivity of AMPK isoforms, and to improve the design of new drugs with an improved therapeutic profile against AMPK, including natural endogenous metabolites. To attain these aims, a detailed understanding of the activation mechanism is mandatory, making it necessary to identify the molecular factors that govern ligand binding, the relationships between activator binding and structural/dynamical changes in the

protein, and the impact of these changes in the enzyme activity.

In order to address these fundamental questions about the structural and dynamical changes triggered upon binding of A-769662 to AMPK, we have performed extended MD simulations for three systems: the apo form of the enzyme, the species bound to A-769662, and the complex formed by AMPK with both A-769662 in the allosteric drug and metabolite (ADaM) binding pocket and ATP in the ATP-binding site. The choice of these systems was motivated by the interest in unveiling the mechanistic determinants of the allosteric regulation played by A-769662 and, particularly, the changes in the network of interactions between key structural elements implicated in the activation of this cellular energy sensor. This knowledge will contribute to the development of new drugs for the treatment of metabolic disorders, such as type 2 diabetes and obesity, as well as cancer.

METHODS

Molecular Dynamics Simulations. The analysis of the structural and dynamic properties of the simulated systems was performed by means of extended molecular dynamics (MD) simulations. To this end, the simulated systems were built up using the X-ray structure of AMPK bound to the thienopyridone derivative A-769662 (PDB entry 4CFF), which contains the structure of chains α (isoform 2), β (isoform 1), and γ (isoform 1). In addition, it contains pThr172 in the α -subunit and pSer108 in the β -subunit. A-769662 is bound to the activation site located at the interface of the α - and β -subunits. Finally, the X-ray structure of $\alpha 2\beta 1\gamma 1$ -AMPK also contains the kinase inhibitor staurosporine at the ATP-binding site, and two AMP molecules bound to the γ -subunit.

For our purposes here, the γ -subunit was not explicitly considered in the simulated system because it does not directly participate in the direct activation triggered by A-769662, while this facilitates a reduction in the cost of the MD simulations.²² Furthermore, the C-terminal tails of the α - and β -subunits were not explicitly treated in the simulated systems due to the lack of structural information in the X-ray crystallographic data (Figure 1B), and staurosporine was deleted. Therefore, the simulated system comprises 368 amino acid residues (subunit α , 272 residues; subunit β , 96 residues), which were utilized to simulate the wild-type protein (*apo*), the pSer108 enzyme interacting with the A-769662 activator (*holo*), and the pSer108 enzyme bound to both A-769662 and ATP. Hereafter, these systems will be referred to as *apo*, *holo*, and *holo+ATP* species, respectively.

The simulated systems were immersed in a pre-equilibrated octahedral box of TIP3P water molecules.²³ The standard protonation state at physiological pH was assigned to ionizable residues. The final systems contain the model protein, and around 27 500 water molecules, in conjunction with two chloride anions for the *holo* system and four sodium cations for the *holo+ATP* one, leading to simulation systems comprising around 90 000 atoms. Simulations were performed in the NPT ensemble for equilibration and the NVT ensemble for production runs.²⁴ The systems were simulated with periodic boundary conditions and Ewald sums (grid spacing 1 Å) for treating long-range electrostatic interactions.²⁵ All simulations were performed with the Amber ff99SB-ildn²⁶ force field using Amber14.²⁷ The A-769662 ligand was

parametrized with the GAFF force field in conjunction with RESP charges calculated at the B3LYP/6-31G(d) level.²⁸

The initial system was minimized using a multistep protocol. First, the position of all hydrogen atoms in the protein was refined by energy minimization (2000 cycles of steepest descent + 8000 cycles of conjugate gradient). Then, this scheme was used to minimize the position of water molecules and counterions. Finally, the position of all atoms in the system was energy minimized (4000 cycles for steepest descent + 1000 cycles of conjugate gradient). The equilibration process was performed in six steps. The first step involved the heating of the system from 0 to 100 K in 20 ps (NVT ensemble) and was followed by four steps of thermalization to increase the temperature from 100 to 300 K (50 ps/step, NPT conditions). Finally, a last step of 5 ns was run to equilibrate the density of the system at a constant temperature (300 K) and pressure (1 atm). The final structure of the equilibration process was used as a starting point for MD simulations.

Apo, *holo*, and *holo+ATP* systems were simulated in triplicate, each simulation covering a production time of 1 μ s, in order to check the convergence of the dynamical behavior of the systems.

Essential Dynamics. Since the structural flexibility of proteins is crucial to gaining insight into the structure–function relationship,^{29,30} essential dynamics (ED)³¹ was utilized to identify the most important motions from the structural variance sampled in MD simulations.^{32,31} This approach has the advantage that the dynamics along the individual modes can be inspected and visualized separately, thereby allowing the filtering of the main collective motions. To this end, the positional covariance matrix, \tilde{C} , is built (eq 1) and diagonalized to obtain the collective structural deformation modes (eigenvectors), whereas the eigenvalues account for the contribution of each motion to the structural variance of the protein. These properties can then be used to extract information about the frequency and nature of the conformational motions in the system.

$$\tilde{C} = \langle \Delta \mathbf{x} \Delta \mathbf{x}^T \rangle \quad (1)$$

where $\Delta \mathbf{x} = \mathbf{x} - \mathbf{x}_{\text{ref}}$, \mathbf{x} is a vector that collects the position of the atoms chosen to characterize a given state of the system and \mathbf{x}_{ref} denotes the position of the atoms in a reference structure, usually the average structure sampled along the trajectory, $\Delta \mathbf{x}^T$ is the transpose of $\Delta \mathbf{x}$, and the angle brackets stand for the average value over the ensemble of snapshots.

Diagonalization was performed for the set of $C\alpha$ atoms after superposition of the selected snapshots from the trajectory, thus yielding a set of $3N - 6$ (N being the number of $C\alpha$ atoms) eigenvectors and their associated eigenvalues. This choice was motivated by the interest in examining the effect of ligand binding on the major deformation of the protein skeleton. The set eigenvectors are orthogonal and account for the structural variance of the system, the magnitude of the eigenvalue being a measure of the fraction of a given eigenvector to the total variance.

ED analysis was performed for a subset of 25 000 snapshots taken from the last 500 ns of each trajectory, considering only backbone atoms. An average structure was used as a reference template for snapshot superposition. Calculations were performed using the PCAsuite program.³³

As a final remark, let us note that this analysis is expected to be able to capture the major changes in protein dynamics triggered upon binding of the activator and ATP, at least from

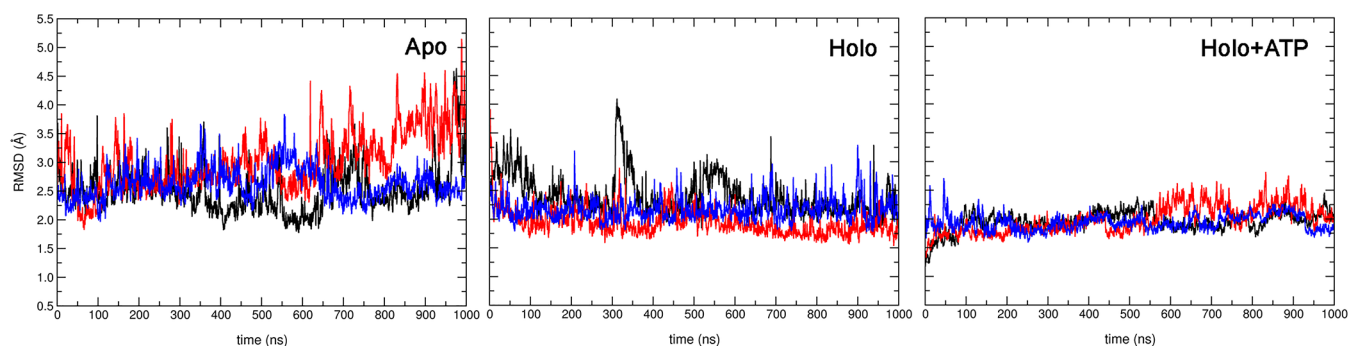


Figure 2. RMSD (Å) along the 1 μ s of simulation for three independent replicas (colored in black, blue, and red) for apo, holo, and holo+ATP using the holo+ATP system (averaged from the last 200 ns of the three independent simulations) as a reference structure and aligned over the C α atoms.

a qualitative point of view. However, comparison of refined techniques that account for nonlinear correlations in the protein motions, such as nonlinear principal component analysis and full correlation analysis based,³⁴ may be better suited for comparative analysis of the conformational changes induced by different activators or in different isoforms.

Conformational Entropy. The conformational entropy was calculated using the quasi-harmonic approximation proposed by Schlitter.³⁵ In this context, the conformational entropy associated with a given eigenvector can be determined as noted in eq 2.

$$S = \frac{1}{2} k_B \sum_i \ln \left(1 + \left(\frac{k_B T e}{\hbar \omega_i} \right)^2 \right) \quad (2)$$

where k_B and \hbar stand for Boltzmann and Planck constants, respectively, T is temperature, and ω_i is the eigenvector i obtained from diagonalization of the mass-weighted covariance matrix.

Since these calculations are very sensitive to simulation time and sampling,³⁶ we have estimated the conformational entropy at infinite time, S^∞ , based on the Harris approach (see eq 3).³⁷ Accordingly, we have calculated the entropy $S(t)$ at regular interval times along the trajectories, and the whole set of entropy values determined was fitted to the following function:

$$S(t) = S^\infty - \left(\frac{a}{t^b} \right) \quad (3)$$

where S^∞ represents the entropy of the system for a simulation of infinite length and a and b are adjustable parameters by curve fitting.

Dynamical Cross-Correlation Analysis. Dynamical cross-correlation matrices (DCCM) have been calculated using the module available in the AMBER package for the whole set of MD simulations.²⁷ To this end, all the snapshots saved along a given trajectory were aligned by means of least-squares fitting of the C α atoms of the whole protein to the equilibrated starting configuration. The cross-correlation matrix is defined as follows:

$$C_{ij} = \frac{c_{ij}}{c_{ii}^{1/2} c_{jj}^{1/2}} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{[\langle r_i^2 \rangle \langle r_j^2 \rangle - \langle r_i \rangle \langle r_j \rangle]^2} \quad (4)$$

where the position vectors of two atoms i and j fitted in the structure at time t are $r_i(t)$ and $r_j(t)$, respectively.

DCCM provides information about the correlation motion of residues along a given trajectory, but not about the magnitude of the motion, thus complementing the ED analysis.³⁸ Cross-correlation coefficients range from -1 to $+1$ values, which are indicative of anticorrelated and correlated motions.

Interaction Energy Network. Analysis of the allosteric network was performed by using the weighted implementation of suboptimal paths (WISP) method.³⁹ This technique permits the carrying out of a dynamic network analysis to understand how the binding of a ligand in an allosteric cavity could affect another binding site. This analysis was performed separately for the structural ensembles sampled for the three simulated systems, taking into account the snapshots collected along the last 500 ns of MD simulations.

RESULTS AND DISCUSSION

Structural Analysis. As noted above, AMPK was simulated in three distinct states: (i) the apo form, (ii) the complex formed by A-769622 with phosphorylated pSer108 AMPK (holo), and (iii) the complex formed by the pSer108 phosphorylated enzyme with both A-769622 and ATP (holo+ATP). The analysis of the trajectories sampled for these species should be valuable in exploring the changes induced by A-769622 on the dynamical properties of the enzyme, keeping in mind that they represent the transition between the basal state of AMPK to the activated species capable of performing the phosphorylation of specific substrates.

The results point out that the subsequent binding of A-769622 and ATP leads to an overall reduction in the conformational flexibility of the enzyme, as indicated by the time evolution of the root-mean square deviation (RMSD) profile determined for the three MD simulations (Figure 2) and the average RMSD values, which are shown in Table 1.

Table 1. RMSD (Å) and Standard Deviation Determined for apo, holo, and holo+ATP States of AMPK Using the holo+ATP System (Averaged from the Last 200 ns of the Three Independent Simulations) As a Reference Structure^a

system	replica 1	replica 2	replica 3
apo	2.5 \pm 0.5	3.0 \pm 0.6	2.6 \pm 0.3
holo	2.3 \pm 0.4	1.9 \pm 0.3	2.2 \pm 0.3
holo + ATP	1.9 \pm 0.2	2.0 \pm 0.3	1.9 \pm 0.2

^aValues determined for the three independent simulations run for apo and ligand-bound states.

Thus, there is a larger reduction of 0.2–1.1 Å in the RMSD of the protein C α atoms upon binding of A-769622, whereas a lower decrease (between 0.1 and 0.4 Å) is observed for the AMPK complex with both A-769662 and ATP. These trends can be primarily attributed to the decrease in the structural flexibility of the N-lobe of the α -subunit and the interacting domain of the β -subunit, which were found to be much more flexible compared to the C-lobe of the α -subunit (see below).

The pattern of residue fluctuations that emerges from the independent simulations run for every protein species is similar (data not shown). Therefore, Figure 3 shows the per-residue

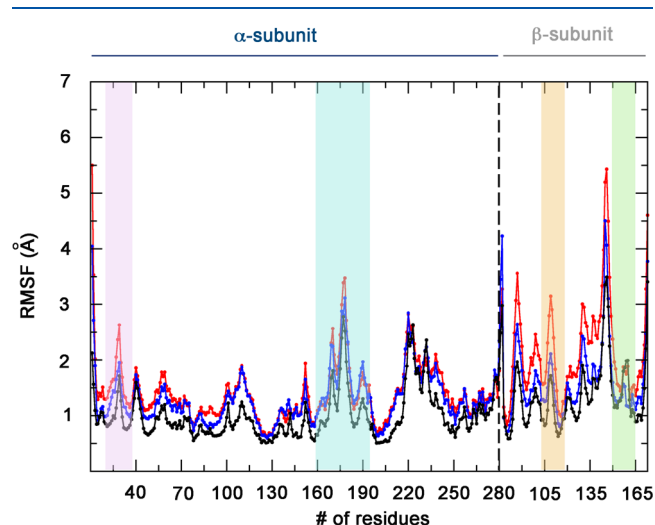


Figure 3. RMSF (Å) for the average of the three independent replicas for each simulated system, apo (red), holo (blue), and holo+ATP (black). The regions highlighted correspond to P-loop (purple), activation loop (cyan), CBM domain (orange), and C-interacting helix (green).

fluctuation averaged from the three replicas run for every simulated (apo, holo, and holo+ATP) system. The region with the highest flexibility in the α -subunit corresponds to the activation loop (highlighted in cyan in Figure 3), and to a lesser extent to the α -helix between residues 210 and 230, located at the bottom of the α -subunit. Comparatively, the magnitudes of the structural fluctuations in the β -subunit, which includes both the CBM domain (containing the phosphorylatable Ser108) and the C-interacting helix (highlighted in orange and green, respectively, in Figure 3), are larger than the RMSF values determined for the α -subunit, which tends to be more rigid.

The largest fluctuations are generally found in the apo state, and the lowest RMSF values are found in the holo+ATP complex, reflecting the increased stiffness arising by filling of the binding pockets occupied by both activator and ATP. Nevertheless, this trend is more relevant in specific regions of the α and β subunits, which are more sensitive to the presence of the ligands. Thus, the magnitude of the fluctuations in the CBM domain is generally larger in the apo species compared to both holo and holo+ATP systems. A similar trait is also observed in the N-terminus region of the β -subunit, which interacts with the N-lobe of the α -subunit. Finally, in this latter subunit, a decrease in the RMSF values is also observed for the region that encloses the P-loop (highlighted in purple in Figure 3) upon binding of the activator, since A-769662 is located at the interface formed by the CBM domain and the P-loop and

by ATP, which fills the cavity located between the P-loop and the helical domain of the α -subunit.

Essential Dynamics (ED). In order to gain deeper insight into the changes in the dynamics of the enzyme promoted upon binding of A-760662 and ATP, ED was used to identify the differences in the major conformational motions of the protein in the different states. As noted in Table 2, the first

Table 2. Contribution of the Essential Motion (%) to the Structural Variance of Different AMPK Systems and the Total Contribution of the First Four Projections

systems	proj. 1	proj. 2	proj. 3	proj. 4	total _(p1–p4)
apo	43.5	10.3	7.9	4.2	65.9
holo	29.0	11.7	7.8	6.3	54.8
holo+ATP	22.5	12.5	6.6	6.2	47.8

essential motion (EM) accounts for almost 44% of the whole structural variance of the protein C α atoms, whereas the contribution of the second EM is reduced to 10%. In the apo form, the main EM is concentrated in two regions: the core of α -helices in the α -subunit, and the β -sheets in the β -subunit. These regions exhibit a bending motion that tends to bring the two domains closer and move them apart in a concerted motion (see Figure 4A). On the other hand, the N-lobe domain of the α -subunit, particularly the P-loop, as well as the α C- and the C-interacting helix from the β -subunit are less flexible, remaining slightly affected by the structural change arising from the first EM.

Comparison of the results obtained for holo and holo+ATP systems also reveals that the first EM has a major contribution to the overall flexibility of the protein, as it accounts for almost 30% and 23% of the structural variance in the complexes with A-769662 and A-769662+ATP, respectively (Table 2), they being 3- and 2-fold larger than the contribution due to the second EM. Furthermore, inspection of Figure 4 shows the existence of a large resemblance between the first EM of apo and holo systems, although it reveals that the extent of the conformational motion observed in the core of α -helices in the α -subunit and the β -sheets in the β -subunit decreases upon binding of the activator. Remarkably, it can be noticed that the first EM in the holo system also includes both the P-loop and the α C- and C-interacting helices, which show a synchronous motion with the structural fluctuations observed in the β -subunit. Since the P-loop and α C- and C-interacting helices contribute to shape the ATP binding-site, these results suggest that the activator could act as “molecular glue” between the β - and the α -subunits, transferring the flexibility from one region to the other (see Figure 4B).

Finally, the holo+ATP system shows a significant reduction in the overall flexibility of the protein, especially regarding the P-loop and the α C- and C-interacting helices, which are much less flexible compared to the holo state. This effect can be attributed to the simultaneous interactions formed by A-769662 and ATP with the residues in the ADaM and ATP-binding sites, which are separated by the P-loop.

The increased stiffness of the protein arising upon binding of the activator to the ADaM site can be viewed as a mechanism to translate the binding energy into changes in the protein dynamics that should facilitate the enzymatic catalysis of AMPK. This concept has been recently reviewed by Richard,⁴⁰ who examined the relationships between ligand-driven conformational changes and the shaping of stiff active sites

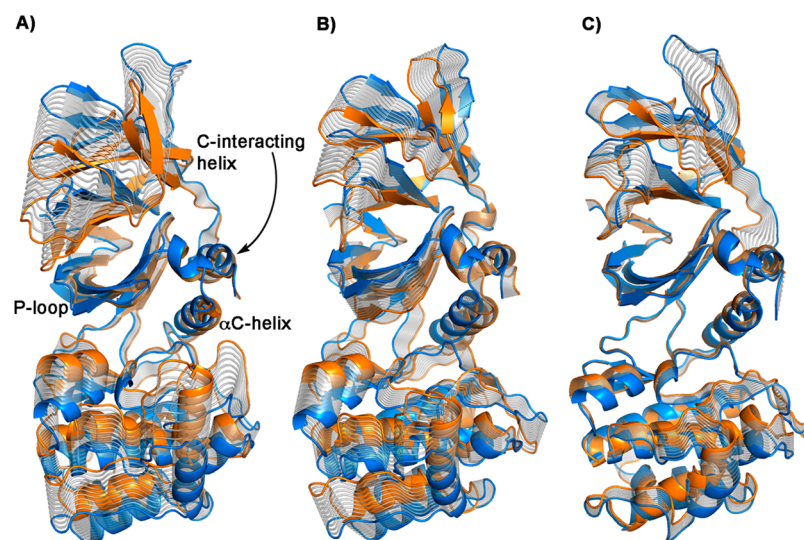


Figure 4. Essential dynamics analysis of the 1 μ s MD simulations run for (A) apo, (B) holo, and (C) holo+ATP forms. Only the first essential motion of the C α atoms is shown.

as a regulation mechanism of the catalytic efficiency of enzymes. In particular, it is stated that the “existence of many enzymes in flexible, entropically rich, and inactive ground states provides a mechanism for utilization of ligand-binding energy to mold these catalysts into stiff and active forms.” In this context, present results allow us to propose that the significant effect played by the binding of activators to the ADaM site on the AMPK conformational flexibility should also contribute to the proper preorganization of the catalytic site.

Conformational Entropy. To estimate the extent of the dynamical differences observed between apo, holo, and holo+ATP systems, the conformational entropy of the protein was evaluated considering the backbone and side chains of residues 5–360, in order to avoid the noise of the N- and C-terminal residues. As expected, the results (Table 3) confirm that holo

Table 3. Entropy Values (kcal K⁻¹ mol⁻¹) for the apo, holo, and holo+ATP Systems of AMPK Determined Considering Heavy Atoms in Both Backbone and Side Chain of Residues 5–360^a

system	replica			mean (SD)
	1	2	3	
S_{∞} apo	46.4	44.1	46.8	45.8 (1.4)
S_{∞} holo	33.8	33.7	39.2	35.5 (3.1)
S_{∞} holo+ATP	32.8	34.5	34.1	33.8 (0.9)
ΔS (holo–apo)	–12.6	–10.4	–7.6	–10.2
ΔS (holo+atp–apo)	–13.6	–9.6	–12.7	–12.8
ΔS (holo+atp – holo)	–1.0	0.8	–5.1	–1.8

^aValues determined for the three independent simulations run for apo and ligand-bound states using eq 1 (see above).³⁷

systems are less flexible than apo ones, as noted in a reduction of the conformational entropy of 10.3 kcal K⁻¹ mol⁻¹. Compared to the holo system, binding of ATP gives rise to an additional decrease (by 1.7 kcal K⁻¹ mol⁻¹) in the conformational entropy of the protein. These results indicate that binding of the activator has a sizable influence on the overall conformational flexibility of the protein domains, leading to a significant increase in the stiffness of the enzyme.

Dynamic Cross-Correlation Analysis (DCC). With the aim to gain deeper insight into the alterations triggered upon binding of A-769662 and ATP on the dynamics of key structural elements, a DCC analysis was performed for the set of trajectories run for apo, holo, and holo+ATP species. It is worth noting that the extreme values indicate the regions that fluctuate in a correlated (+1) or anticorrelated (–1) sense, showing which regions of the protein move in the same or opposite direction, respectively, whereas values close to zero indicate the absence of correlated motions. Figure 5 shows the results obtained for representative trajectories of the apo, holo, and holo+ATP systems (note that similar trends were observed for the other two replicas run for every simulated system; see the Supporting Information).

Inspection of the DCC matrix for the apo system (Figure 5A) reveals the correlation between several regions implicated in the enzyme activation. Thus, there is a high correlation between the α C-helix and C-interacting helix, which are dynamically shifted in a correlated way along the MD simulations, and to a lesser extent with the P-loop. In contrast, the P-loop is highly correlated with the α C-helix. On the other hand, there is a lack of correlation in the fluctuations of the CBM domain and the structural elements mentioned above (i.e., C-interacting helix, α C-helix, and P-loop). This is reflected in the trends derived from the ED analysis (see Figure 4A), as the negligible motion observed for both the P-loop and the C-interacting and α C-helices are in contrast with the large structural deformation found for the CBM domain. Finally, it is also worth noting the lack of significant correlation between the residues in the A-loop and the other structural elements implicated in the enzyme activation.

Figure 5B (holo) and C (holo+ATP) permit the examination of the impact of ligand (A-769662, ATP) binding on the dynamical behavior of the enzyme. In particular, these plots reveal that although the gross features of the DCC matrix obtained for the apo species can be recognized in the matrices obtained for holo and holo+ATP systems, there is a significant reduction in the extent of the correlated motions. This is in agreement with the increased stiffness of the protein backbone upon binding of both activator and ATP (see above). However, fine inspection of Figure 5B and C also points out

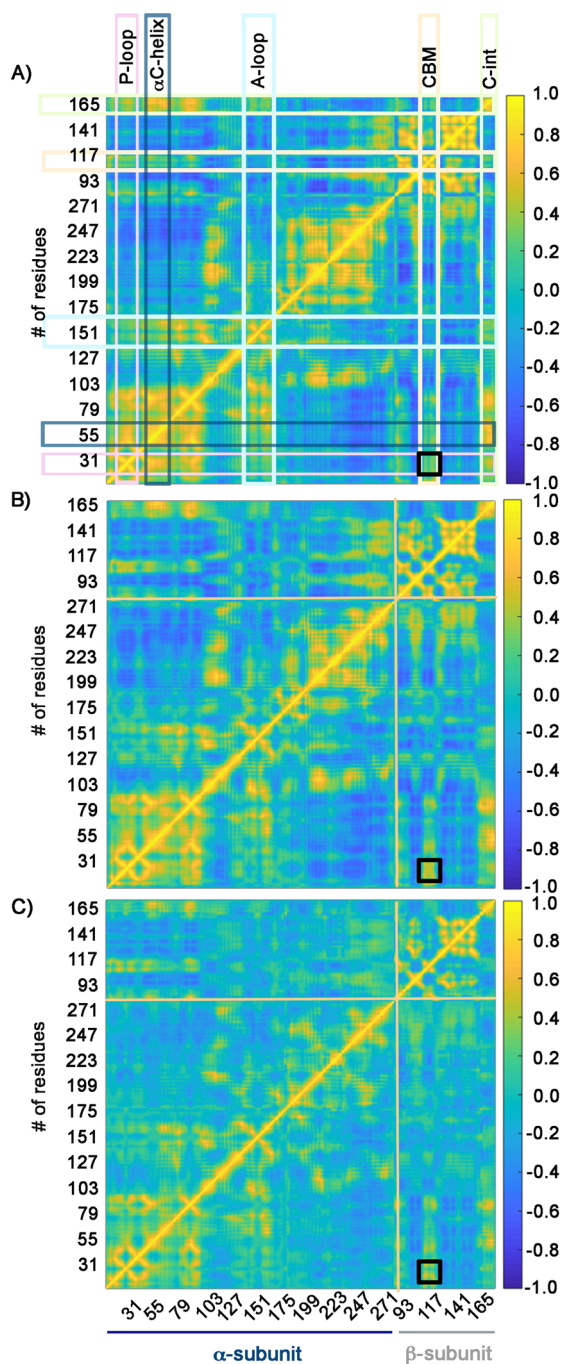


Figure 5. Dynamic cross-correlation (DCC) matrices for (A) apo, (B) holo, and (C) holo+ATP systems considering 1 μ s of simulation. Regions colored in yellow show high correlated fluctuations (+1), while regions colored in blue show high-anticorrelated fluctuations (-1). The structural subunits and the most important regions of AMPK systems are indicated in squares for P-loop (pink), α C-helix (blue), A-loop (cyan), CBM domain (orange), and C-interacting helix (green).

that binding of A-769662 and ATP to AMPK reinforces the correlation between the motion of the CBM domain and the P-loop (highlighted as black square in Figure 5), which suggests that binding of the activator may have a direct influence on the preorganization of the ATP binding site. In particular, it can be hypothesized that the conformational rearrangement of the P-loop may be relevant for enabling the

adoption of a proper conformation well suited for recognition and binding of ATP.

To check this hypothesis, we have determined the structural resemblance of the ATP-binding site along the trajectories obtained for apo, holo, and holo+ATP systems with regard to the conformation corresponding to the ATP-bound form (see Figure 6). For the holo-ATP species, the residues that shape

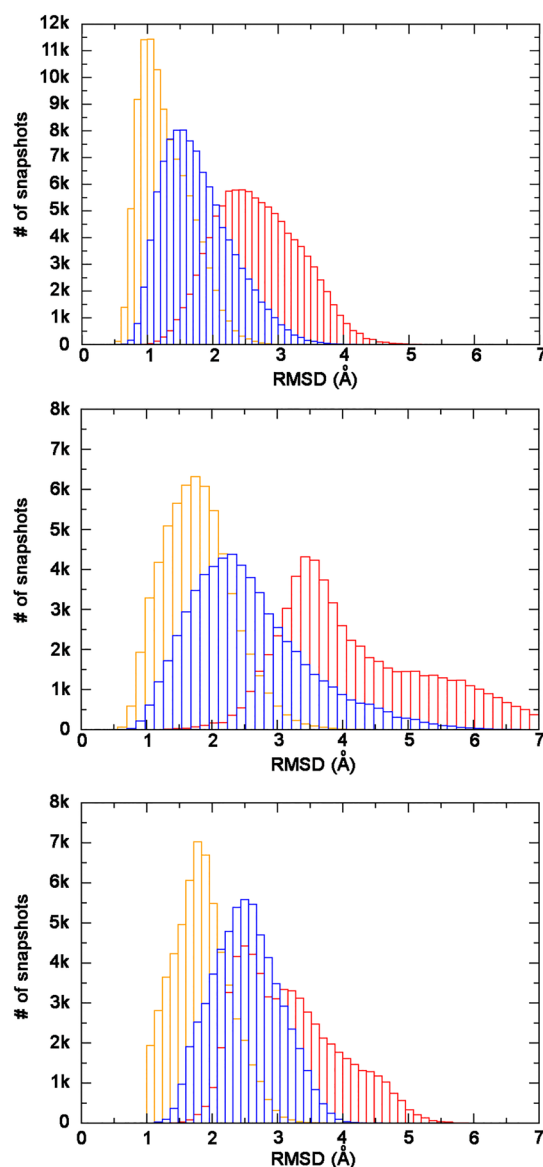


Figure 6. Distribution of the positional deviation (RMSD; Å) of the structures sampled along the trajectories run for apo (red), holo (blue), and holo+ATP (orange) for structural elements implicated in enzyme activation: (A) ATP-binding site, (B) α C- and C-interacting helices, and (C) activation loop. All the systems have been previously aligned over the reference structure, i.e., the average structure of the last 200 ns of the three replicas of the holo+ATP system.

the ATP-binding site, which includes the P-loop, sample a reduced conformational space, with a peak in the narrow population distribution centered at a positional root-mean square deviation of 1 Å (Figure 6A). Not surprisingly, a broader distribution is found for the structures sampled in the apo state, the peak in the distribution profile being located at 2.5 Å. In contrast, analysis of the structures sampled in the

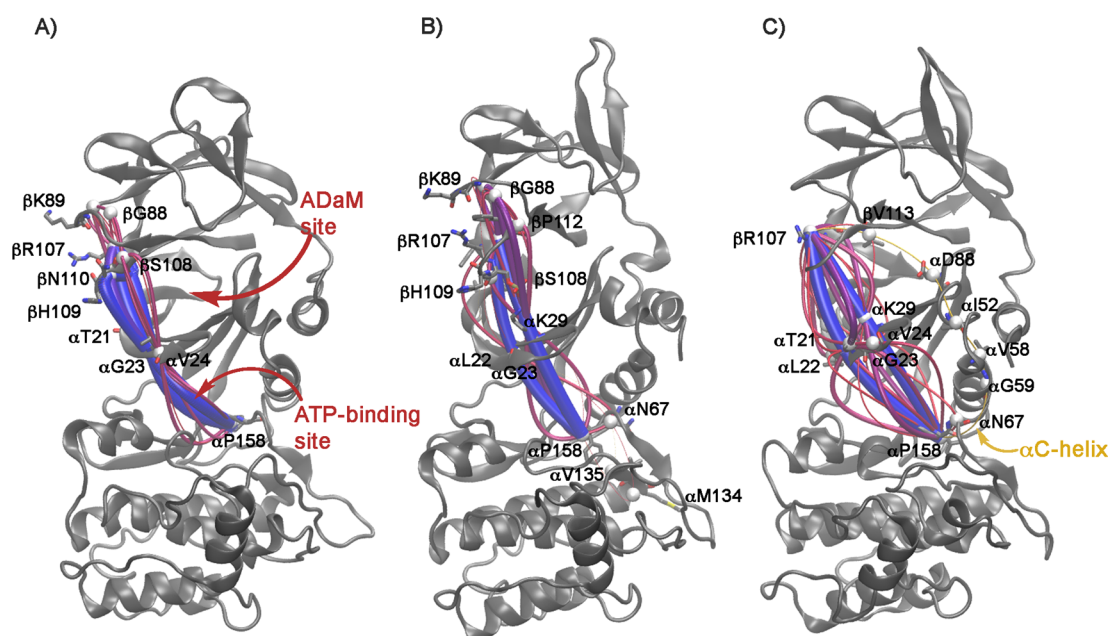


Figure 7. Representation of the optimal pathways indicating the residue interaction network between the ADaM and ATP-binding sites for the simulations of (A) apo, (B) holo, and (C) holo+ATP. The wide blue lines indicate the shortest pathway, and thinner red and orange lines display the longest pathways.

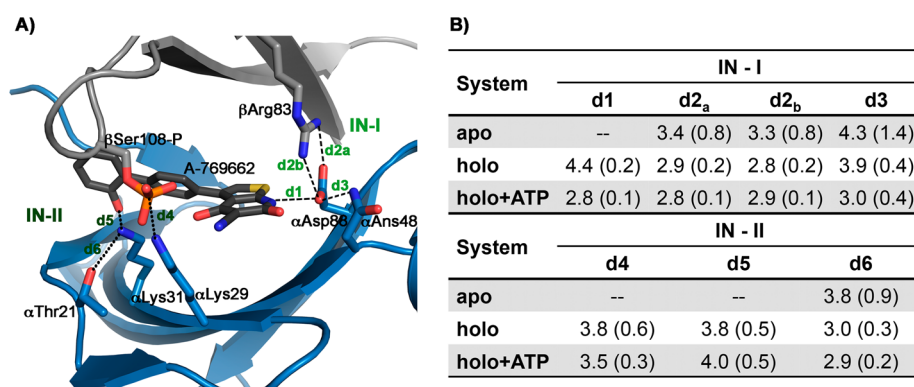


Figure 8. (A) Representation of key hydrogen-bond interactions in the ADaM site (interaction networks IN-I and IN-II). The α and β subunits are displayed in blue and gray, respectively. (B) Mean (standard deviation) values (Å) for selected distances corresponding to interactions in IN-I and IN-II networks for apo, holo, and holo+ATP systems.

holo trajectory reveals that the population distribution is shifted to the left, approaching the distribution obtained for the holo+ATP system, as noted in the presence of the peak of the population distribution at 1.6 Å and in the narrowing of the distribution profile with regard to the apo species.

A similar but less marked effect is observed in the spatial arrangement of the α C- and C-interacting helices (Figure 6B). Compared to the ATP-binding site, the distribution of structures reveals a wider range of structural fluctuation, with a peak centered at a positional deviation of 1.7 Å. In contrast, the distribution profile of the apo structure reveals a much larger structural fluctuation, which is however notably reduced upon binding of A-769662 to the ADaM site (note that peak located at 3.5 Å for the apo form is shifted to a value of 2.3 Å in the holo species).

Finally, it is worth noting that the preorganization effect promoted upon binding of the activator is less apparent in the activation loop (Figure 6C). Although the distribution profile determined for the holo species is narrower than the

distribution obtained for the apo form, there is a significant overlap, and indeed the position of the peaks (centered at around 2.5 Å) is rather similar for both apo and holo species.

Overall, these findings reinforce the concept that the activator is acting as a molecular glue, which connects the intrinsic fluctuations of the CBM domain in the β -subunit with the conformational flexibility of the P-loop, increasing the tendency to preorganize key structural elements, particularly the P-loop, toward conformations well suited to bind ATP.

Interaction Network Analysis. How is the preorganization of the P-loop achieved through binding of A-769662? To answer this question, we performed a network analysis by means of the weighted implementation of suboptimal paths (WISP) tool,³⁹ which permit the identification of protein residues that may cooperatively interact to facilitate the structural arrangement of the ATP-binding site.

The results obtained for apo, holo, and holo+ATP (Figure 7) disclose networks of residues that directly connect the β - and α -subunits through the ADaM site (i.e., the binding site of

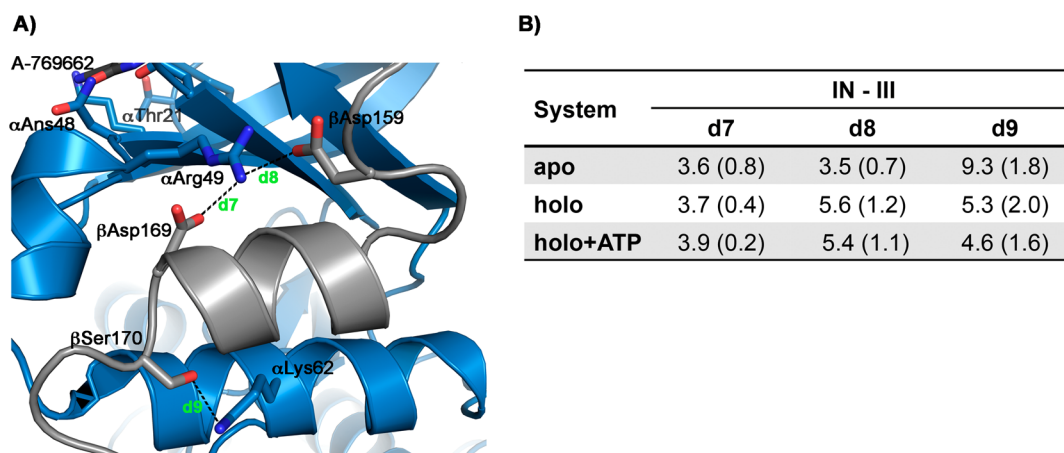


Figure 9. (A) Representation of key hydrogen-bond interactions related to the α C and C-interacting helices from the α and β subunits, respectively. (B) Mean (standard deviation) values (Å) for selected distances corresponding to interactions for apo, holo, and holo+ATP systems.

the activator). Comparison of Figure 7A (apo) and B (holo) reveals that the number of residues from the α -subunit involved in these networks is notably increased in the presence of the activator, in conjunction with the number of paths that connect ADaM and ATP-binding sites (red thin lines in Figure 7B). However, the most interesting finding is found in the holo+ATP system, as this analysis permits the identification of an additional network (orange thin line in Figure 7C) that involves the connection between β - and α -subunits through the α C-helix. This can be viewed as the result of the better preorganization of the P-loop and the α C- and C-interacting helices synergistically induced by the presence of A-769662 and ATP.

The analysis of the interaction networks also permits the identification of key residues participating in the allosteric regulation of AMPK by A-769662 (Figures 8 and 9). Most of the contacts comprise van der Waals interactions between the activator and hydrophobic residues, but the binding is also assisted by the formation of hydrogen bonds. In particular, α Asp88 is engaged in hydrogen bond interactions with α Asn48 and A-769662, as noted in average distance of 2.8 and 3.0 Å (distances d1 and d3 in Figure 8) in the holo+ATP system. Furthermore, the arrangement of α Asp88 is also assisted by a salt bridge formed with β Arg83, as noted in the average values of the N \cdots O distances, 2.8 (d_{2a}) and 2.9 (d_{2b}) Å, in the holo+ATP system, which can be expected to exert a cooperative reinforcement of the hydrogen-bond interactions described above. In this regard, the analysis of the distances for these interactions (Figure 8B) reveals a progressive shortening from apo to holo to holo+ATP, suggesting that the interaction between α Asp88 and β Arg83 (d_{2a} and d_{2b} in Figure 8A) is crucial to maintaining the position of α Asp88, which is the direct connector with the activator. At this point, it is worth noting that the relevant contribution of β Arg83 was demonstrated by Xiao and co-workers¹⁶ in several biochemical assays performed for the β R83A variant of the enzyme, a mutation that was shown to abolish the enzyme activation. In particular, these studies also demonstrated that the β R83A mutation increased the equilibrium dissociation constant (K_d ; μ M) for the binding of A-769662 from 0.51 for the wild type α 1 β 1 γ 1 isoform to 8.6 for the R83A mutated species (see Table 2 in ref 16).

Another relevant interaction network (IN-II; Figure 8) is formed by the interactions between residues α Thr21, α Lys29

and α Lys31 and β pSer108. In particular, the stability of the A-769662-bound complex is assisted by the electrostatic contact formed between the phosphate group of β pSer108 and α Lys29 in the holo species, and further strengthened in the holo+ATP system, as noted in the reduction of the distance between the interacting residues and the smaller standard deviation in the holo+ATP complex. Thus, the interaction between β pSer108 and α Lys29 can be viewed as an effective mechanical constraint that traps A-769662 in the ADaM site, reinforcing the “gluing” effect between the CBM domain and the P-loop. Worthwhile, this interaction has been experimentally demonstrated to be crucial for the activation mechanism of AMPK, since the mutations α 2(K29A/K31A) β 1 γ 1 and α 2 β 1(S108A)- γ 1 lead to a marked reduction in the activity of AMPK, which shows no sensitivity to activation by A-769662, without altering the enzyme regulation by AMP. Indeed, these site-directed mutagenesis assays also demonstrated that mutations significantly reduced the dissociation constant of A-769662, a trait also found for the binding of activator 991, which also binds at the activator binding site.¹⁶ Overall, these interactions appear to be crucial for the formation of the allosteric network implicated in the AMPK activation by A-769662.

Finally, a detailed analysis of the residue interaction network also reveals that the preorganization of the CBM and P-loop domains triggered upon activator binding is also accompanied by a conformational arrangement of the α C and C-interacting helices. In fact, the results reveal the identification of a network involving residues in these two structural elements, which are required to adopt a proper arrangement to sustain the optimal activity of AMPK, in the holo+ATP system, while such a network is not apparent in both apo and holo systems (Figure 7). Figure 9 shows a representative snapshot of the interaction between selected residues of the C-interacting helix from the β -subunit and α C of the α -subunit. The interaction between α Arg49 and β Asp169 (d7 in Figure 9) is maintained stable for all the systems; however, the interaction between α Arg49 and β Asp159 (d8 in Figure 9) at the beginning of the C-interacting helix is weakened upon transition from apo to holo systems as a result of the interaction of the ligand in the ADaM site. This trait is maintained in the holo+ATP system. Conversely, the interaction between α Lys62 and β Ser170, which directly connects the α C and C-interacting helices, decreases from apo to holo+ATP systems, showing the correlated movement that exists due to the preorganization of the ATP-binding site. Let

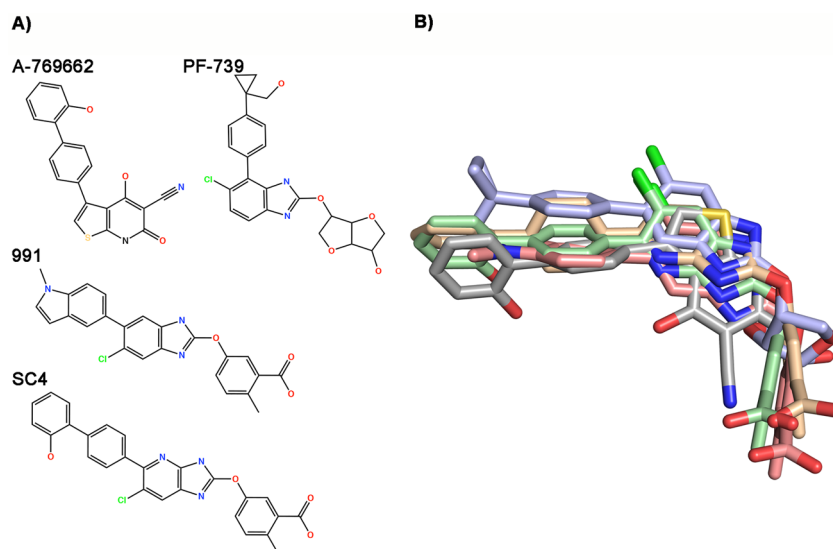


Figure 10. (A) 2D representation of direct AMPK activators. (B) Superposition of the molecular structure of activators in the X-ray structures for A-769662 (gray, PDB: 4CFE),¹⁶ 991 (pink, PDB: 4CFE),¹⁶ PF-379 (blue, PDB: SUFU),⁴¹ and SC4 (green, PDB: 6B1U and yellow, PDB: 6B2E).⁴²

us note that these interactions, and the concomitant spatial arrangement of the α C and C-interacting helices, support the findings observed in the ED (Figure 4) and DCCM (Figure 5) analysis shown above.

The preceding results provide a molecular basis to relate the specific structure of the ADaM site with the intermolecular interactions that mediate the binding of A-769662 and the activation of AMPK. It then may be questioned whether this molecular mechanism may underlie the activity of other direct activators, such as compounds PF-739, 991, and SC4, whose chemical structures are shown in Figure 10A. Inspection of the X-ray structures available for their AMPK complexes reveals that there is a nice correspondence in the molecular overlays of these compounds (Figure 10B), which can fill the ADaM site mimicking the arrangement of A-769662. It is noteworthy that the hydrogen-bond interaction between the activator and α Asp88 is well preserved in all cases, which reinforces the key role of this interaction for the efficient enzyme activation. Similarly, the interaction between β pSer108 and α Lys29 can *a priori* be stably formed in all complexes. However, this raises the challenging question about the distinct sensitivity of $\alpha\beta\gamma$ isoforms to the activation triggered by these compounds.¹⁶ For instance, it is unclear whether the replacement of β Asn111 in $\alpha 2\beta 1\gamma 1$ by Asp in $\alpha 2\beta 2\gamma 1$ may alter the trade-off between electrostatic and hydrophobic interactions formed by the activator in the ADaM site, and hence the efficiency of the activation mechanism, or whether these differences may be attributed to more subtle differences that affect the dynamical response of the protein upon ligand binding and the preorganization of the structural elements directly implicated in the catalytic reaction. Further studies should be valuable for shedding light into the origin of the isoform-dependent catalytic activation by these compounds.

CONCLUSIONS

In this work, we have studied the activation mechanism of AMPK by the direct activator A-769662, which binds to the ADaM site. To the best of our knowledge, this specific binding cavity is not found in other kinases, which makes the

therapeutic value of AMPK to highly attractive as a way to modulate the enzyme activity, especially in the context of pathological dysfunctions. The presence of the activator affects the protein flexibility, particularly regarding the motion of the N-terminus of the catalytic α -subunit and the regulatory domain of the β -subunit. The net effect is that the shape and the size of the ATP-binding pocket is allosterically regulated by the binding of A-769662 and tends to adopt a topology well suited for ATP binding. Thus, it can be hypothesized that the activator A-769662 might act as a glue, filling the space between the β -subunit and the N-terminal domain of the α -subunit, making an effective connection between β - and α -subunits that preorganizes the ATP-binding site, favoring the binding of ATP, and explaining the increase of the AMPK activity.

The structural analysis has disclosed key residues required for the formation of the allosteric network that connects the ADaM and ATP-binding sites through interactions with the A-769662 ligand. In particular, the interactions between α Asp88 and β Arg83 and between β pSer108 and α Lys29 are found to be critical for the enzyme activation through the analysis of the interaction networks. In turn, these findings pave the way to explore the structural features that underlie the different sensitivity of AMPK isoforms to A-769662, as the occurrence of specific residues in different isoforms may affect the efficiency of the allosteric signaling mediated by direct activators. As an example, these studies could be valuable to discern why A-769662 is only active in the $\alpha 2\beta 1\gamma 1$ isoform, while compound 991 is also active with AMPK complexes containing the isoform $\beta 2$.¹⁶

Finally, the understanding of the direct activation mechanism of AMPK described in this work opens new opportunities not only for the rational development of small compounds that might modulate the activity of specific isoforms of this cellular energy sensor in different tissues but also for selecting guidelines that enable the screenings of endogenous metabolites that might regulate the enzyme activity under physiological conditions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.8b00890.

Dynamic cross-correlation analysis obtained for replicas of apo, holo, and holo+ATP systems of AMPK (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: fjlunque@ub.edu. Tel.: +34 934033788.

*E-mail: cestarellas@ub.edu. Tel.: +34 934033793.

ORCID

Jeronimo Lameira: 0000-0001-7270-1517

F. Javier Luque: 0000-0002-8049-3567

Carolina Estarellas: 0000-0002-0944-9053

Author Contributions

These authors have equally contributed to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Spanish Ministerio de Economía y Competitividad (SAF2017-88107-R, and Maria de Maetzu MDM-2017-0767), the Generalitat de Catalunya (2017SGR1746), and Fundación Eugenio Rodríguez Pascual for financial support and the Barcelona Supercomputing Center (BSC-CNS: BCV-2014-3-0011 and BCV-2019-1-0009) and the Consorci de Serveis Universitaris de Catalunya (CSUC) for providing access to computational resources. E.A. thanks AGAUR (Generalitat of Catalunya; 2018FI_B1_00001) and G.M. thanks the National Council for the Improvement of Higher Education (CAPES) for their fellowships.

■ REFERENCES

- (1) Xiao, B.; Sanders, M. J.; Underwood, E.; Heath, R.; Mayer, F. V.; Carmena, D.; Jing, C.; Walker, P. A.; Eccleston, J. F.; Haire, L. F.; Saiu, P.; Howell, S. A.; Aasland, R.; Martin, S. R.; Carling, D.; Gamblin, S. J. Structure of Mammalian AMPK and its Regulation by ADP. *Nature* **2011**, *472*, 230–233.
- (2) Carling, D. AMPK Signalling in Health and Disease. *Curr. Opin. Cell Biol.* **2017**, *45*, 31–37.
- (3) Hardie, D. G. AMP-Activated Protein Kinase: a Cellular Energy Sensor with a Key Role in Metabolic Disorders and in Cancer. *Biochem. Soc. Trans.* **2011**, *39*, 1–13.
- (4) Penfold, L.; Woods, A.; Muckett, P.; Nikitin, A. Y.; Kent, T. R.; Zhang, S.; Graham, R.; Pollard, A.; Carling, D. CAMKK2 Promotes Prostate Cancer Independently of AMPK via Increased Lipogenesis. *Cancer Res.* **2018**, *78*, 6747–6761.
- (5) Hardie, D. G. AMP-Activated Protein Kinase as a Drug Target. *Annu. Rev. Pharmacol. Toxicol.* **2007**, *47*, 185–210.
- (6) Li, X.; Wang, L.; Zhou, X. E.; Ke, J.; de Waal, P. W.; Gu, X.; Tan, M. H. E.; Wang, D.; Wu, D.; Xu, H. E.; Melcher, K. Structural Basis of AMPK Regulation by Adenine Nucleotides and Glycogen. *Cell Res.* **2015**, *25*, 50–66.
- (7) Hardie, D. G. Organismal Carbohydrate and Lipid Homeostasis. *Cold Spring Harbor Perspect. Biol.* **2012**, *4*, a006031.
- (8) Andersson, U.; Filipsson, K.; Abbott, C. R.; Woods, A.; Smith, K.; Bloom, S. R.; Carling, D.; Small, C. J. AMPK-Activated Protein Kinase Plays a Role in the Control of Food Intake. *J. Biol. Chem.* **2004**, *279*, 12005–12008.
- (9) Carling, D.; Thornton, C.; Woods, A.; Sanders, M. J. AMP-Activated Protein Kinase: New Regulation, New Roles? *Biochem. J.* **2012**, *445*, 11–27.
- (10) Chen, L.; Wang, J.; Zhang, Y. Y.; Yan, S. F.; Neumann, D.; Schlattner, U.; Wang, Z. X.; Wu, J. W. AMP-Activated Protein Kinase Undergoes Nucleotide-Dependent Conformational Changes. *Nat. Struct. Mol. Biol.* **2012**, *19*, 716–718.
- (11) Carling, D.; Mayer, F. V.; Sanders, M. J.; Gamblin, S. J. AMP-Activated Protein Kinase: Nature's Energy Sensor. *Nat. Chem. Biol.* **2011**, *7*, 512–518.
- (12) Calabrese, M. F.; Rajamohan, F.; Harris, M. S.; Caspers, N. L.; Magyar, R.; Withka, J. M.; Wang, H.; Borzilleri, K. A.; Sahasrabudhe, P. V.; Hoth, L. R.; Geoghegan, K. F.; Han, S.; Brown, J.; Subashi, T. A.; Reyes, A. R.; Frisbie, R. K.; Ward, J.; Miller, R. A.; Landro, J. A.; Londregan, A. T.; Carpino, P. A.; Cabral, S.; Smith, A. C.; Conn, E. L.; Cameron, K. O.; Qiu, X.; Kurumbail, R. G. Structural Basis for AMPK Activation: Natural and Synthetic Ligands Regulate Kinase Activity from Opposite Poles by Different Molecular Mechanism. *Structure* **2014**, *22*, 1161–1172.
- (13) Cool, B.; Zinker, B.; Chiou, W.; Kifle, L.; Cao, N.; Perham, M.; Dickinson, R.; Adler, A.; Gagne, G.; Iyengar, R.; Zhao, G.; Marsh, K.; Kym, P.; Jung, P.; Camp, H. S.; Frevert, E. Identification and Characterization of a Small Molecule AMPK Activator that Treats Key Components of Type 2 Diabetes and the Metabolic Syndrome. *Cell Metab.* **2006**, *3*, 403–416.
- (14) Li, J.; Li, S.; Wang, F.; Xin, F. Structural and Biochemical Insights into the Allosteric Activation Mechanism of AMPK. *Chem. Biol. Drug Des.* **2017**, *89*, 663–669.
- (15) Cameron, K. O.; Kurumbail, R. G. Recent Progress in the Identification of Adenosine Monophosphate-Activated Protein Kinase (AMPK) Activators. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 5139–5148.
- (16) Xiao, B.; Sanders, M. J.; Carmena, D.; Bright, N. J.; Haire, L. F.; Underwood, E.; Patel, B. R.; Heath, R. B.; Walker, P. A.; Hallen, S.; Giordanetto, F.; Martin, S. R.; Carling, D.; Gamblin, S. J. Structural Basis of AMPK Regulation by Small Molecule Activators. *Nat. Commun.* **2013**, *4*, 3017.
- (17) Sanders, M. J.; Ali, Z. S.; Hegarty, B. D.; Heath, R.; Snowden, M. A.; Carling, D. Defining the Mechanism of Activation of AMP-Activated Protein Kinase by the Small Molecule A-769662, a Member of the Thienopyridone Family. *J. Biol. Chem.* **2007**, *282*, 32539–32548.
- (18) Scott, J. W.; van Denderen, B. J.; Jorgensen, S. B.; Honeyman, J. E.; Steinberg, G. R.; Oakhill, J. S.; Iseli, T. J.; Koay, A.; Gooley, P. R.; Stapleton, D.; Kemp, B. E. Thienopyridone Drugs Are Selective Activators of AMP-Activated Protein Kinase Beta1-containing Complexes. *Chem. Biol.* **2008**, *15*, 1220–1230.
- (19) Dite, T. A.; Ling, N. X. Y.; Scott, J. W.; Hoque, A.; Galic, S.; Parker, B. L.; Ngoei, K. R. W.; Langendorf, C. G.; O'Brien, M. T.; Kundu, M.; Viollet, B.; Steinberg, G. R.; Sakamoto, K.; Kemp, B. E.; Oakhill, J. S. The Autophagy Initiator ULK1 Sensitizes AMPK to Allosteric Drugs. *Nat. Commun.* **2017**, *8*, 571.
- (20) Willows, R.; Sanders, M. J.; Xiao, B.; Patel, B. R.; Martin, S. R.; Read, J.; Wilson, J. R.; Hubbard, J.; Gamblin, S. J.; Carling, D. Phosphorylation of AMPK by Upstream Kinases Is Required for Activity in Mammalian Cells. *Biochem. J.* **2017**, *474*, 3059–3073.
- (21) Scott, J. W.; Ling, N.; Issa, S. M.; Dite, T. A.; O'Brien, M. T.; Chen, Z. P.; Galic, S.; Langendorf, C. G.; Steinberg, G. R.; Kemp, B. E.; Oakhill, J. S. Small Molecule Drug A-769662 and AMP Synergistically Activate Naive AMPK Independent of Upstream Kinase Signalling. *Chem. Biol.* **2014**, *21*, 619–627.
- (22) Sanbonmatsu, K. Y.; Tung, C. S. High Performance Computing in Biology: Multimillion Atom Simulations of Nanoscale Systems. *J. Struct. Biol.* **2007**, *157*, 470–480.
- (23) 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.* **1983**, *79*, 926–935.
- (24) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (25) Leach, A. R. *Molecular Modelling: Principles and Applications*, 2nd ed.; Pearson Education Limited, 2001.

(26) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950–1958.

(27) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R. *AMBER 14*; University of California, San Francisco, 2014.

(28) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: the RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(29) Daniel, R. M.; Dunn, R. V.; Finney, J. L.; Smith, J. C. The Role of Dynamics in Enzyme Activity. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 69–92.

(30) Yang, L. W.; Bahar, I. Coupling Between Catalytic Site and Collective Dynamics: A Requirement for Mechanochemical Activity of Enzymes. *Structure* **2005**, *13*, 893–904.

(31) Amadei, A.; Linssen, A. B.; Berendsen, H. J. Essential Dynamics of Proteins. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 412–425.

(32) Garcia, A. E. Large-Amplitude Nonlinear Motions in Proteins. *Phys. Rev. Lett.* **1992**, *68*, 2696–2699.

(33) PCAsuite. <http://mmb.pcb.ub.es/software/pcasuite/pcasuite.html> (accessed June 20, 2018).

(34) Daidone, I.; Amadei, A. Essential Dynamics: Foundations and Applications. *WIREs Comput. Mol. Sci.* **2012**, *2*, 762–770.

(35) Schlitter, J. Estimation of Absolute and Relative Entropies of Macromolecules Using the Covariance Matrix. *Chem. Phys. Lett.* **1993**, *215*, 617–621.

(36) Baron, R.; Hünenberger, P. H.; McCammon, J. A. Absolute Single-Molecule Entropies from Quasi-Harmonic Analysis of Microsecond Molecular Dynamics: Correction Terms and Convergence Properties. *J. Chem. Theory Comput.* **2009**, *5*, 3150–3160.

(37) Harris, S. A.; Gavathiotis, E.; Searle, M. S.; Orozco, M.; Laughton, C. A. Cooperativity in Drug-DNA Recognition: a Molecular Dynamics Study. *J. Am. Chem. Soc.* **2001**, *123*, 12658–63.

(38) Hunenberger, P.; Mark, A.; van Gunsteren, W. Fluctuation and Cross-Correlation Analysis of Protein Motions Observed in Nano-second Molecular Dynamics Simulations. *J. Mol. Biol.* **1995**, *252*, 492–503.

(39) Van Wart, A. T.; Durrant, J.; Votapka, L.; Amaro, R. E. Weighted Implementation of Suboptimal Paths (WISP): An Optimized Algorithm and Tool for Dynamical Network Analysis. *J. Chem. Theory Comput.* **2014**, *10*, 511–517.

(40) Richard, J. P. Protein Flexibility and Stiffness Enable Efficient Enzymatic Catalysis. *J. Am. Chem. Soc.* **2019**, *141*, 3320–3331.

(41) Cokorinos, E. C.; Delmore, J.; Reyes, A. R.; Albuquerque, B.; Kjobsted, R.; Jorgensen, N. O.; Tran, J. L.; Jatkar, A.; Cialdea, K.; Esquejo, R. M.; Meissen, J.; Calabrese, M. F.; Cordes, J.; Moccia, R.; Tess, D.; Salatto, C. T.; Coskran, T. M.; Opsahl, A. C.; Flynn, D.; Blatnik, M.; Li, W.; Kindt, E.; Foretz, M.; Viollet, B.; Ward, J.; Kurumbail, R. G.; Kalgutkar, A. S.; Wojtaszewski, J. F. P.; Cameron, K. O.; Miller, R. A. Activation of Skeletal Muscle AMPK Promotes Glucose Disposal and Glucose Lowering in Non-human Primates and Mice. *Cell Metab.* **2017**, *25*, 1147–1159.

(42) Ngoei, K. R. W.; Langendorf, C. G.; Ling, N. X. Y.; Hoque, A.; Varghese, S.; Camerino, M. A.; Walker, S. R.; Bozikis, Y. E.; Dite, T. A.; Ovens, A. J.; Smiles, W. J.; Jacobs, R.; Huang, H.; Parker, M. W.; Scott, J. W.; Rider, M. H.; Foitzik, R. C.; Kemp, B. E.; Baell, J. B.; Oakhill, J. S. Structural Determinants for Small-Molecule Activation of Skeletal Muscle AMPK alpha 2 beta 2 gamma 1 by the Glucose Importer SC4. *Cell Chem. Biol.* **2018**, *25*, 728–737.

Supporting Information

Understanding the Mechanism of Direct Activation of AMP-kinase: Towards a Fine Allosteric Tuning of the Kinase Activity

Elnaz Aledavood,^{†,#} Gleiciane Moraes,^{‡,#} Jeronimo Lameira,[‡] Ana Castro,[§] F. Javier

Luque,^{†,*} Carolina Estarellas^{†,*}

[†] Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicine (IBUB) and Institute of Theoretical and Computational Chemistry (IQTCUB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain

[‡] Faculdade de Ciências Naturais, Campus Universitário do Marajó-Breves, Universidade Federal do Pará (CUMB-UFPA), Breves, Brasil.

§ Instituto de Química Médica, Consejo Superior de Investigaciones Científicas (IQM-CSIC), 28006 Madrid, Spain.

These authors have equally contributed to this work.

Corresponding Authors: * E-mail: fjluque@ub.edu; * E-mail: cestarellas@ub.edu.

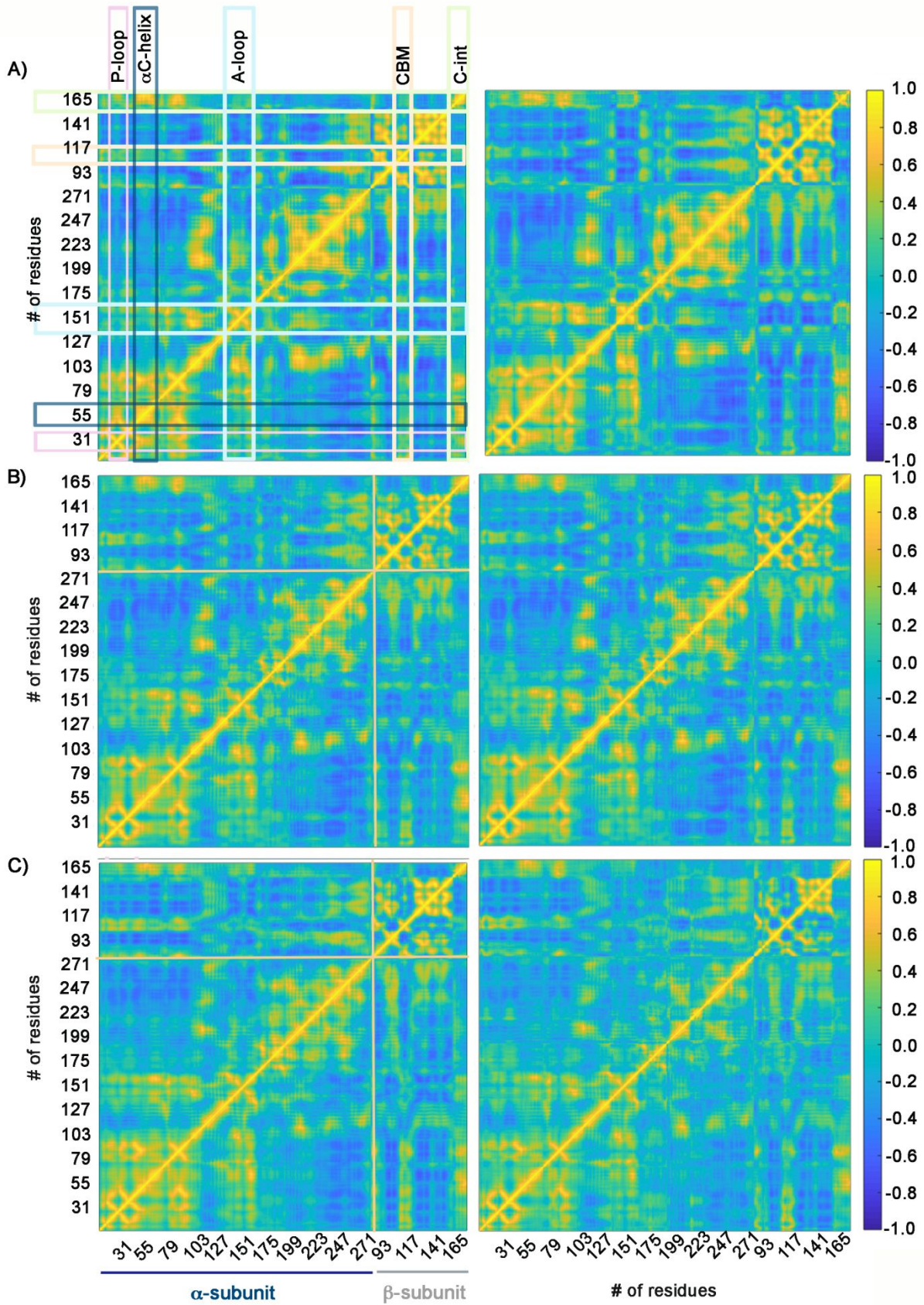


Figure S1. Dynamic cross-correlation (DCC) matrices for two independent 1 μ s MD simulations (left and right columns) of A) apo, B) holo and C) holo+ATP systems. Regions colored in yellow show high correlated fluctuations, while regions colored in blue show high-anticorrelated fluctuations. The structural subunits, together with the most important regions of AMPK systems, are indicated in squares for P-loop (pink), α C-helix (blue), A-loop (cyan), CBM domain (orange) and C-interacting helix (green).

3.1.2. Paper 2: “Structural Basis of the Selective Activation of Enzyme Isoform: Allosteric Response to Activators of β 1- and β 2-Containing AMPK Complexes”

Elnaz Aledavood,[†] Alessia Forte,[†] Carolina Estarellas,^{*,†} and F. Javier Luque^{*,†}

[†]Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicina (IBUB) and Institute of Theoretical and Computational Chemistry (IQTCUB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain

Published in Computational and Structural Biotechnology Journal, **2021**, *19*, 3394-3406



Structural basis of the selective activation of enzyme isoforms: Allosteric response to activators of β 1- and β 2-containing AMPK complexes



Elnaz Aledavood¹, Alessia Forte, Carolina Estarellas^{*,3}, F. Javier Luque^{*,2}

Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicine (IBUB) and Institute of Theoretical and Computational Chemistry (IQTCUB), University of Barcelona, Av. Prat de la Riba 171, Santa Coloma de Gramenet 08921, Spain

ARTICLE INFO

Article history:

Received 3 March 2021

Received in revised form 30 May 2021

Accepted 30 May 2021

Available online 05 June 2021

Keywords:

Enzyme isoforms

Allostery

Enzyme activation

AMPK

Structure-function relationships

Protein dynamics

ABSTRACT

AMP-activated protein kinase (AMPK) is a key energy sensor regulating the cell metabolism in response to energy supply and demand. The evolutionary adaptation of AMPK to different tissues is accomplished through the expression of distinct isoforms that can form up to 12 complexes, which exhibit notable differences in the sensitivity to allosteric activators. To shed light into the molecular determinants of the allosteric regulation of this energy sensor, we have examined the structural and dynamical properties of β 1- and β 2-containing AMPK complexes formed with small molecule activators A-769662 and SC4, and dissected the mechanical response leading to active-like enzyme conformations through the analysis of interaction networks between structural domains. The results reveal the mechanical sensitivity of the α 2 β 1 complex, in contrast with a larger resilience of the α 2 β 2 species, especially regarding modulation by A-769662. Furthermore, binding of activators to α 2 β 1 consistently promotes the pre-organization of the ATP-binding site, favoring the adoption of activated states of the enzyme. These findings are discussed in light of the changes in the residue content of β -subunit isoforms, particularly regarding the β 1Asn111 \rightarrow β 2Asp111 substitution as a key factor in modulating the mechanical sensitivity of β 1- and β 2-containing AMPK complexes. Our studies pave the way for the design of activators tailored for improving the therapeutic treatment of tissue-specific metabolic disorders.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Protein isoforms provide complexity to the structural and functional space of the human proteome [1]. Diverse mechanisms, such as genetic changes in protein-coding regions, alternative splicing and post-translational alterations, mediate the formation of isoforms, enriching the functional diversity by modulating enzymatic activities, molecular interactions and subcellular localizations [2,3]. In this context, disclosing the molecular basis of the structural variation between isoforms is critical for understanding their functional adaptation to different tissues and organs, and for the design of therapeutic approaches tailored to the spatio-temporal context of diseases [4–7].

The evolutionary adaptation of protein isoforms is a key aspect of mammalian adenosine monophosphate-activated protein kinase (AMPK), which can form up to 12 different combinations according to the tissue-dependent expression of isoforms identified for the three structural components required for full activity [8–10]. This can be understood from the role played as a fuel-sensing enzyme in preserving the cellular energy homeostasis [11,12], since AMPK activation reduces the rate of anabolic pathways and up-regulates catabolic processes, resulting in increased levels of ATP. Furthermore, AMPK is implicated in metabolic disorders such as obesity and type 2 diabetes, cardiovascular diseases and cancer [13–15], attracting widespread interest as a therapeutic target [16].

AMPK is a heterotrimeric complex consisting of a catalytic α -subunit and two regulatory (β and γ) components, which are encoded by multiple genes, including two α (α 1, α 2), two β (β 1, β 2), and three γ (γ 1, γ 2, γ 3) isoforms [17,18]. The catalytic subunit has a Ser/Thr kinase domain at the N-terminus and its C-terminus is necessary for the formation of the complex with the other components. The β -subunit has a central carbohydrate-binding module (CBM) that mediates AMPK interaction with glycogen, and the C-terminal region acts as a scaffold for the heterotrimeric assembly.

* Corresponding authors.

E-mail addresses: cestarellas@ub.edu (C. Estarellas), fjluque@ub.edu (F. Javier Luque).

¹ ORCID: 0000-0002-4837-7849.

² ORCID: 0000-0002-8049-3567.

³ ORCID: 0000-0002-0944-9053.

The γ -subunit has four tandem repeats of the cystathionine β -synthase (CBS) domain and contains four adenine nucleotide binding sites, which mediate the allosteric activation of the kinase activity by AMP, enabling AMPK to react to subtle fluctuations in the AMP/ATP ratio [19,20].

Due to its role in energy homeostasis, AMPK is finely regulated by different mechanisms [10], such as phosphorylation of α 2Thr172 in the activation loop of the kinase domain. Thus, binding of AMP to the γ -subunit, in conjunction with Thr172 phosphorylation by upstream kinases such as LKB1 and CaMKK β , converts the inactive enzyme into an active species, which is several thousand-fold more active [20,21]. AMPK can also be indirectly activated by compounds such as metformin, phenformin and oligomycin [22], which increase the intracellular levels of AMP. Finally, a direct activation of AMPK can be triggered by modulators such as the thienopyridone drug A-769662 (Table 1) [23], which binds to the so-called allosteric drug and metabolite (ADaM) site located at the interface between α - and β -subunits [24,25], promoting an activation mechanism independent of Thr172 phosphorylation. Indeed, binding of A-769662 can increase the AMPK activity >90-fold when Ser108 in the CBM domain of the β -subunit is phosphorylated [26], and may also protect against dephosphorylation of Thr172 [21].

Several direct AMPK activators have been reported in the last few years, showing in some cases a marked isoform selectivity, [27] as can be noticed upon inspection of the biochemical data collected in Table 1 for selected small molecule AMPK activators. A-769662 is active in α 2 β 1 γ 1 but not in α 2 β 2 γ 1, at least up to concentrations of 10 μ M [24], demonstrating selectivity toward β 1-containing AMPK complexes. Likewise, the enzyme activation is ~7-fold larger in α 2 β 1 γ 1 relative to α 2 β 2 γ 1. A mild β 1-selective activation is found for 991 and SC4, as they activate β 1- and β 2-containing AMPK complexes, although a higher (~2-fold) activation is observed in the former case [24,28,29]. The activation also seems to be slightly larger for the enzyme containing the α 2 isoform. While PF-249 is an activator selective for β 1-containing complexes, PF-739 is a pan-activator that activates α 2 β 1 γ 1 and α 2 β 2 γ 1 complexes, though it still exhibits a larger affinity for the β 1-containing isoforms (EC₅₀ ratio of ~ 8 and ~ 15 between α 2 β 1 γ 1/ α 2 β 2 γ 1 and α 1 β 1 γ 1/ α 1 β 2 γ 1, respectively) [30]. PF-739 seems also to have a larger effect on the α 2-containing AMPK complexes, as the ratio of the EC₅₀ values is 1.8–3.2 larger for α 1 β 1 γ 1 and α 1 β 2 γ 1 relative to the α 2-containing enzymes. Finally, MT47-100, a structural analogue of A-769662 possessing a dihydroquinoline ring instead of the thienopyridone core, activates both α 1 β 1 γ 1 and α 2 β 1 γ 1, but promotes the inhibition of β 2-containing complexes [31].

Understanding the molecular mechanisms that underlie the regulatory effect of direct activators, particularly targeting selectively a given isoform, is of utmost relevance for gaining insight into the puzzling tissue-dependent modulation of AMPK, and to disclose drugs active against specific pathological disorders. Given the sensitivity of the AMPK enzyme to the precise combination of different isoforms, as exemplified by the distinctive trends observed in the activation of complexes with α 1/ α 2 and β 1/ β 2 isoforms (see above), we have adopted a 'divide-and-conquer' strategy in order to explore the molecular basis of the selective isoform activation of AMPK, specifically focusing on the role played by β 1- and β 2-subunits. To this end, extended molecular dynamics (MD) simulations have been performed for the apo forms, the holo complexes formed with A-769662 and SC4 bound to the ADaM site, and the ternary (holo + ATP) complexes formed by the enzyme bound to both activator (in the ADaM site) and ATP (in the ATP-binding site). The analysis of the structural and dynamical properties of AMPK complexes and the identification of the interaction networks between α - and β -subunits have unveiled distinctive

molecular determinants of the allosteric regulation exerted by direct activators on β 1- and β 2-containing enzymes.

2. Material and methods

2.1. Molecular dynamics simulations

Extended molecular dynamics (MD) simulations were utilized to analyze the structural and dynamical characteristics of the simulated system. For this purpose, the α 2 β 1 γ 1 systems were built up using the complexes with A-769662 (PDB entry 4CFF) [24] and SC4 (PDB entry 6B1U) [29]. On the other hand, the system related to the complex of α 2 β 2 γ 1 bound to SC4 (PDB entry 6B2E) [29] was also used as a template to model the complexes with A-769662. Following our previous studies [32], the γ -subunit was not considered in MD simulations due to several reasons. First, the ADaM site is shaped only by α - and β -isoforms. Furthermore, the lack of precise structural information about stretches of both α - and β -subunits, particularly regarding the C-terminal regions, which are located close to the γ -subunit, would introduce an additional level of uncertainty, opening the way to potential artefacts in the simulations. Finally, inclusion of the γ -subunit would have required a larger computational cost in order to guarantee a proper sampling of the dynamical motions of the three isoforms. Accordingly, following the 'divide-and-conquer' strategy outlined above, the simulated systems comprise only α - and β -subunits. Specifically, simulations were performed for residues 8–278 of the α 2 isoform, and residues 78–173 and 77–171 of the β 1- and β 2-isoforms, which were solved without disruptions in the X-ray structures. Finally, these structures were used to model the apo protein, the complexes of the activators bound to the phosphorylated Ser108 (pSer108)-containing isoforms (holo), and the corresponding holo + ATP complexes with both activator in the ADaM site and ATP in the ATP-binding site.

Simulations were performed using the AMBER18 package [33] and the Amber ff99SB-ILDN force field [34] for the protein, whereas the ligands (A-769662, SC4) were parameterized using the GAFF [35] force field in conjunction with restrained electrostatic potential-fitted (RESP) partial atomic charges derived from B3LYP/6-31G(d) calculations [36]. The standard protonation state at physiological pH was assigned to ionisable residues, and a capping group (N-methyl) was added to the C-terminus of the α -subunit. The simulated systems were immersed in an octahedral box of TIP3P water molecules [37]. Counterions atoms were added to maintain the neutrality of the simulated systems [38]. The final systems included the AMPK protein, around 25,000–28,000 water molecules, and a variable number of Na⁺ and Cl⁻ ions, leading to simulated systems that contain between 81,000–90,000 atoms (specific details are given in Table S1).

Simulations were performed in the NPT ensemble for equilibration and NVT for MD productions using periodic boundary conditions and Ewald sums (grid spacing 1 Å) for treating long-range electrostatic interactions. Apo, holo and holo + ATP systems were simulated in triplicate. The minimization of the systems was performed refining the position of hydrogen atoms in the protein (2000 cycles of steepest descent algorithm followed by 8000 cycles of conjugate gradient), subsequently minimizing the position of water molecules (using again the previous scheme), and finally minimization of the whole system (4000 cycles for steepest descent and 1000 cycles of conjugate gradient). Later, the temperature of the system was gradually raised from 100 to 300 K in 5 steps, 50 ps each using the NVT ensemble. In this process, suitable restraints (5 kcal mol⁻¹ Å⁻²) were imposed to keep the ligand (activator, ATP) in the binding pocket and prevent artefactual rearrangements along the equilibration stage. In order to equilibrate

Table 1
Activation of AMPK isoforms by selected direct activators.

Activator	Property ^a	Isoform			
		$\alpha 2\beta 1\gamma 1$	$\alpha 2\beta 2\gamma 1$	$\alpha 1\beta 1\gamma 1$	$\alpha 1\beta 2\gamma 1$
A-769662	Activation fold ^b K_d (μM) ^b	14.3 0.40	NA ^c –	2.0 0.51	NA ^c 14.5
991	Activation fold ^b K_d (μM) ^b	12.2	5.4	4.8	2.4
		7.7 ^{d,e}	5.7 ^d		
		0.06 ^f	–	0.06	0.51
		0.085		0.078	1.18
SC4	Activation fold ^g EC_{50} (nM) ^g	5.2	2.9	2.4	1.2
		–	17.2	5.1	–
PF-249	EC_{50} (nM) ^h	8.04	> 40 μM	9.37	> 40 μM
PF-739	EC_{50} (nM) ^h	5.23	42.4	8.99	136
MT47-100	Activation fold ⁱ	2.3	0.5	2.5	0.4

^a Activation fold measured relative to the basal activity of the enzyme. Dashed lines stand for systems where no experimental data have been reported, to the best of our knowledge. ^b Ref 24. ^c No activation detected up to a concentration of 10 μM . ^d Ref. 28. ^e Determined for $\alpha 2\beta 1\gamma 2$. ^f Values obtained by using either biolayer interferometry or circular dichroism. ^g Ref. 29. ^h Ref. 30. ⁱ Ref. 31.

the density of the system an additional 5 ns step performed in the NPT ensemble. In addition, the restraints were progressively eliminated in this later step. Production MD simulations were run for 1 μs per replica, leading to a total simulation time of 24 μs for the ligand-bound AMPK complexes, and 6 μs for the two apo species of AMPK.

2.2. Essential dynamics

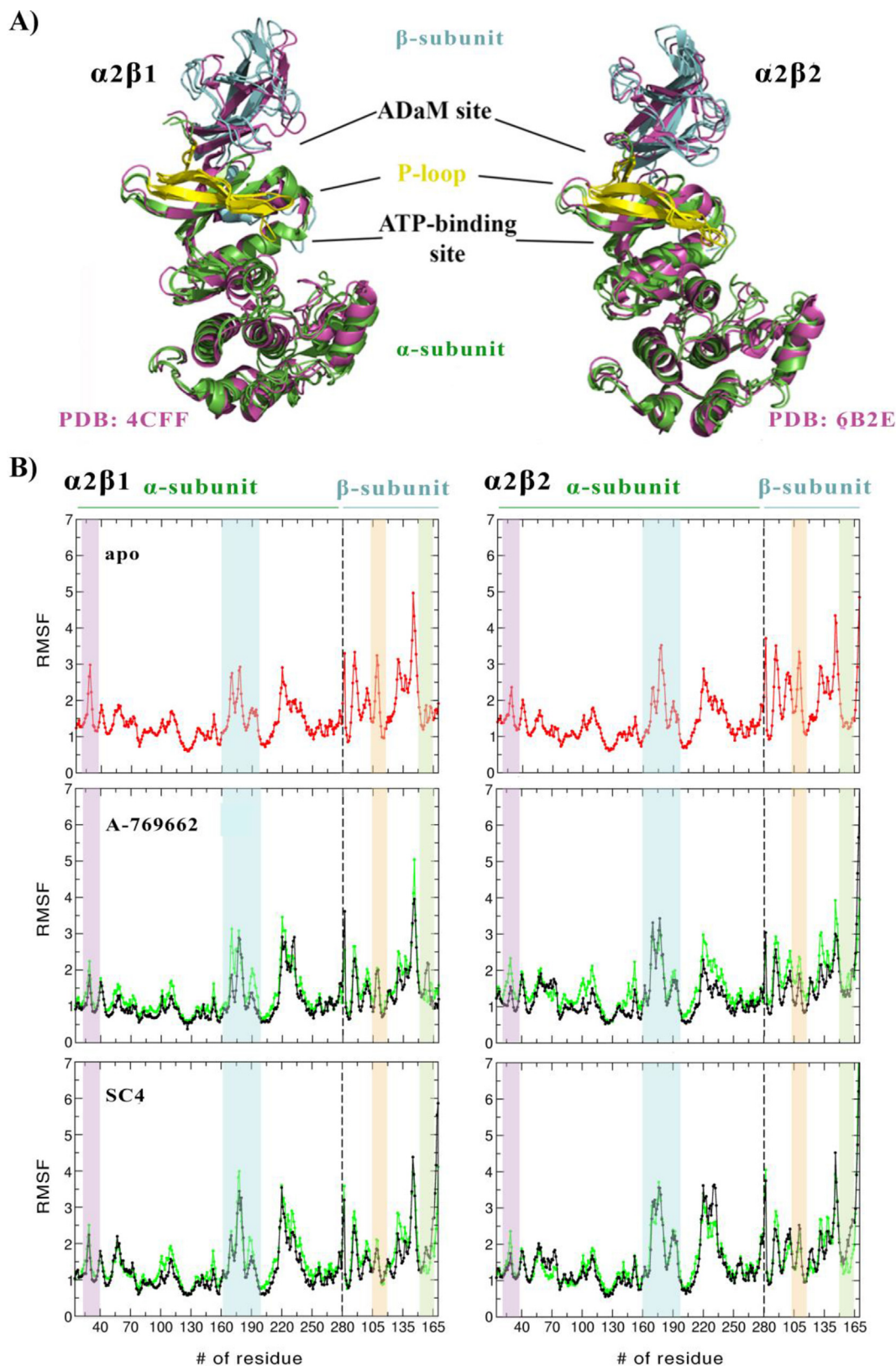
This method was utilized to specify the most important motions from the structural variance sampled in MD simulations. In essential dynamics (ED) [39], the dynamics along the individual modes can be studied and visualized separately, so we can filter the main collective motions during our simulations. Therefore, the positional covariance matrix is created and diagonalized in order to achieve the collective deformation modes, i.e. the eigenvectors, while the eigenvalues account for the contribution of each motion to the structural variance of the protein. ED analysis was done for 25,000 snapshots from the last 500 ns of each simulation, taking into account only the backbone atoms and the calculations were performed with PCAsuite program (available at <http://www.mmb.irbbarcelona.org/software/pca-suite/pca-suite.html>), which is integrated in the pyPCcazip program, a suite of tools for compression and analysis of molecular simulations [40].

2.3. Conformational entropy

The quasi-harmonic approximation proposed by Schlitter [41] was used to calculate the conformational entropy sampled for the different simulation systems. Due to the sensitivity of these calculations to simulation time and sampling [42], the conformational entropy determined at regular interval times ($S(t)$) along the MD simulation was used to estimate the conformational entropy at infinite time (S^∞) [43]. In particular, S^∞ was determined by fitting the asymptotic behaviour of the curve according to the expression $S(t) = S^\infty \frac{t}{t+\alpha}$, where α is an adjustable parameter by curve fitting. In order to enable a direct comparison between the conformational entropies of the different $\alpha 2\beta 1$ and $\alpha 2\beta 2$ systems, only the set of atoms of the protein backbone were considered in entropy calculations, thus avoiding a bias originated from the residue substitutions between $\beta 1$ - and $\beta 2$ -isoforms in the calculated entropies (note that a different choice of atoms, including both backbone and side chains, was used in the comparison of the conformational entropies reported for $\alpha 2\beta 1$ bound to A-769662 in a previous study [32]).

2.4. Dynamical cross-correlation analysis

To complement the information gained from the ED analysis, dynamical cross-correlation (DCC) was used to examine the corre-



lation motion of residues along a given trajectory. To this end, all the snapshots were aligned by means of least-square fitting of C_α atoms of the whole protein to the equilibrated starting configuration. Then, the DCC matrix was determined as noted in Eq. (1).

$$C_{ij} = \frac{C_{ij}}{C_{ii}^{1/2} C_{jj}^{1/2}} = \frac{\langle \mathbf{r}_i \mathbf{r}_j \rangle - \langle \mathbf{r}_i \rangle \langle \mathbf{r}_j \rangle}{\left[\left(\langle \mathbf{r}_i^2 \rangle \langle \mathbf{r}_i \rangle^2 \right) \left(\langle \mathbf{r}_j^2 \rangle - \langle \mathbf{r}_j \rangle^2 \right) \right]^{1/2}} \quad (1)$$

where the position vectors of two C_α atoms i and j fitted in the structure at time t are denoted as $\mathbf{r}_i(t)$ and $\mathbf{r}_j(t)$, respectively.

The cross-correlation coefficients range from -1 to $+1$, which represent anticorrelated and correlated motions, respectively, whereas values close to zero indicate the absence of correlated motions [44]. This analysis was performed using the module available in AMBER package. The similarity between the DCC matrices computed for the three replicas run for apo, holo and holo + ATP systems was estimated using the Tanimoto similarity index [45]. This parameter is a distance metrics used to quantify the degree of similarity between two sets of data. While this index is widely adopted to compare the descriptors that characterize the chemical structure of molecules, in this study it is used to compare the correlated motions determined for pairs of residues in the AMPK complexes.

2.5. Interaction energy network

Networks of local interactions are intrinsically linked to the structural response of proteins to external factors [46]. For our purposes, Weighted Implementation of Suboptimal Path (WISP) [47] was utilized to analyze the allosteric network. This method enabled us to perform a dynamic network analysis to understand how the binding of a ligand in an allosteric cavity could affect another binding site. In particular, WISP relies on the dynamical interdependence among the protein residues. To this end, each amino acid is treated as a node, which was located at the residue center-of-mass, and the interdependence among nodes is represented as a connecting edge with an associated numeric value that reflects its strength. The interdependence is determined from an $N \times N$ matrix C (N is the number of nodes) with values corresponding to the weights of each edge, reflecting the correlated motion among node–node pairs. Finally, the weight between the edge that connects nodes i and j is expressed as $w_{ij} = -\log(|C_{ij}|)$, so that highly correlated or anticorrelated motions are characterized by small values of w_{ij} . This analysis was performed for the last 500 ns of the MD simulations.

Table 2

RMSD and standard deviation (Å) determined for the protein backbone of the snapshots sampled along the last 500 ns of MD simulations performed for all systems (apo, holo and holo + ATP states of AMPK isoforms $\alpha 2\beta 1$ and $\alpha 2\beta 2$). Values were determined using the energy-minimized holo + ATP species averaged for the last 200 ns of each simulation system as reference structure.

Activator	System	Replica 1	Replica 2	Replica 3	Average
$\alpha 2\beta 1$ A-769662	apo	2.5 ± 0.5	3.1 ± 0.6	2.6 ± 0.3	2.7
	holo	2.3 ± 0.4	1.9 ± 0.3	2.2 ± 0.3	2.1
	holo + ATP	1.9 ± 0.2	2.0 ± 0.3	1.9 ± 0.2	1.9
SC4	apo	3.2 ± 0.5	2.9 ± 0.5	3.1 ± 0.4	3.1
	holo	2.9 ± 0.4	2.2 ± 0.3	2.3 ± 0.6	2.5
	holo + ATP	2.0 ± 0.2	2.3 ± 0.2	2.2 ± 0.2	2.2
$\alpha 2\beta 2$ A-769662	apo	2.7 ± 0.3	3.3 ± 0.3	3.9 ± 0.4	3.3
	holo	4.2 ± 0.5	3.2 ± 0.4	4.5 ± 0.4	4.0
	holo + ATP	2.1 ± 0.2	2.2 ± 0.2	1.9 ± 0.1	2.1
SC4	apo	3.0 ± 0.3	2.8 ± 0.2	3.4 ± 0.4	3.1
	holo	3.3 ± 0.3	2.9 ± 0.3	3.1 ± 0.3	3.1
	holo + ATP	2.3 ± 0.2	2.5 ± 0.2	2.7 ± 0.2	2.5

3. Results

MD simulations were performed to examine the structural and dynamical changes triggered by A-769662 and SC4 through the analysis of the apo form of AMPK considering both $\alpha 2\beta 1$ and $\alpha 2\beta 2$ isoforms, and their holo and holo + ATP complexes, as well as the interaction networks implicated in the dynamical response to activator binding. Three independent replicas (1 μ s/replica) were run for each system, leading to a total simulation time of 6 μ s for the apo species and 24 μ s for the ligand-bound complexes.

3.1. Global structural analysis of AMPK complexes

The effect of activator and ATP binding on the structural features of AMPK complexes was examined from the positional root-mean square deviation (RMSD) of the protein backbone along the simulation. The average structure of the holo + ATP species sampled in the last 200 ns of the three simulations run for each activator was used as reference. There is a large structural resemblance between the averaged holo + ATP species, as noted in the superposition of the protein backbone for the complexes of $\alpha 2\beta 1$ (Fig. 1A left) and $\alpha 2\beta 2$ (Fig. 1A right) bound to ATP and either A-769662 or SC4. The structural resemblance of the energy-minimized averaged holo + ATP structures is also noted in RMSD values of 0.8–1.6 Å, which supports the preservation of the overall fold of the protein upon binding of both activator and ATP. Preservation of the overall structure of holo + ATP complexes is also supported by the small fluctuations observed in RMSD profiles (Fig. S1 and S2), as noted by the values of 1.9–2.3 Å for $\alpha 2\beta 1$ and 1.9–2.7 Å for $\alpha 2\beta 2$ (Table 2). For the $\alpha 2\beta 1$ systems, the RMSD of the sampled structures is reduced from apo (2.5–3.2 Å) to holo (1.9–2.9 Å) to holo + ATP (1.9–2.3 Å), revealing that binding of activator and ATP stabilizes the overall fold of the complex (Table 2). This effect is less apparent for the $\alpha 2\beta 2$ complexes, as the RMSD values of the apo species (2.7–3.9 Å) are similar or even lower than those determined for the holo form (2.9–4.5 Å), suggesting that the activator exerts a weaker structural stabilization upon binding to $\alpha 2\beta 2$ isoform.

On the other hand, the per-residue mean square fluctuation (RMSF) profile is similar for the two isoforms, as noted in the resemblance of the fluctuation pattern obtained by averaging the RMSF of the three replicas run for every system (Fig. 1B). The highest fluctuations in the α -subunit mainly originate from residues in the activation loop (residues 165–185) and the α -helix between residues 210–230. The largest fluctuations in the β -subunit correspond to residues in the CBM domain, which contains Ser108, and regions proximal to the C-interacting helix. Compared to the

apo forms, binding of activators generally lead to subtle changes in the RMSF profile, such as enhanced fluctuations in the activation loop, and the α -helix mentioned above, while a reduction in the fluctuation of residues in the β -subunit is observed. In contrast, the RMSF profiles remain generally unaltered upon subsequent binding of ATP, as noted in the comparison of the results obtained for holo and holo + ATP systems.

3.2. Dynamics of AMPK complexes

The dynamical behavior of proteins exerts a major role in their function at multiple time scales, affecting processes such as enzyme catalysis [48–50]. Here the effect of activator on the dynamics of AMPK was examined by analyzing the essential dynamics (ED) and conformational entropy of the protein back-

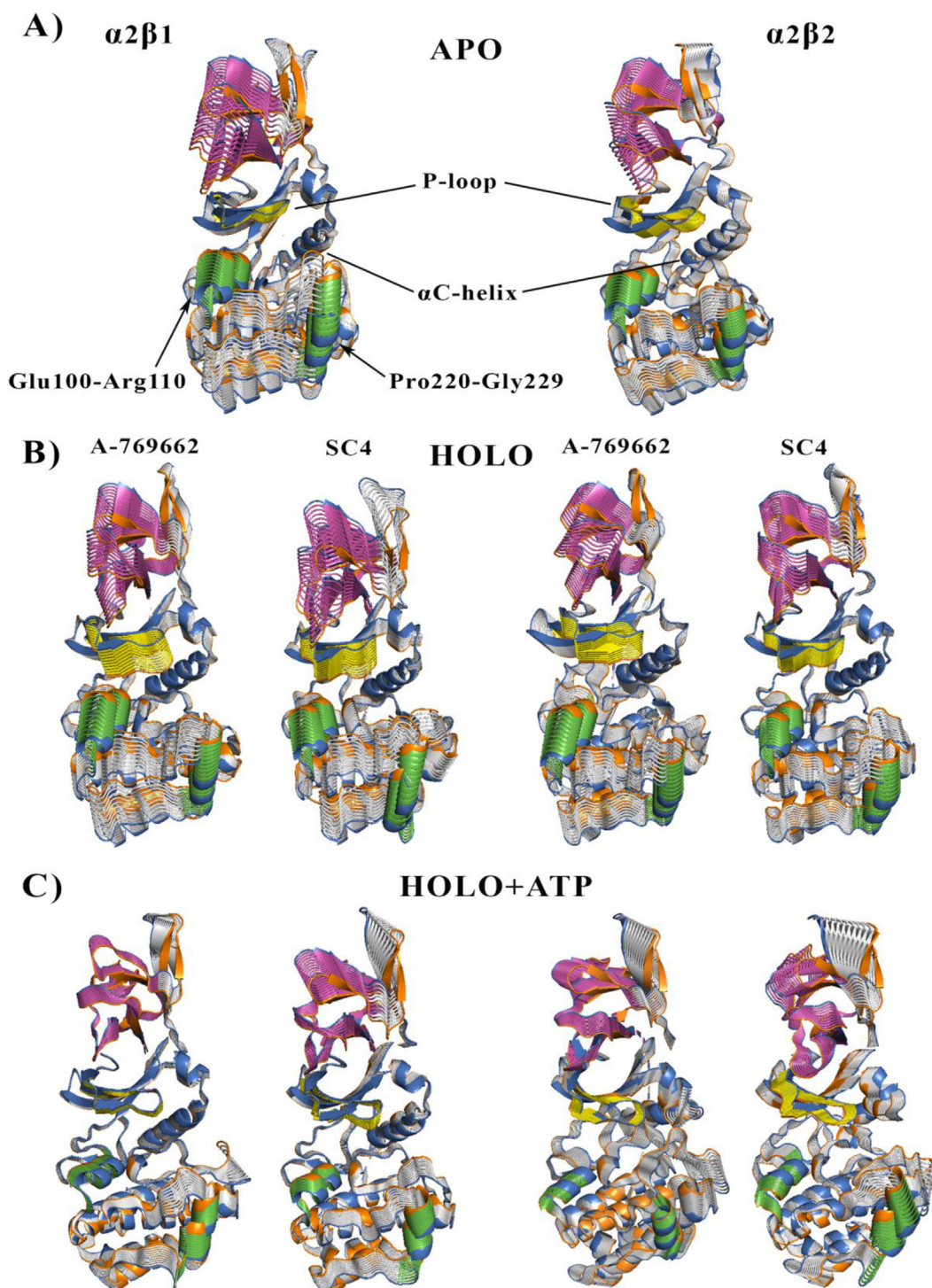


Fig. 2. Representation of the first essential motion derived from the ED analysis of the protein backbone for A) the $\alpha 2\beta 1$ and $\alpha 2\beta 2$ apo species, and their B) holo and C) holo + ATP complexes with A-769662 and SC4 determined from the snapshots sampled along the last 500 ns of simulations. The P-loop is shown in yellow, the helices formed by residues 100–110 and 220–229 in the α -subunit in green, and the CBM domain in magenta. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

bone in conjunction with the cross-correlation mapping at the residue level.

ED was used to characterize the effect of ligand binding on the major motions of the protein backbone. For the two apo species, the first essential motion involves a concerted bending that brings the core of α -helices in the α -subunit and the β -sheets in the β -subunit closer and then moves them apart (Fig. 2A). Nevertheless, there are slight differences in the extent of this motion for certain structural elements, such as the deformation of the CBM unit and the P-loop, which is located at the interface of ADaM and ATP-binding sites (Fig. S3). On average, the first motion accounts for 41% of the structural variance in $\alpha 2\beta 1$, and 31% in $\alpha 2\beta 2$. The difference in the conformational variance observed in the first essential motion is also reflected in the cumulative variance explained by the three main essential motions, which amounts to 66% and 57% of the structural flexibility in $\alpha 2\beta 1$ and $\alpha 2\beta 2$, respectively (Fig. S4). This highlights the significant contribution of the first essential motion to the conformational flexibility of the AMPK complexes. On the other hand, the conformational entropy of the two apo forms is basically the same (~ 19 kcal K⁻¹ mol⁻¹; Table 3).

Binding of activator and ATP reduces the conformational variance originated from the first three essential motions in $\alpha 2\beta 1$ complexes with A-769662 and SC4 (from 66% to 58% to 47% from apo to holo to holo + ATP in the case of A-769662), reflecting the trend observed for the first essential motion (Fig. S4). This leads to a progressive decrease in the conformational entropy of the protein skeleton (Table 3). In contrast to the apo species, the activator promotes the synchronous motion of the P-loop and the CBM domain, which bends toward the α -subunit (Fig. 2 and S5). Subsequent binding of ATP in the kinase domain, however, reduces the overall flexibility of the protein, especially regarding the P-loop and most of the helices in the α -subunit, but also elements of the β -subunit proximal to the ADaM site, reflecting the synergy between activator and ATP in increasing the stiffness of the AMPK complex.

The dynamical behaviour observed for $\alpha 2\beta 1$ is less regular in $\alpha 2\beta 2$. Thus, the structural variance explained by the first essential motion in the apo form remains unaltered or is even slightly increased in the presence of the activator (Fig. S4). Furthermore, although the activator increases the conformational flexibility of the P-loop, the response of other structural elements is more variable, as noted for instance in the deformations observed for certain helices in the α -subunit, such as those formed by residues Glu100-Arg110 and Pro220-Gly229, or the β -sheets in the N-terminus of the β -subunit (Fig. 2 and S5), this effect being generally more pronounced for SC4-bound complexes. The lack of a uniform, consistent pattern of structural flexibility is also found in the

holo + ATP systems, as well as in the distinct changes observed in the conformational entropy for A-769662 and SC4 (Table 3).

To complement the preceding results, dynamical cross-correlation (DCC) analysis was performed to examine the correlated motions of residues in AMPK complexes. Results for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ complexes interacting with A-769662 are shown in Fig. 3. The overall trends are similar to those obtained for the SC4-bound systems (Fig. S6). Binding of activator and ATP promotes a general decrease in the dynamical correlation of residues. This is reflected in the weakening of the correlated regions in the DCC map (Fig. 3B-C and S6B-C), as one may notice a progressive reduction in the number and intensity of the areas that exhibit a pronounced correlation (shown in yellow and blue for highly correlated and anticorrelated fluctuations between residues in plots B and C, which correspond to holo and holo + ATP species). Furthermore, one may note that this effect is more noticeable for the $\alpha 2\beta 1$ isoform. The apo systems (Fig. 3A and S6A) exhibit a correlation between the P-loop, the α -helix formed by residues 100–110 and the α C-helix (as noted by the yellow marks), which pertain to the α -subunit. In turn, they are correlated with the C-interacting helix in the β -subunit. Binding of activator and ATP reinforces the correlation between the motion of the CBM domain and the P-loop (black square in Fig. 3 and S6). This effect is more pronounced for the $\alpha 2\beta 1$ isoform, which shows a more consistent dynamical response to activator binding, as noted in similarity indexes of 0.91 (A-769662) and 0.84 (SC4) obtained from the comparison of DCC matrices derived for the simulations of $\alpha 2\beta 1$ (Table S2), whereas the similarity is reduced to 0.75 for $\alpha 2\beta 2$. Remarkably, this difference is even larger for the dynamical coupling between α - and β -subunits, as noted in similarity indexes of 0.89 (A-769662) and 0.83 (SC4) for $\alpha 2\beta 1$, which are reduced to 0.59 and 0.66 for $\alpha 2\beta 2$, respectively.

Despite the overall gross similarity in the dynamical behavior of both $\alpha 2\beta 1$ and $\alpha 2\beta 2$ isoforms, the preceding results reveal the occurrence of subtle differences between $\beta 1$ - and $\beta 2$ -containing AMPK complexes. In particular, the results suggest that the $\alpha 2\beta 2$ complex has a larger resilience to the structural modulation exerted by the activator, whereas the $\alpha 2\beta 1$ isoform is more sensitive to the conformational adaptation induced upon activator binding to the ADaM site, enhancing the stiffness of protein backbone for the $\beta 1$ -containing complex, especially for A-769662.

3.3. Pre-organization of the ATP-binding site

To examine the implications of the differences in the dynamical response to binding of activator, we determined the structural resemblance of the ATP-binding site along the simulations obtained for apo, holo, and holo + ATP systems. To this end, we decided to use the average structure of the holo + ATP complex as reference, this choice being motivated by two major reasons: i) the large structural resemblance of the energy-minimized averaged holo + ATP structures for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ complexes with A-769662 and SC4 (see Fig. 1A), and ii) the notable reduction in the conformational flexibility of the ATP binding site for the holo + ATP species (see plots C in Fig. 2).

For the holo + ATP species, the residues in the ATP-binding site sample a reduced conformational space, with a peak generally centred at a positional RMSD of 1.0–1.6 Å for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ isoforms (Fig. 4). In contrast, the apo form exhibits a wider distribution, showing either a single or a bimodal RMSD profile, with values ranging from 1.0 to 5.5 Å. This reflects that the ATP-binding site in the apo species is more conformationally mobile than in the holo + ATP state. Due to this conformational flexibility, the ATP-binding site can often adopt conformations close to those populated in the holo + ATP state, but also visit more dissimilar conformational regions. The most distinctive features emerge from the

Table 3

Entropy values at infinite simulation time (S^∞ ; kcal K⁻¹ mol⁻¹) determined for the protein backbone of apo, holo and holo + ATP systems of AMPK. Changes in conformational entropy (ΔS^∞) are determined relative to the apo species. Values averaged for the three independent simulations run for the systems (standard deviation in parenthesis).

Species	$\alpha 2\beta 1$	$\alpha 2\beta 2$
S^∞ (apo)	19.1 (0.2)	18.8 (0.3)
A-769662		
S^∞ (holo)	18.9 (0.1)	18.6 (0.2)
ΔS^∞ (holo)	-0.2	-0.2
S^∞ (holo + ATP)	18.5 (0.1)	18.4 (0.1)
ΔS^∞ (holo + ATP)	-0.6	-0.4
SC4		
S^∞ (holo)	18.8 (0.4)	19.0 (0.2)
ΔS^∞ (holo)	-0.3	+0.2
S^∞ (holo + ATP)	18.5 (0.3)	19.0 (0.6)
ΔS^∞ (holo + ATP)	-0.6	+0.2

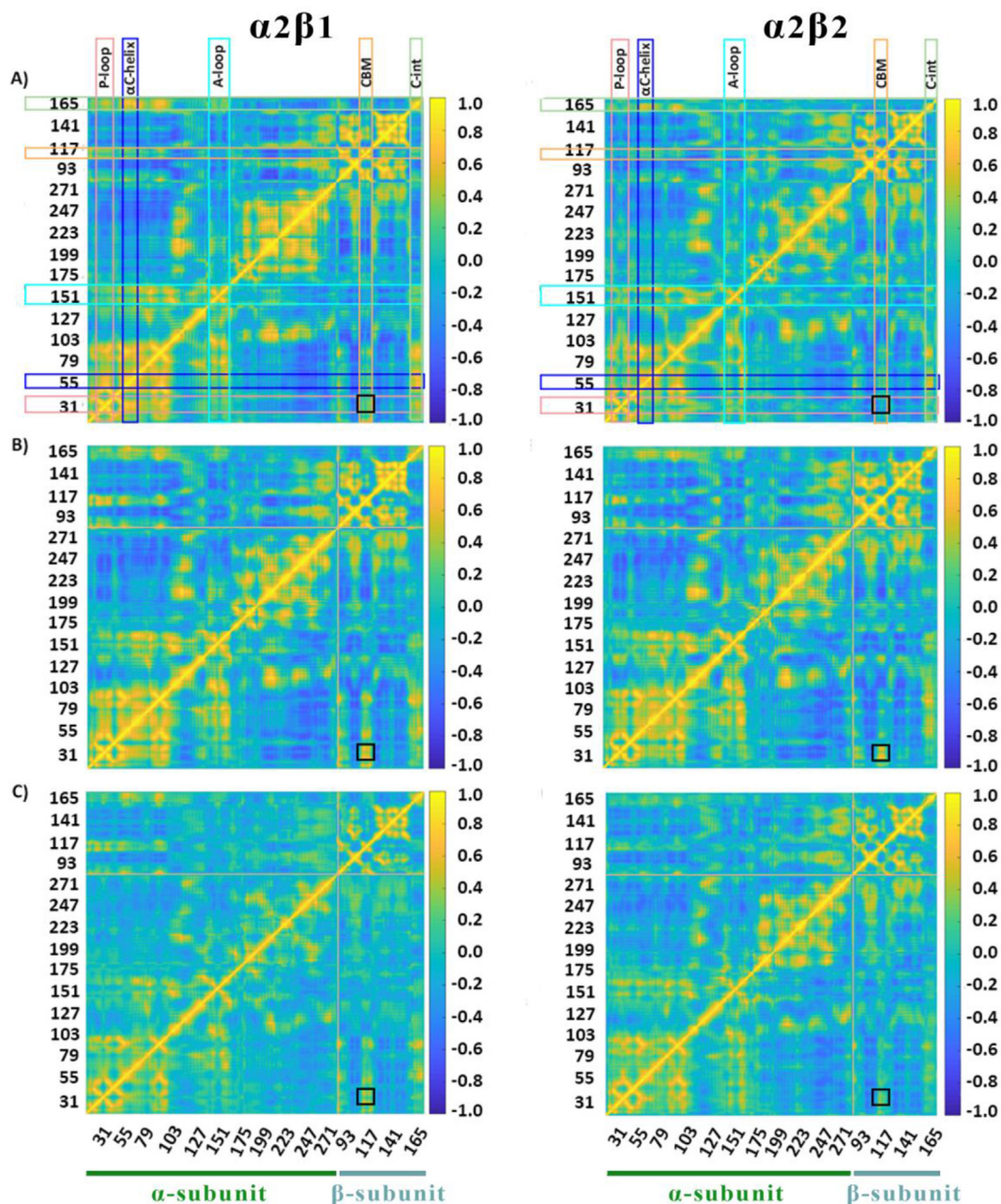


Fig. 3. Dynamic cross-correlation (DCC) matrices for A) apo, B) holo and C) holo + ATP complexes of (left) $\alpha 2\beta 1$ and (right) $\alpha 2\beta 2$ with A-769662. The x- and y-axis denote the numbering of residue in the α - and β -subunit. The residues pertaining to the α - and β -subunits are indicated at the bottom of the plot (green and blue lanes for α - and β -subunits, respectively). Specific regions of AMPK are highlighted with squares at the top of the plot: P-loop (pink), α C-helix (blue), activation loop (A-loop; cyan), CBM unit (orange) and C-interacting helix (C-int; green). Regions colored in yellow/blue show high correlated/anticorrelated fluctuations. The black square in the plots highlights the motion between the CBM domain and the P-loop. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

distribution profiles of the holo species. For the $\alpha 2\beta 1$ complexes, binding of the activator shifts the holo distribution toward the holo + ATP profile. This is clearly observed in the complexes with A-769662, where the peak in the apo distribution located at 2.8 Å is shifted to 1.5 Å for the holo species, and SC4, where the broad distribution observed for the apo species is converted into a narrower band that matches well the distribution of the holo + ATP system (Fig. 4). In contrast, the RMSD profile obtained for $\alpha 2\beta 2$ complexes reveals that the activator leads to structures where the ATP-binding site has less resemblance to the holo + ATP system (Fig. 4). The analysis of the distinct structural elements that shape the ATP-binding site shows that the differences observed for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ holo species mainly arise from

the conformational changes of the P-loop rather than from residues in the helical domain of the α -subunit (data not shown), suggesting that the activator is more effective in inducing the remodeling of the P-loop toward active-like conformations in $\beta 1$ -containing complexes.

These results are in agreement with the larger activation exerted by A-769662 for the $\alpha 2\beta 1$ isoform, while a weaker effect is observed for $\alpha 2\beta 2$ isoform (Table 1), supporting the hypothesis that the activator acts as a molecular glue in $\alpha 2\beta 1$ that connects the fluctuations of the CBM domain in the β -subunit with the P-loop, facilitating the arrangement of key structural elements, particularly the P-loop, toward conformations favouring the binding of ATP. These trends can be extrapolated to SC4, although the

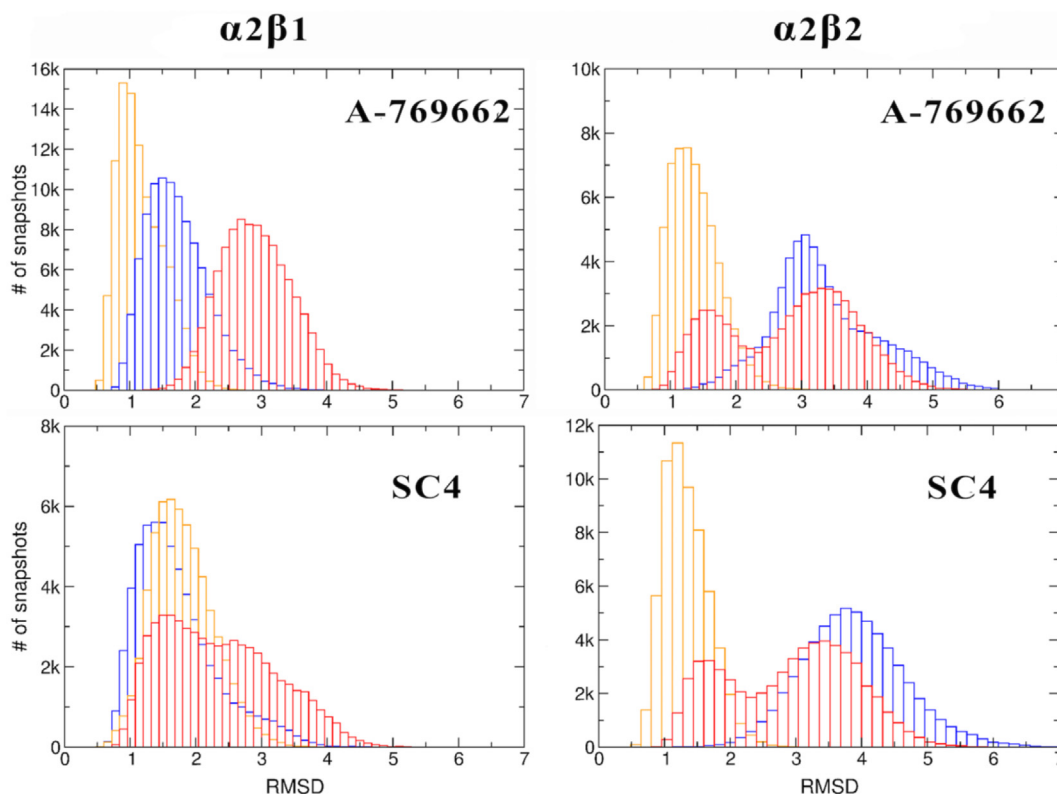


Fig. 4. Distribution of the positional deviation (RMSD; Å) of the structures sampled along the trajectories run for apo (red), holo (blue), and holo + ATP (orange) for the residues that shape the ATP-binding site (residues $\alpha 22$ – $\alpha 32$, $\alpha 42$ – $\alpha 46$, $\alpha 75$ – $\alpha 79$, $\alpha 142$ – $\alpha 147$, and $\alpha 153$ – $\alpha 157$). A total of 60,000 snapshots taken from the last 500 ns of MD simulations were considered for each system in the analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

difference between $\alpha 2\beta 1$ and $\alpha 2\beta 2$ is less apparent compared to A-769662 (Table 1).

3.4. Structural basis of the selective $\beta 1$ -containing AMPK activation by A-769662 and SC4

Accordingly, to explore the structural basis of the selective activation of $\beta 1$ -containing AMPK by A-769662, we examined the network of interactions formed between the activator with both P-loop and CBM unit, paying attention to the occurrence of residue substitutions between $\beta 1$ - and $\beta 2$ -isoforms (Fig. S7).

At the inner edge of the ADaM site in holo $\alpha 2\beta 1$, $\beta \text{Arg}83$ consistently forms a salt bridge with $\alpha \text{Asp}88$ (average distance of 3.2 ± 0.2 Å from the guanidinium unit to the carboxylate group), which is assisted by a hydrogen bond between $\alpha \text{Asp}88$ and the thienopyridone ring of A-769662 (2.8 ± 0.2 Å; Fig. 5A). At the outer edge of the ADaM site, the phosphorylated $\beta \text{Ser}108$ ($\beta \text{Ser}108\text{p}$) adopts two arrangements that enable interactions with $\alpha \text{Lys}31$ (4.3 ± 1.4 Å) and $\alpha \text{Lys}29$ (3.5 ± 0.8 Å). Binding of A-769662 is also assisted by hydrophobic contacts of the ligand with $\alpha \text{Val}11$, $\alpha \text{Leu}18$, $\alpha \text{Ile}46$, $\alpha \text{Phe}90$ and $\beta \text{Val}113$ (not shown in Fig. 5A for the sake of clarity in the graphical display), forming a tightly packed cage around A-769662.

This interaction pattern is drastically altered in holo $\alpha 2\beta 2$, as noticed in the loss of the $\beta 2\text{Arg}82$ – $\alpha \text{Asp}88$ interaction (N...O distance > 7.1 Å; Fig. 5B). This can be attributed to the replacement $\beta 1\text{Asn}111 \rightarrow \beta 2\text{Asp}111$, as this latter residue competes with $\alpha \text{Asp}88$ in forming electrostatic interactions with $\beta 2\text{Arg}82$. In fact, $\beta 2\text{Arg}82$ adopts a distinct conformation that enables a direct salt bridge with $\beta \text{Ser}108\text{p}$, which maintains a fixed orientation in contrast to the dual arrangement found in holo $\alpha 2\beta 1$. This weakens the

role of A-769662 in mediating the interaction between the CBM domain and the P-loop. Indeed, the impact of the $\beta 1\text{Asn}111 \rightarrow \beta 2\text{Asp}111$ change is reflected in the major interaction pathways identified from WISP analysis for the two holo species. In holo $\alpha 2\beta 1$, the CBM domain and the P-loop are connected via three pathways, implying in one case the passage through $\beta \text{Ser}108\text{p}$, and the involvement of the ligand, especially via the triad $\beta \text{Arg}83$ – $\alpha \text{Asp}88$ –A-769662, in the other two (Fig. 5C; close up view shown in Fig. S8). In contrast, only the $\beta \text{Ser}108\text{p}$ pathway is found in holo $\alpha 2\beta 2$, which suggests a weaker ability of A-769662 to act as a gluing molecule between the $\beta 2$ -subunit and the P-loop compared to holo $\alpha 2\beta 1$ (Fig. 5D; see also Fig. S8).

Compared to A-769662, SC4 exhibits a mild preference for activation of $\beta 1$ -containing AMPK (Table 1), raising questions about the interaction network formed by SC4 in the holo species. The analysis reveals a more complex scenario, since two arrangements are found for $\beta \text{Arg}83$ ($\beta \text{Arg}82$) in holo $\alpha 2\beta 1$ ($\alpha 2\beta 2$) (Fig. 6A–B). In one case, $\beta \text{Arg}83$ ($\beta \text{Arg}82$) forms a salt bridge with $\alpha \text{Asp}88$ (3.1 ± 0.2 Å), which also accepts a hydrogen bond from the imidazo [4,5-b]pyridine ring of SC4 (3.4 ± 0.2 Å), whereas $\alpha \text{Lys}29$ interacts with $\beta \text{Ser}108\text{p}$ (3.3 ± 0.7 Å) and is hydrogen bonded with the pyridine-like/imidazole-like nitrogen atoms of SC4 (3.6 ± 0.6 Å, Fig. 6A–B, residues in cyan). Alternatively, $\beta \text{Arg}83$ ($\beta \text{Arg}82$) interacts with $\beta \text{Ser}108\text{p}$ (Fig. 6A–B, residues in grey), forming an additional cation– π interaction with the imidazo[4,5-b]pyridine ring, which in turn is hydrogen bonded to $\alpha \text{Lys}29$. Remarkably, in holo $\alpha 2\beta 2$ this latter arrangement of $\beta \text{Arg}82$ is also stabilized by the interaction with $\beta \text{Asp}111$.

In contrast to A-769662, the occurrence of this dual interaction pattern reflects the electrostatic influence of the carboxylate group present in SC4, which would facilitate the adoption of the two

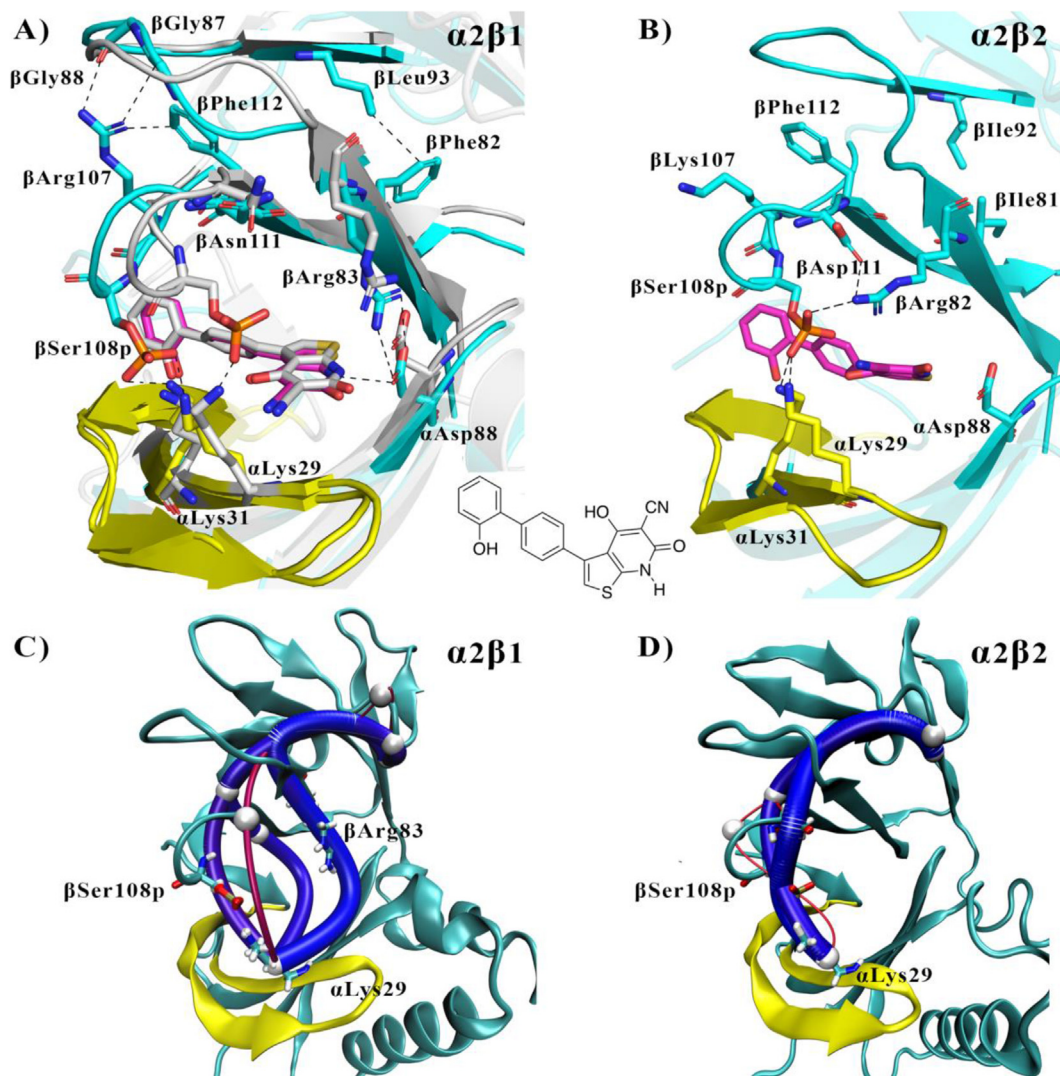


Fig. 5. Representation of selected interactions between the CBM, P-loop and A-769662 in holo A) $\alpha 2\beta 1$ and B) $\alpha 2\beta 2$. The two arrangements of β Ser108p in holo $\alpha 2\beta 1$ are shown in cyan and grey colors (P-loop in yellow). A-769662 is shown as coloured (carbon atoms in magenta) sticks in the ADaM site, and a 2D representation of the chemical structure is shown for the sake of clarity. Selected interactions between residues are highlighted as dashed lines. Major interaction networks obtained from WISP analysis for C) $\alpha 2\beta 1$ and D) $\alpha 2\beta 2$ systems (close up views available in Fig. S8). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

arrangements of β Arg83 (β Arg82). When β Arg83 forms a salt bridge with α Asp88 in holo $\alpha 2\beta 1$, the CBM unit and the P-loop are then connected by two main pathways, which reveal the involvement of either β Ser108p or the triad β Arg83-SC4- α Lys29 (Fig. 6C left; close up view shown in Fig. S8). However, when the β Arg83- α Asp88 salt bridge is lost (Fig. 6C right, and Fig. S8), the CBM and P-loop are linked via β Asn111, the phosphate unit of β Ser108p and α Lys31 as well as via direct interactions between β Arg83 and the β Ser108p phosphate unit. The interaction network observed in holo $\alpha 2\beta 2$ exhibits a more diverse pattern, likely reflecting the larger flexibility observed in this species relative to holo $\alpha 2\beta 1$. Thus, besides the β Arg82-SC4- α Lys29 triad, which defines the main pathway, an alternate network involves the correlated motion of the pair formed by α Lys29 and the β Ser108p phosphate unit, whereas the loss of the β Arg82- α Asp88 salt bridge primarily reinforces the connection via β Ser108p, which is also assisted by another path where β Arg82 interacts with both β Asp111 and the β Ser108p phosphate group. Compared to the behavior found for A-769662 bound to $\alpha 2\beta 1$ and $\alpha 2\beta 2$ (Fig. 5C and D), these findings provide a basis to understand the mild selectivity of SC4 between $\alpha 2\beta 1$ and $\alpha 2\beta 2$ isoforms.

4. Discussion and conclusion

Unveiling the molecular factors that underlie the structure–function relationships of AMPK isoforms is necessary to rationalize the tissue-dependent expression of AMPK complexes, enabling the design of tailored compounds active against specific metabolic disorders. This question, however, is challenged by the high structural and dynamical behavior observed for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ complexes, although our results have revealed subtle differences in the dynamical response of AMPK to the binding of activators, showing a larger mechanical sensitivity of the $\alpha 2\beta 1$ complex and a larger resilience of the $\alpha 2\beta 2$ species, especially regarding modulation by A-769662.

The different interaction networks observed upon binding of A-769662 to $\alpha 2\beta 1$ supports the effective role of the activator in assisting synergistically the fluctuations of the CBM unit and the P-loop, facilitating the adoption of pre-organized conformations that favor the binding of ATP, shifting the enzyme towards an active state. This effect is weakened in holo $\alpha 2\beta 2$, which results in decreased efficacy of A-769662 in acting as a molecular glue between the CBM unit and the P-loop, making the enzyme less

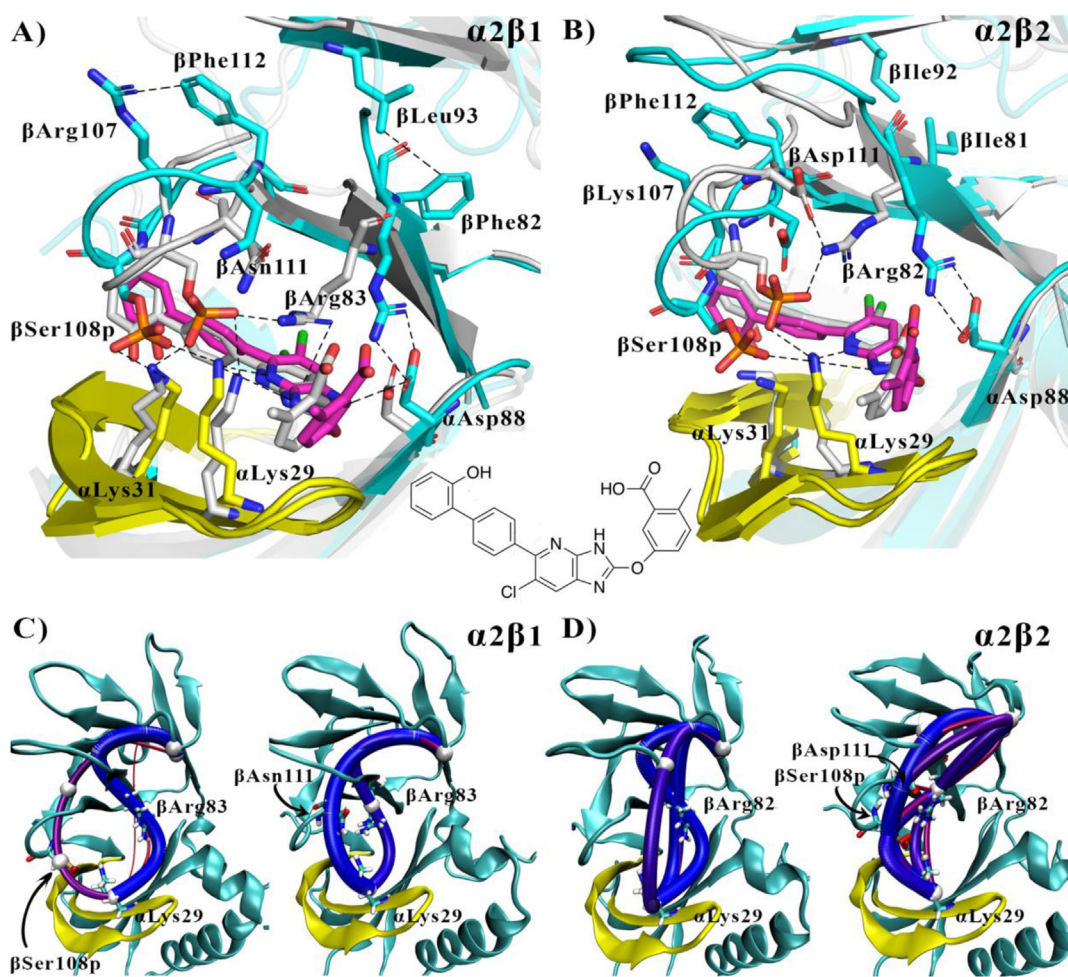


Fig. 6. Representation of selected interactions between the CBM, P-loop and SC4 in holo A) $\alpha 2\beta 1$ and B) $\alpha 2\beta 2$. The distinct conformational arrangements are shown in cyan and grey colors (P-loop in yellow). SC4 is shown as coloured (carbon atoms in orange) sticks in the ADaM site, and a 2D representation of the chemical structure is shown for the sake of clarity. Selected interactions between residues are highlighted as dashed lines. Major interaction networks obtained from WISP analysis for C) $\alpha 2\beta 1$ and D) $\alpha 2\beta 2$ systems (close up views available in Figs. S9 and S10). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

prone to adopt conformations suited to ATP binding. We speculate that this effect might be related, at least in part, to the occurrence of the $\beta 1\text{Asn111} \rightarrow \beta 2\text{Asp111}$ substitution, which would perturb the pattern of interactions formed by the activator upon binding to the ADaM site.

Disentangling the functional relevance of the perturbation caused by the $\beta 1\text{Asn111} \rightarrow \beta 2\text{Asp111}$ substitution is challenging in light of the presence of other substitutions present in $\beta 1$ and $\beta 2$ isoforms, which collectively may dictate the sensitivity of the AMPK complex to direct activators. In particular, changes of residues $\beta 1\text{Phe82}$, $\beta 1\text{Leu93}$ and $\beta 1\text{Leu103}$ to $\beta 2\text{Ile81}$, $\beta 2\text{Ile92}$ and $\beta 2\text{Ile103}$, respectively, which are embedded in a hydrophobic cluster in the CBM unit (Fig. S11), result not only in enlarged fluctuations of the hydrophobic core in holo $\alpha 2\beta 2$, as the RMSF of these residues increases on average from 1.27 ± 0.05 ($\alpha 2\beta 1$) to 1.57 ± 0.14 ($\alpha 2\beta 2$) Å, but also to residue rearrangements, such as the conformational change observed for $\beta 2\text{Tyr125}$ and $\beta 2\text{Phe96}$, leading to local remodeling of the CBM unit (Fig. S12). At this point, it is worth noting that previous experimental studies reported that the double substitution $\beta 1\text{Phe82} \rightarrow \text{Ile}$ and $\beta 1\text{Leu93} \rightarrow \text{Ile}$ sufficed to annihilate the activation of $\alpha 1\beta 1\gamma 1$ by MT47-100 [31]. On the other hand, the linkage between $\beta \text{Ser108p}$ in the CBM unit and residues αLys29 and αLys31 in the P-loop is reinforced by secondary interactions formed by other residues, such as $\beta 1\text{Thr106}$

and $\beta 1\text{Arg107}$, which are changed to $\beta 2\text{Ile106}$ and $\beta 2\text{Lys107}$, respectively. $\beta 1\text{Arg107}$ forms a cation- π interaction with βPhe112 (4.3 ± 0.9 Å), supplemented by a hydrogen bond with the carbonyl oxygen of βGly88 . These interactions are lost in holo $\alpha 2\beta 2$, where $\beta 2\text{Lys107}$ is exposed to the aqueous solvent. Moreover, $\beta 1\text{Thr106}$ is transiently hydrogen bonded to the carbonyl oxygen of αGly9 , favouring the interaction of αArg10 with αGlu34 , which is located at the end of the P-loop.

In the case of SC4, the resemblance of the dual interaction networks that connect the CBM unit and the P-loop via $\beta \text{Ser108p}$ and particularly the $\beta \text{Arg83-SC4-}\alpha \text{Lys29}$ triad provides a basis to understand the mild selectivity between $\beta 1$ - and $\beta 2$ -containing AMPK. This can be attributed to the electrostatic influence of the carboxylate group present in SC4, which would facilitate the adoption of the two arrangements of βArg83 (βArg82), tempering the effect of the $\beta 1\text{Asn111} \rightarrow \beta 2\text{Asp111}$ substitution. Noteworthy, this may explain the slightly larger activation observed for $\beta 1$ -containing AMPK by 991, keeping in mind the similar scaffold of SC4 and 991, particularly regarding the presence of the carboxylate group (Table 1). On the other hand, these findings explain the moderate impact observed experimentally for the mutation $\beta \text{Ser108} \rightarrow \text{Ala}$, which increase the EC_{50} value for SC4 activation by only ~ 4 -fold (determined for $\alpha 1\beta 1\gamma 1$) [29], since this substitution can be counterbalanced by the $\beta \text{Arg83-SC4-}\alpha \text{Lys29}$ network.

Further, this also explains the ~6-fold reduction in the activation potency (determined for $\alpha 2\beta 2\gamma 1$) found upon substitution of the SC4 pyridine-like nitrogen in the imidazo[4,5-b]pyridine ring for carbon [24], since this replacement should annihilate the hydrogen bond formed between SC4 and α Lys29, altering the triad α Lys29-imidazo[4,5-b]pyridine- α Asp88. At this point, let us note that this triad cannot be found in the X-ray structure of the complex between $\alpha 2\beta 2\gamma 1$ and SC4, since the crystallization experiment was performed at pH 6.0, where the imidazo[4,5-b]pyridine was protonated, leading to a different interaction pattern.

In summary, we were able to discern key molecular features that mediate the selective activation of $\alpha 2\beta 1$ and $\alpha 2\beta 2$ isoforms, explaining the different selectivity of A-769662 and SC4. These findings pave the way to understand the 'mechanically sensitive' role of specific residues in the ADaM site that are able to modulate the direct activation mechanism of $\beta 1$ - and $\beta 2$ -containing AMPK complexes. Further studies will be valuable to discern the structural basis of the different sensitivity of AMPK complexes formed by distinct α -subunits to the regulation by direct activators, thus enabling the design of tissue-selective modulators of this important energy sensor.

Funding

We thank the Spanish Ministerio de Economía y Competitividad (SAF2017-88107-R, and Maria de Maetzu MDM-2017-0767, AEI/FEDER), and the Generalitat de Catalunya (2017SGR1746) for financial support and the Barcelona Supercomputing Center (BCV-2014-3-0011 and BCV-2019-1-0009) and the Consorci de Serveis Universitaris de Catalunya (CSUC) for computational resources. E.A. thanks AGAUR (Generalitat of Catalunya; 2018FI-B1-00001) for a fellowship.

CRedit authorship contribution statement

Elnaz Aledavood: Formal analysis, Investigation, Visualization, Writing - original draft. **Alessia Forte:** Formal analysis, Investigation, Visualization, Writing - original draft. **Carolina Estarellas:** Conceptualization, Methodology, Investigation, Supervision, Writing - review & editing. **F. Javier Luque:** Conceptualization, Methodology, Investigation, Supervision, Writing - review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2021.05.056>.

References

- [1] Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, et al. Towards a knowledge-based human protein atlas. *Nat Biotech* 2010;28:1248–50.
- [2] Uhlen M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A. Tissue-based map of the human proteome. *Science* 2015;347:1260419.
- [3] Yang X, Coulombe-Huntington J, Kang S, Sheynkman GM, Hao T, Richardson A, et al. Widespread expansion of protein interaction capabilities by alternative splicing. *Cell* 2016;164:805–17.
- [4] Eksi R, Li H-D, Menon R, Wen Y, Omenn GS, Ketzler M, et al. Systematically Differentiating Functions for Alternatively Spliced Isoforms Through Integrating RNA-seq Data. *PLoS Comput Biol* 2013;9:e1003314.

- [5] Sulakhe D, S'Souza M, Wang S, Balasubramanian S, Athri P, Xie B, et al. Exploring the functional impact of alternative splicing of human protein isoforms using available annotation sources. *Brief Bioinform* 2019;20:1754–68.
- [6] Menon R, Roy A, Mukherjee S, Belkin S, Zhang Y, Omenn GS. Functional implications of structural predictions for alternative splice proteins expressed in her2/neu-induced breast cancers. *J Proteome Res* 2011;10:5503–11.
- [7] Chen H, Shaw D, Zeng J, Bu D, Jiang T. DIFFUSE: Predicting Isoform Functions From Sequences and Expression Profiles Via Deep Learning. *Bioinformatics* 2019;35:284–94.
- [8] Thornton C, Snowden MA, Carling D. Identification of a Novel AMP-Activated Protein Kinase Beta Subunit Isoform That Is Highly Expressed in Skeletal Muscle. *J Biol Chem* 1998;273:12443–50.
- [9] Cheung CF, Salt IP, Davies A, Hardie DG, Carling D. Characterization of AMP-activated protein kinase gamma subunit isoforms and their role in AMP binding. *Biochem J* 2000;346:659–69.
- [10] Mahlapuu M, Johansson C, Lindgren K, Hjälml G, Barnes BR, Krook A, et al. Expression Profiling of the γ -Subunits Isoforms of AMP-Activated Protein Kinase Suggests a Major Role for $\gamma 3$ in White Skeletal Muscle. *Am J Physiol Endocrinol Metab* 2004;286:E194–200.
- [11] Carling D, Mayer FV, Sanders MJ, Gamblin SJ. AMP-activated protein kinase: nature's energy sensor. *Nat Chem Biol* 2011;7:512–8.
- [12] Herzig S, Shaw RJ. AMPK: guardian of metabolism and mitochondrial homeostasis. *Nat Rev Mol Cell Biol* 2018;19:121–35.
- [13] Day EA, Ford RJ, Steinberg GR. AMPK as a therapeutic target for treating metabolic diseases. *Trends Endocrinol Metab* 2017;28:545–60.
- [14] Luo Z, Saha AK, Xiang X, Ruderman NB. AMPK, the Metabolic Syndrome and Cancer. *Trends Pharmacol Sci* 2005;26:69–76.
- [15] Hardie DG. AMP-activated protein kinase as a drug target. *Annu Rev Pharmacol Toxicol* 2007;47:185–210.
- [16] Carling D. AMPK signalling in health and disease. *Curr Opin Cell Biol* 2017;45:31–7.
- [17] Xiao B, Sanders MJ, Underwood E, Heath R, Carmena D, Jing C, et al. Structure of mammalian AMPK and its regulation by ADP. *Nature* 2011;472:230–3.
- [18] Calabrese MF, Rajamohan F, Harris MS, Caspers NL, Magyar R, Withka JM, et al. Structural basis for AMPK activation: natural and synthetic ligands regulate kinase activity from opposite poles by different molecular mechanisms. *Structure* 2014;22:1161–72.
- [19] Carling D, Thornton C, Woods A, Sanders MJ. AMP-activated protein kinase: new regulation, new roles?. *Biochem J* 2012;445:11–27.
- [20] Chen L, Wang J, Zhang YY, Yan SF, Neumann D, Schlattner U, et al. AMP-activated protein kinase undergoes nucleotide-dependent conformational changes. *Nat Struct Mol Biol* 2012;19:716–8.
- [21] Willows R, Sanders MJ, Xiao B, Patel BR, Martin SR, Read J, et al. Phosphorylation of AMPK by upstream kinases is required for activity in mammalian cells. *Biochem J* 2017;474:3059–73.
- [22] Vazquez-Martin A, Vellon L, Quirós PM, Cufi S, De Galarreta ER, Ferraros CO, et al. Activation of AMP-activated protein kinase (AMPK) provides a metabolic barrier to reprogramming somatic cells into stem cells. *Cell Cycle* 2012;11:974–89.
- [23] Cool B, Zinker B, Chiou W, Kifle L, Cao N, Perham M, et al. Identification and Characterization of a Small Molecule AMPK Activator That Treats Key Components of Type 2 Diabetes and the Metabolic Syndrome. *Cell Metab* 2006;3:403–16.
- [24] Xiao B, Sanders MJ, Carmena D, Bright NJ, Haire LF, Underwood E, et al. Structural Basis of AMPK Regulation by Small Molecule Activators. *Nat Commun* 2013;4:1–10.
- [25] Langendorf CG, Kemp BE. Choreography of AMPK activation. *Cell Res* 2015;25:5–6.
- [26] Hardie DG. AMPK - Sensing Energy While Talking to Other Signaling Pathways. *Cell Metab* 2014;20:939–52.
- [27] Olivier S, Foretz M, Viollet B. Promise and Challenges for Direct Small Molecule AMPK Activators. *Biochem Pharmacol* 2018;153:147–58.
- [28] Willows R, Navaratnam N, Lima A, Read J, Carling D. Effect of different γ -subunit isoforms on the regulation of AMPK. *Biochem J* 2017;474:1741–54.
- [29] Ngoei KRW, Langendorf CG, Ling NXY, Hoque A, Varghese S, Camerino MA, et al. Structural Determinants for Small-Molecule Activation of Skeletal Muscle AMPK $\alpha 2\beta 2\gamma 1$ by the Glucose Importagoc SC4. *Cell Chem Biol* 2018;25:728–37.
- [30] Cokorinos EC, Delmore J, Reyes AR, Albuquerque B, Kjøbsted R, Jørgensen NO, et al. Activation of Skeletal Muscle AMPK Promotes Glucose Disposal and Glucose Lowering in Non-Human Primates and Mice. *Cell Metab* 2017;25:1147–59.
- [31] Scott JW, Galic S, Graham KL, Foitzik R, Ling NXY, Dite TA, et al. Inhibition of AMP-Activated Protein Kinase at the Allosteric Drug-Binding Site Promotes Islet Insulin Release. *Chem Biol* 2015;22:705–11.
- [32] Aledavood E, Moraes G, Lameira J, Castro A, Luque FJ, Estarellas C. Understanding the Mechanism of Direct Activation of AMP-Kinase: Toward a Fine Allosteric Tuning of the Kinase Activity. *J Chem Inf Model* 2019;59:2859–70.
- [33] Case DA, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham TE, Cruzeiro VWD. AMBER 2018. San Francisco: University of California; 2018.
- [34] Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins* 2010;78:1950–8.

- [35] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. *Development and testing of a general AMBER force field*. *J Comput Chem* 2004;25:1157–74.
- [36] Bayly CI, Cieplak P, Cornell W, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* 1993;97:10269–80.
- [37] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;79:926–35.
- [38] Joung IS, Cheatham TE. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* 2008;112:9020–41.
- [39] Amadei A, Linssen ABM, Berendsen HJC. Essential dynamics of proteins. *Proteins* 1993;17:412–25.
- [40] Shkurti A, Goni R, Andrio P, Breitmoser E, Bethune I, Orozco M, et al. PyPczip: A PCA-based toolkit for compression and analysis of molecular simulation data. *SoftwareX* 2015;5:44–50.
- [41] Schlitter J. Estimation of Absolute And Relative Entropies Of Macromolecules Using The Covariance Matrix. *Chem Phys Lett* 1993;215:617–21.
- [42] Baron R, Hünenberger PH, McCammon JA. Absolute single-molecule entropies from quasi-harmonic analysis of microsecond molecular dynamics: correction terms and convergence properties. *J Chem Theory Comput* 2009;5:3150–60.
- [43] Harris SA, Gavathiotis E, Searle MS, Orozco M, Lughton CA. Cooperativity in drug-DNA recognition: a molecular dynamics study. *J Am Chem Soc* 2001;123:12658–63.
- [44] Hunenberger P, Mark A, van Gunsteren W. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *J Mol Biol* 1995;252:492–503.
- [45] Bajusz D, Rácz A, Héberger K. Why is tanimoto index an appropriate choice for fingerprint-based similarity calculations?. *J. Cheminform.* 2015;7:20.
- [46] O'Rourke KF, Gorman SD, Boehr DD. Biophysical and computational methods to analyze amino acid interaction networks in proteins. *Comput Struct Biotechnol J* 2016;14:245–51.
- [47] Van Wart AT, Durrant J, Votapka L, Amaro RE. Weighted implementation of suboptimal paths (WISP): an optimized algorithm and tool for dynamical network analysis. *J Chem Theory Comput* 2014;10:511–7.
- [48] Dusan P, Lynn KSC. Molecular modeling of conformational dynamics and its role in enzyme evolution. *Curr Opin Struct Biol* 2018;52:50–7.
- [49] Orellana L. Large-scale conformational changes and protein function: breaking the in silico barrier. *Front. Mol. Biosci.* 2019;6:117.
- [50] Waheed SO, Ramanan R, Chaturvedi SS, Lehnert N, Schofield CJ, Christov CZ, et al. Role of Structural Dynamics in Selectivity and Mechanism of Non-Heme Fe(II) and 2-oxoglutarate-Dependent Oxygenases Involved in DNA Repair. *ACS Cent Sci* 2020;6:795–814.

Supplementary Material

Structural Basis of the Selective Activation of Enzyme Isoforms: Allosteric Response of β 1- and β 2-Containing AMPK Complexes

Elnaz Aledavood,[†] Alessia Forte,[†] Carolina Estarellas^{*,†} and F. Javier Luque^{*,†}

[†]Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicine (IBUB) and Institute of Theoretical and Computational Chemistry (IQTUCB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain

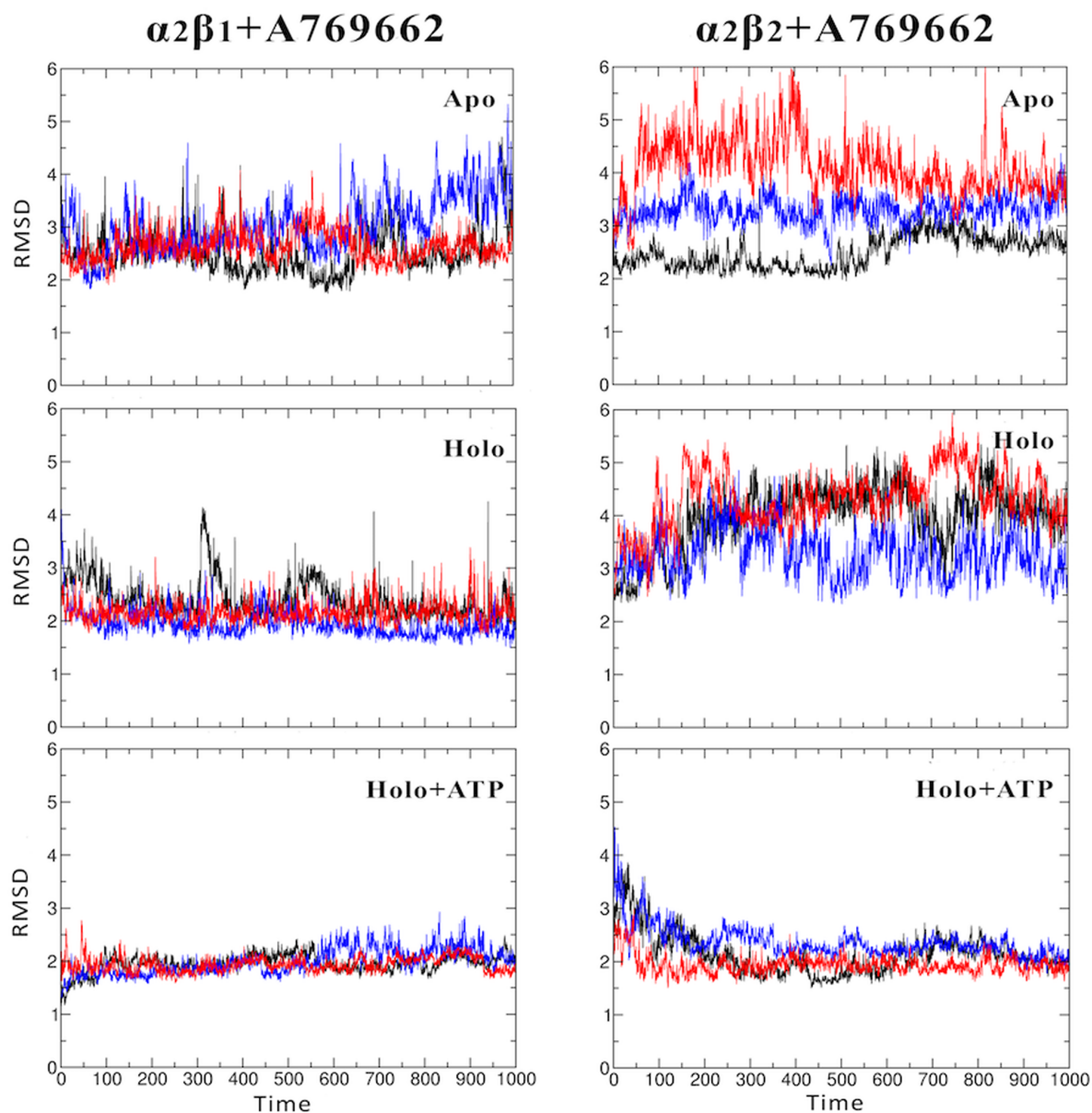


Fig. S1. Root mean squared deviation (RMSD, Å) determined for the protein backbone along the three 1 μ s MD simulations run for the apo, holo and holo+ATP species of AMPK isoforms $\alpha 2\beta 1$ and $\alpha 2\beta 2$ bound to A-769662 (each replica is shown in black, red and blue). For each isoform, the energy-minimized average structure of the holo+ATP sampled in the last 200 ns of the three independent MD simulations was used as reference structure.

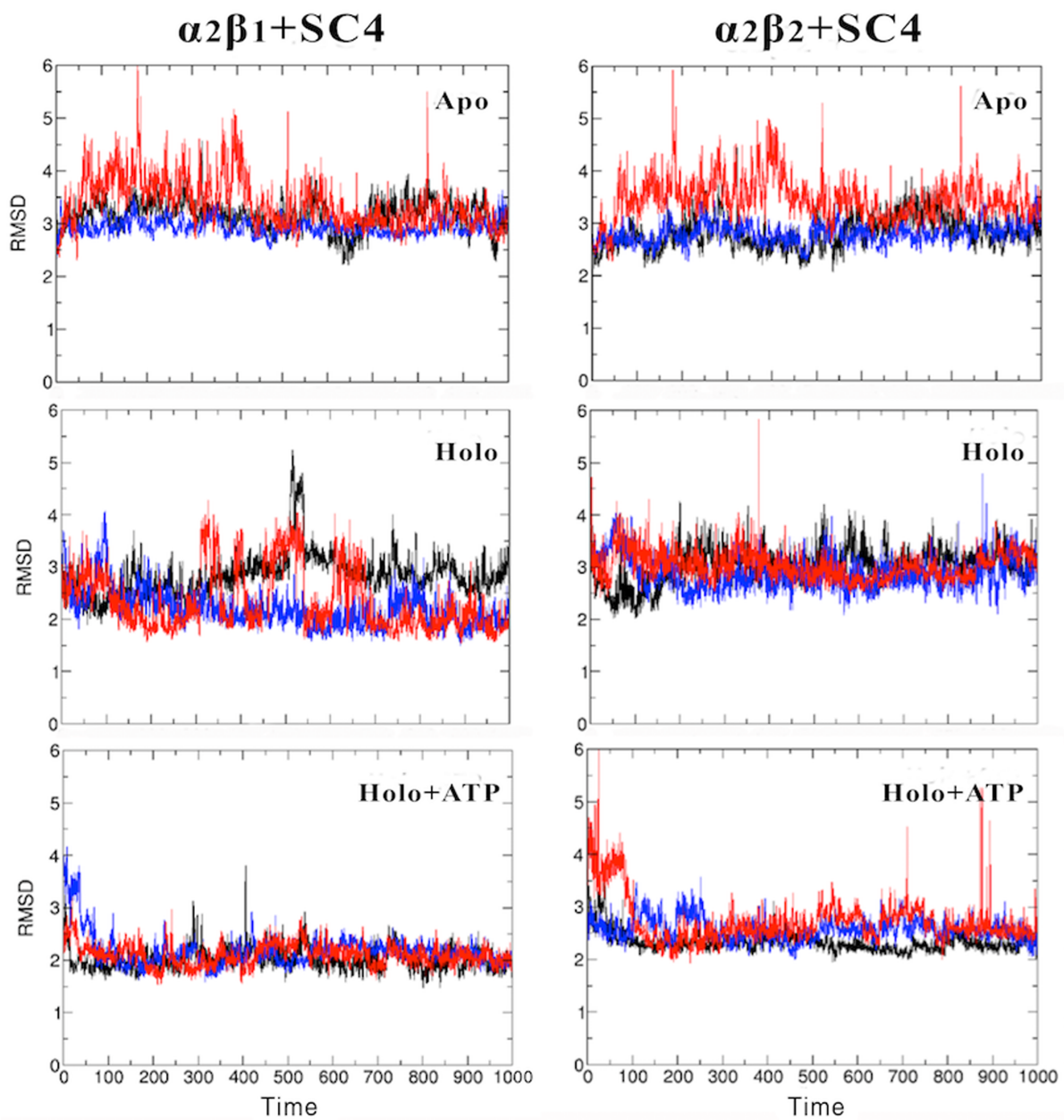


Fig. S2. RMSD (Å) determined for the protein backbone along the three 1 μ s MD simulations run for the apo, holo and holo+ATP species of AMPK isoforms $\alpha 2\beta 1$ and $\alpha 2\beta 2$ bound to SC4 (each replica is shown in black, red and blue). For each isoform, the energy-minimized average structure of the holo+ATP sampled in the last 200 ns of the three independent MD simulations was used as reference structure.

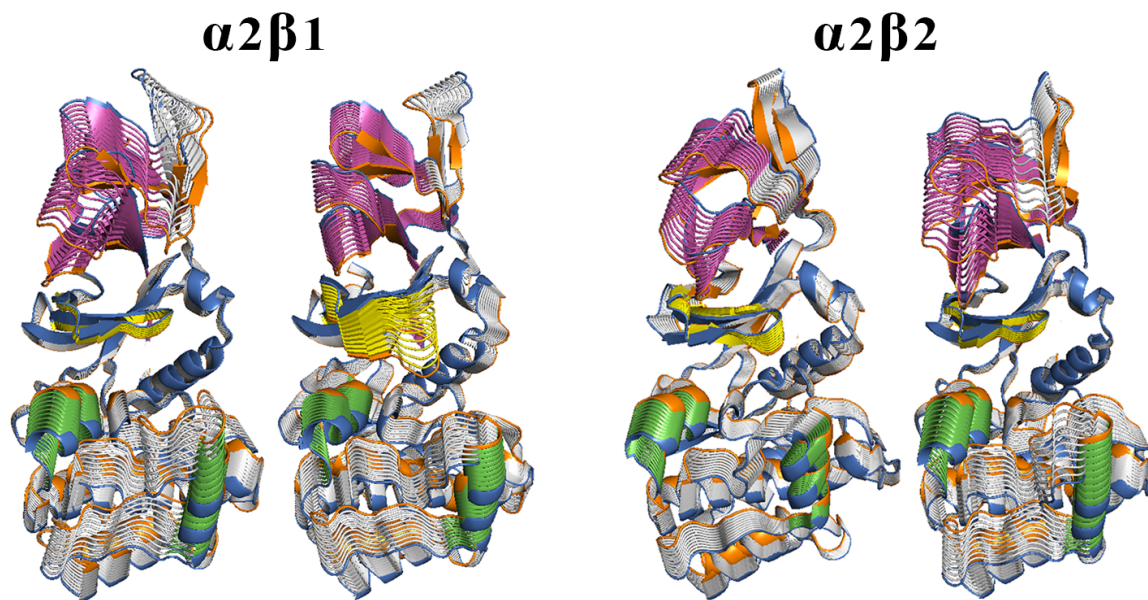


Fig. S3. First essential motion obtained for two additional replicas run for apo species of $\alpha 2\beta 1$ and $\alpha 2\beta 2$ AMPK isoforms. For the sake of clarity, the P-loop is shown in yellow, the helices formed by residues 100-110 and 220-229 in the α -subunit in green (see legend to Figure 2), and β -sheets of the CBM of the β -subunit in magenta.

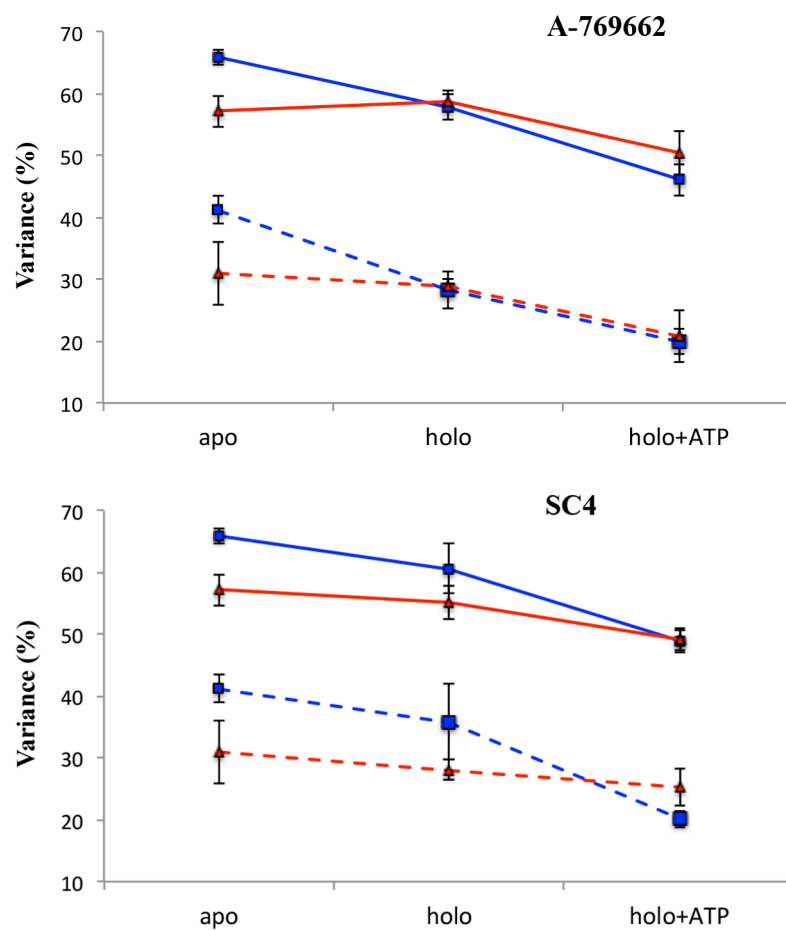


Fig. S4. Contribution to the structural variance (%) of the first essential motion (dashed line) and the sum of the first three essential motions (solid line) of the protein backbone for apo, holo and holo+ATP species of AMPK. Values determined for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ isoforms are shown in blue and red, respectively. Values averaged for the three replicas run for each system (bars denote the standard deviation).

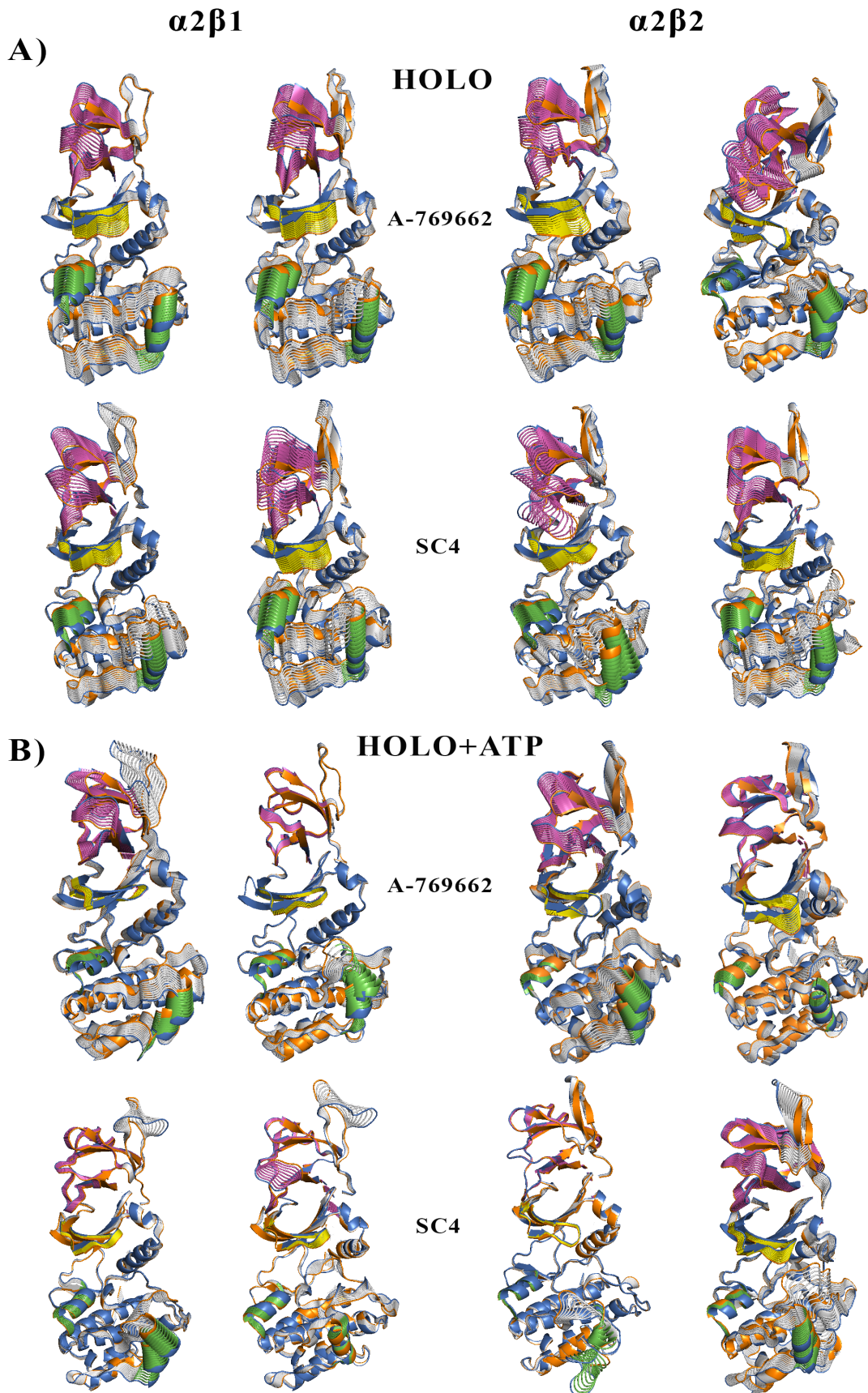


Fig. S5. Representation of the first essential motion obtained for two additional replicas run for holo and holo+ATP species of A) A-769662 and B) SC4 bound to $\alpha 2\beta 1$ and $\alpha 2\beta 2$ isoforms of AMPK. For the sake of clarity, the P-loop is shown in yellow, the helices formed by residues 100-110 and 220-229 in the α -subunit in green (see legend to Figure 2), and β -sheets in the CBM domain of the β -subunit in magenta.

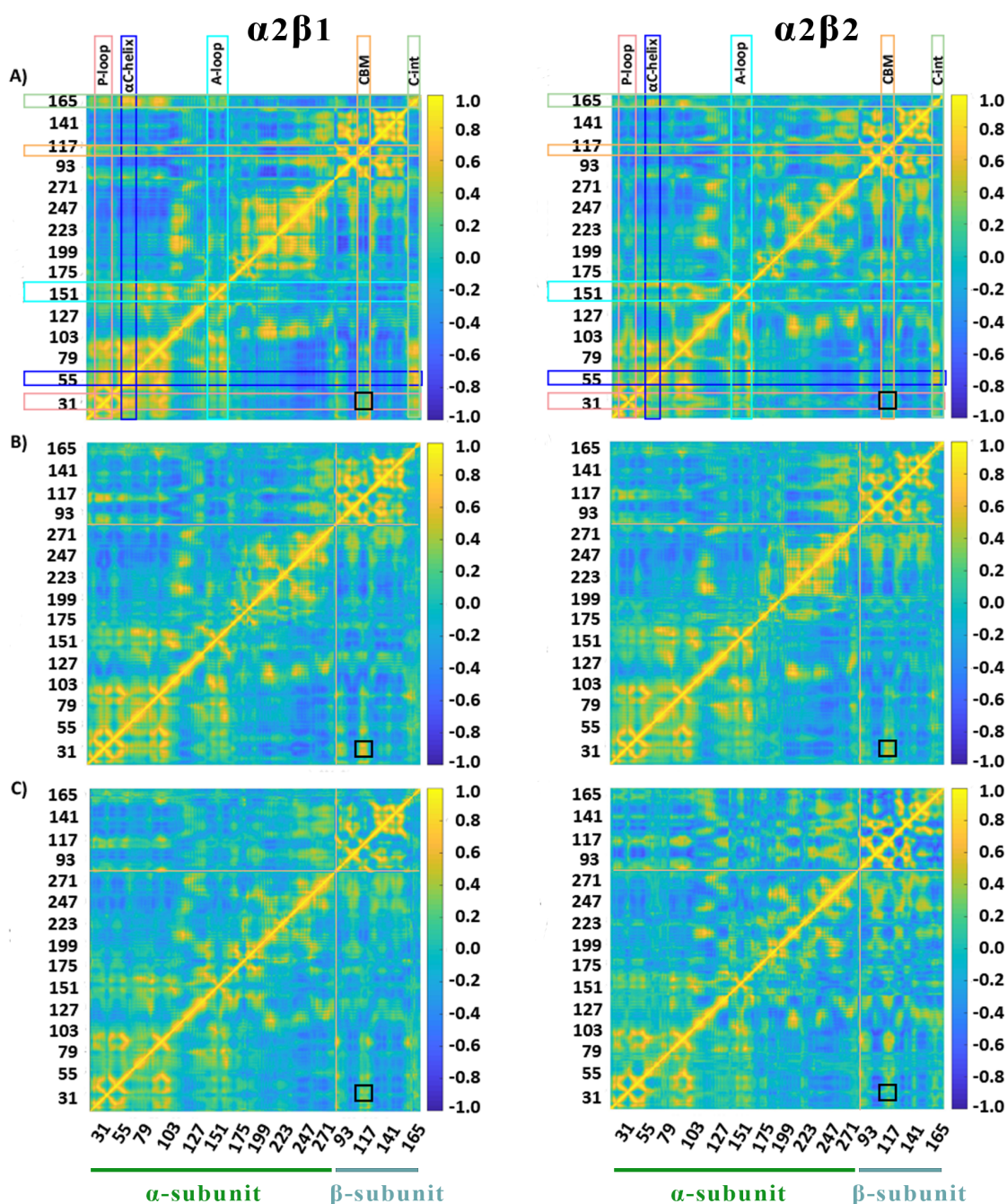


Fig. S6. Dynamic cross-correlation (DCC) matrices for A) apo, B) holo and C) holo+ATP complexes of (left) $\alpha 2\beta 1$ and (right) $\alpha 2\beta 2$ with SC4. The x- and y-axis denote the numbering of residue in the α - and β -subunit. The residues pertaining to the α - and β -subunits are indicated at the bottom of the plot (green and blue lanes for α - and β -subunits, respectively). Specific regions of AMPK are highlighted with squares at the top of the plot: P-loop (pink), α C-helix (blue), activation loop (A-loop; cyan), CBM unit (orange) and C-interacting helix (C-int; green). Regions colored in yellow/blue show

high correlated/anticorrelated fluctuations. The black square in the plots highlights the motion between the CBM domain and the P-loop.

A)

```

CBM
β1  76  QARPTVFRWTGGGKEVYLSGSFNNWS-KLPLTRSHNNFVAILDLPE 120
β2  75  QARPTVIRWSEGGKEVFIISGSFNNWSTKIPLIKSHNDFVAILDLPE 120
β1  121  GEHQYKFFVDGQWTHDPSEPIVTSQLGTVNNIIQV 155
β2  121  GEHQYKFFVDGQWVHDPSEPVVTSQLGTINNLIHV 155

Linker region
β1  156  KKTDFE 161
β2  156  KKSDFE 161

C-interacting helix
β1  162  VFDALMVDSQKCSDVS 177
β2  162  VFDALKLDSMESSETS 177

```

B)

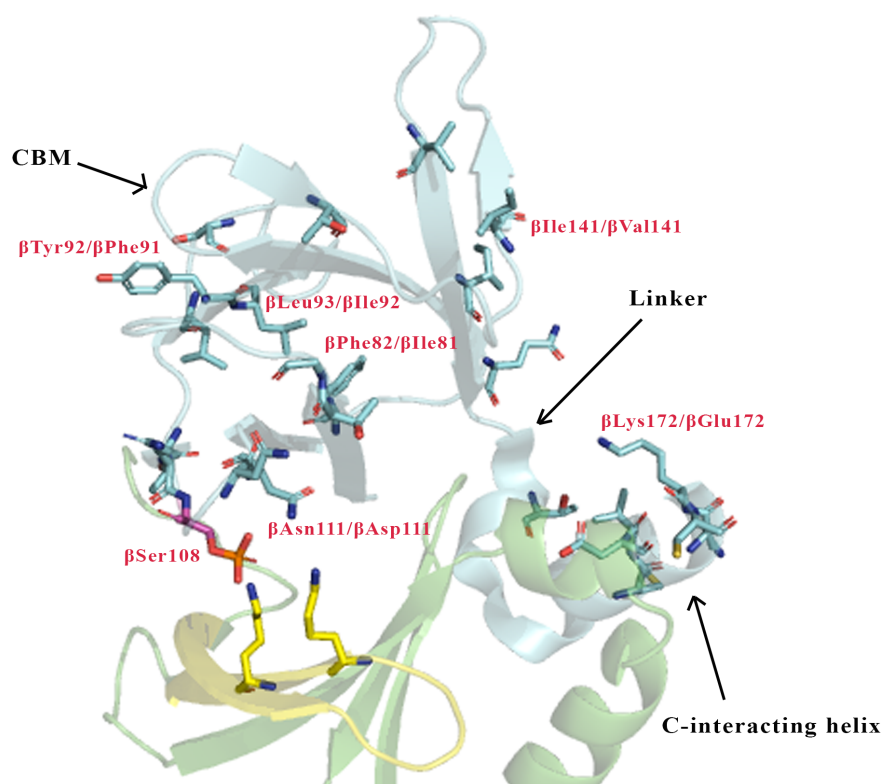


Fig. S7. A) Sequence alignment of regions corresponding to the CBM unit (residues 73-155) and C-interacting helix (residues 162-177) in human AMPK β 1- and β 2-isoforms. Non-conserved residues are displayed in blue. The phosphorylatable Ser108 is highlighted in magenta. B) Graphical display of mutated residues between β 1- and β 2-isoforms (shown as blue-colored cartoon). Residues are highlighted as sticks using the X-ray structure of α 2 β 1 γ 1 AMPK (PDB entry 4CFF). The protein backbone of the α - and β -subunits domain is shown as green and cyan cartoon, respectively. The P-loop is shown in yellow. The main residues involved in activation mechanism of different AMPK isoforms are labelled in red.

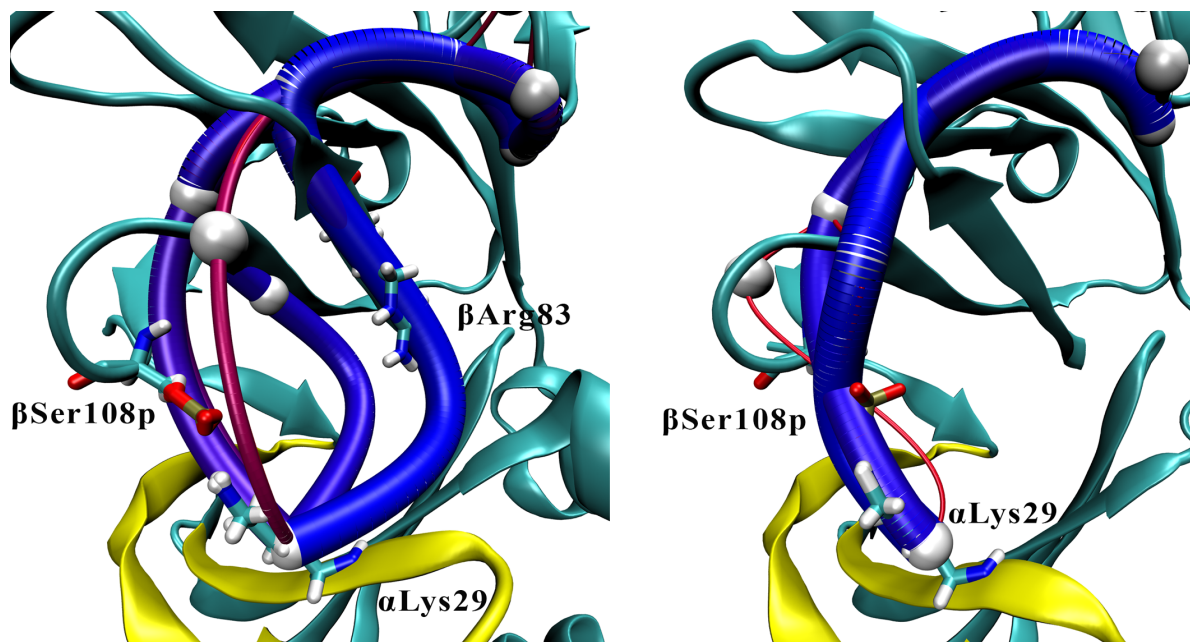


Fig. S8. Close up view of the major interaction networks obtained from WISP analysis for the holo systems of A-769662 bound to (left) $\alpha 2\beta 1$ and (right) $\alpha 2\beta 2$ (see also Figures 5C and 5D).

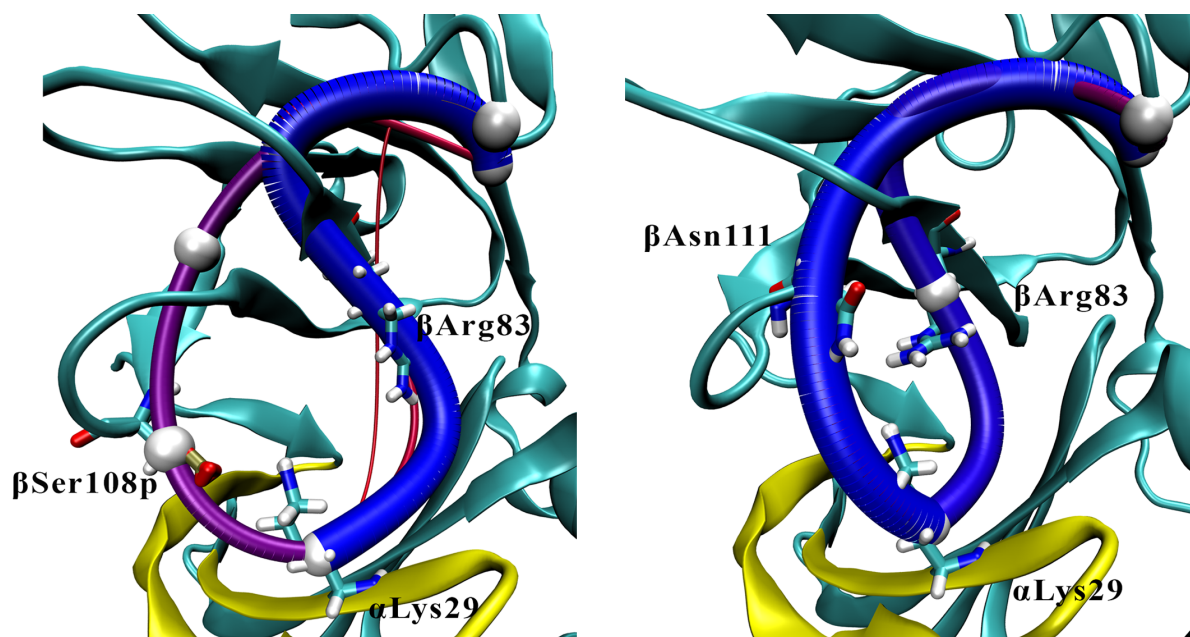


Fig. S9. Close up view of the two interaction networks obtained from WISP analysis for the holo systems of SC4 bound to $\alpha 2\beta 1$ (see also Figure 6C).

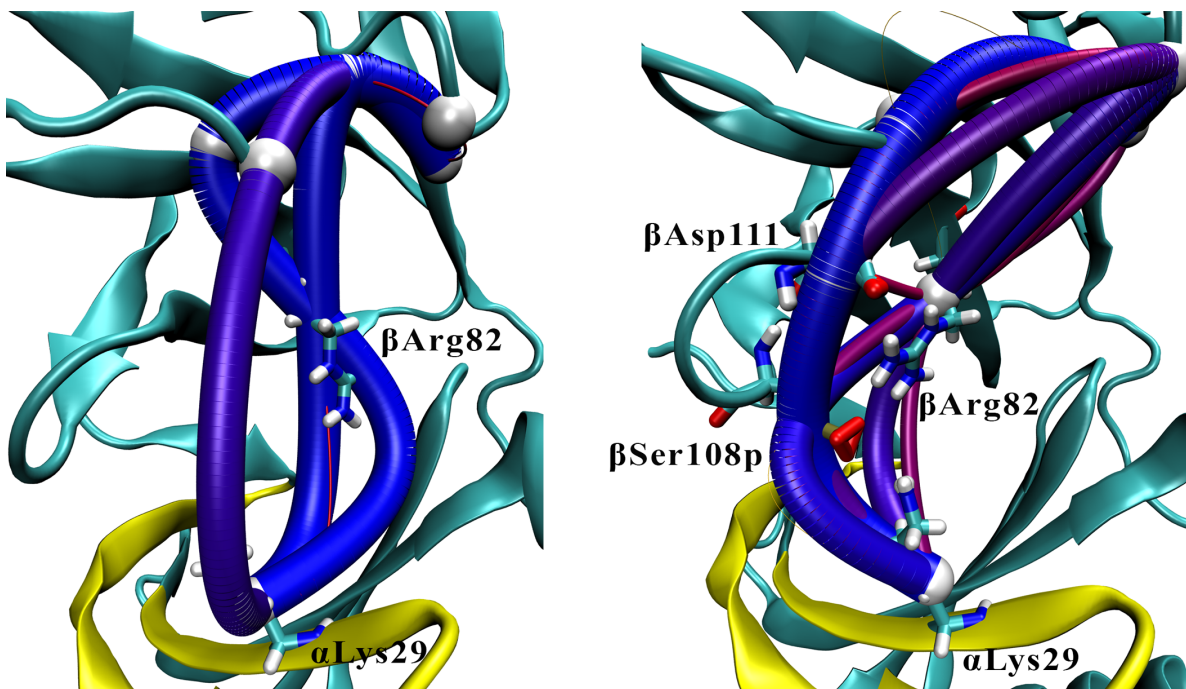


Fig. S10. Close up view of the two interaction networks obtained from WISP analysis for the holo systems of SC4 bound to α 2 β 2 (see also Figure 6D).

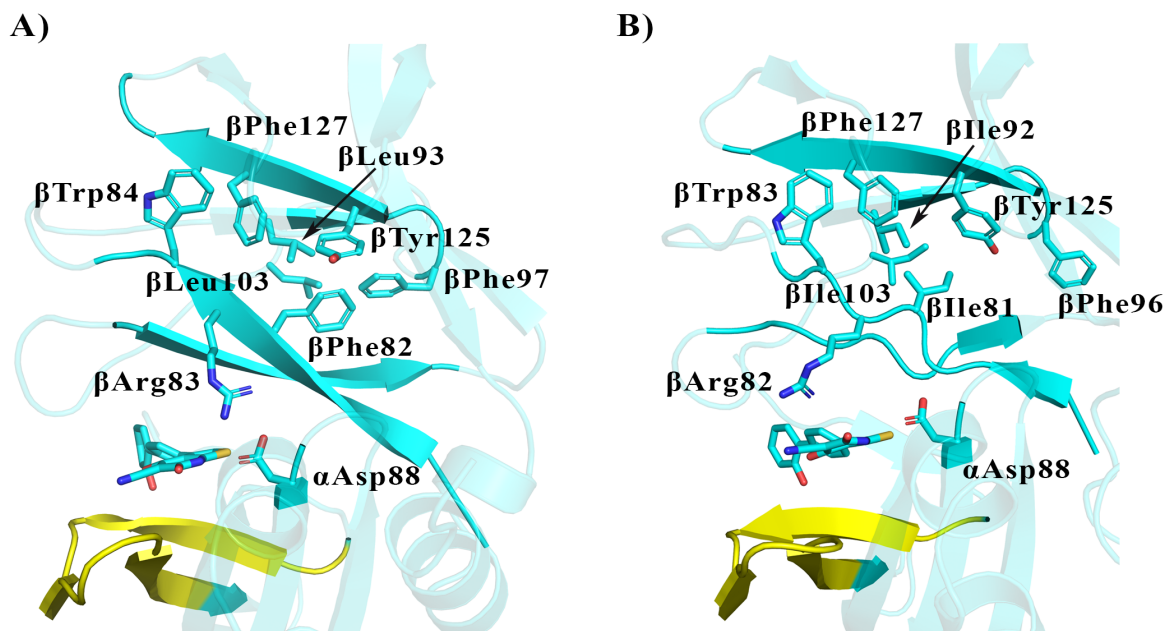


Fig. S11. Hydrophobic cluster formed by aromatic and aliphatic residues in the CBM unit of the β -subunit for A) $\alpha 2\beta 1$ and B) $\alpha 2\beta 2$ isoforms. Note the different arrangement of residues $\beta 1$ Tyr125- $\beta 1$ Phe97 in $\alpha 2\beta 1$ and $\beta 2$ Tyr125- $\beta 2$ Phe96 in $\alpha 2\beta 2$, which alters the packing of the CBM unit.

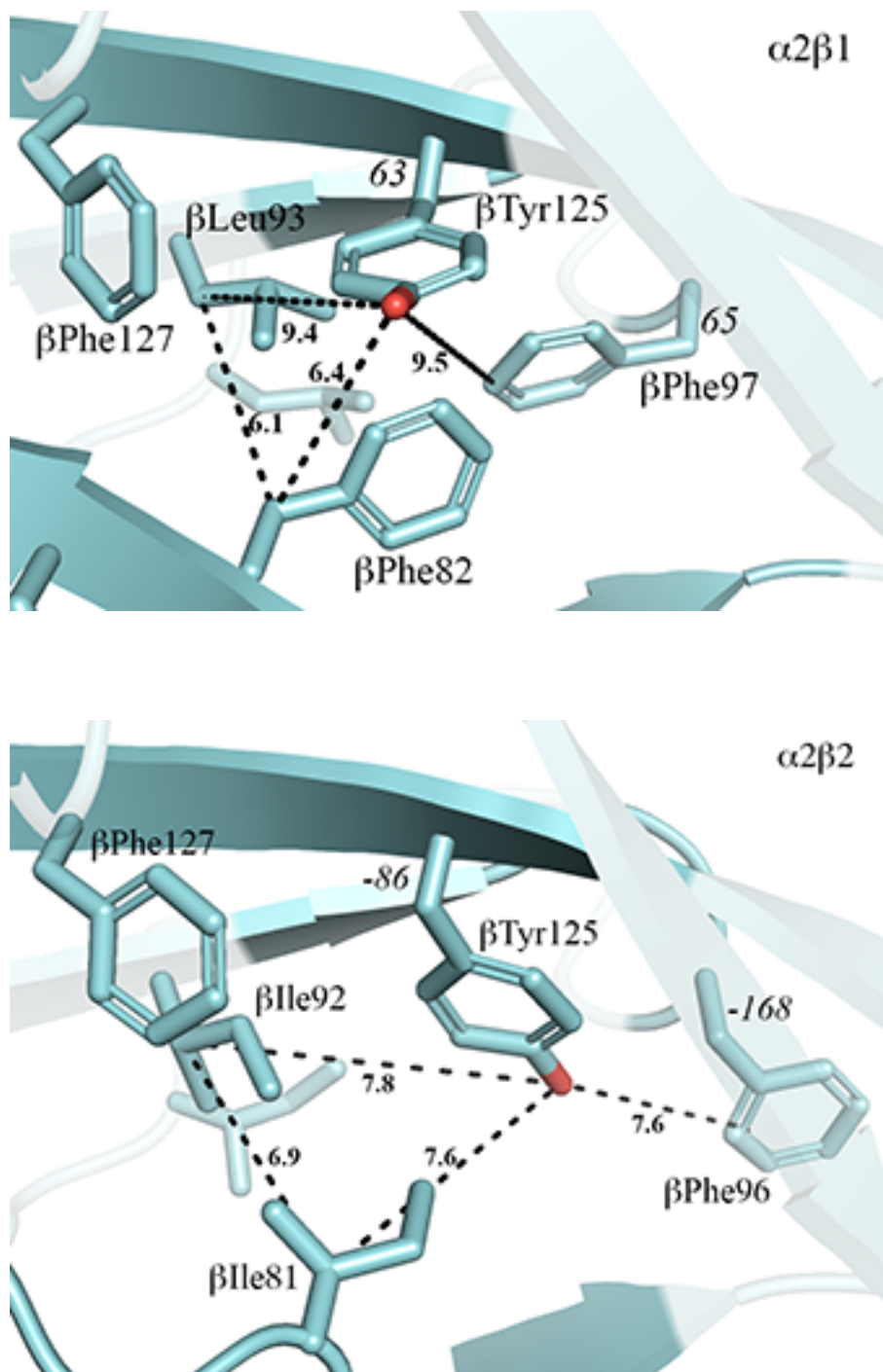


Fig. S12. Close up views of the different arrangement of residues (top) β 1Tyr125 and β 1Phe97 in $\alpha 2\beta 1$ and (bottom) β 2Tyr125 and β 2Phe96 in $\alpha 2\beta 2$ in the CBM unit. Dashed lines highlight selected distances (values in Å). The torsional angle N-C α -C β -C γ of these residues is (in degrees) is shown in italics.

Table S1. Description of the systems used in Molecular Dynamics simulations.

$\alpha 2\beta 1$	Protein	Water	Na+/Cl-	Atoms
Apo	$\alpha 2$: residues 8-278	25,037	--- / 2	81,030
Holo (A-769662)	4388 atoms	24,957	--- / ---	80,829
Holo+ATP (A-769662)	$\beta 1$: residues 78-173	27,830	4 / ---	89,495
Holo (SC4)	1526 atoms	26,626	1 / ---	85,850
Holo+ATP (SC4)		26,620	5 / ---	85,879
$\alpha 2\beta 2$				
Apo	$\alpha 2$: residues 8-278	25,768	--- / ---	83,213
Holo (A-769662)	4388 atoms	25,733	2 / ---	83,151
Holo+ATP (A-769662)	$\beta 2$: residues 77-171	25,729	6 / ---	83,186
Holo (SC4)	1518 atoms	25,728	3 / ---	83,150
Holo+ATP (SC4)		25,710	7 / ---	83,143

Table S2. Similarity indexes determined for the three independent simulations run for apo, holo and holo+ATP species of $\alpha 2\beta 1$ and $\alpha 2\beta 2$ isoforms upon binding with activator (A-769662 and SC4) and ATP. Values stand for the average similarity measured between each pair of replicas (standard deviation is given in parenthesis), and was determined for the whole protein backbone as well as for the correlation between residues within α - and β -subunits.

Species	$\alpha 2\beta 1$				$\alpha 2\beta 2$			
	Global	α - α	β - β	α - β	Global	α - α	β - β	α - β
apo								
---	0.84 (0.02)	0.83 (0.02)	0.94 (0.02)	0.82 (0.02)	0.79 (0.06)	0.79 (0.08)	0.86 (0.03)	0.75 (0.05)
holo								
A-769662	0.91 (0.02)	0.92 (0.01)	0.94 (0.02)	0.89 (0.03)	0.75 (0.07)	0.80 (0.04)	0.83 (0.06)	0.59 (0.10)
SC4	0.84 (0.02)	0.84 (0.03)	0.89 (0.02)	0.83 (0.01)	0.75 (0.03)	0.79 (0.03)	0.76 (0.08)	0.66 (0.05)
holo+ATP								
A-769662	0.81 (0.01)	0.85 (0.01)	0.86 (0.05)	0.73 (0.03)	0.76 (0.02)	0.79 (0.03)	0.80 (0.03)	0.69 (0.04)
SC4	0.66 (0.10)	0.70 (0.12)	0.75 (0.03)	0.52 (0.10)	0.60 (0.05)	0.64 (0.06)	0.69 (0.14)	0.47 (0.06)

3.1.3. Paper 3: “Elucidating the activation mechanism of AMPK by direct pan-activator PF-739”

Manuscript submitted

Elnaz Aledavood,¹ Aria Gheeraert,² Ivan Rivalta,³ F. Javier Luque,¹ Carolina Estarellas,^{1,*}

¹Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicina (IBUB) and Institute of Theoretical and Computational Chemistry (IQTUCB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain

²Laboratory of Mathematics (LAMA), CNRS, University of Savoie Mont Blanc, France

³Dipartimento di Chimica Industriale “Toso Montanari” Università di Bologna, Bologna, Italy

Elucidating the activation mechanism of AMPK by direct pan-activator PF-739

1 Elnaz Aledavood¹, Aria Gheeraert², Ivan Rivalta³, F. Javier Luque¹, Carolina Estarellas^{1,*}

2 ¹Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences,
3 and Institute of Theoretical and Computational Chemistry (IQTCUB), University of Barcelona, Av.
4 Prat de la Riba 171, Santa Coloma de Gramenet 08921, Spain

5 ²Laboratory of Mathematics (LAMA), CNRS, University of Savoie Mont Blanc, France

6 ³Dipartimento di Chimica Industriale “Toso Montanari” Università di Bologna, Bologna, Italy

7 * **Correspondence:**

8 Corresponding Author

9 cestarellas@ub.edu

10 **Keywords: protein dynamics, AMPK, molecular dynamic simulations, activation mechanism,**
11 **pan-activator, enzyme isoforms.**

12 Abstract

13 AMP-activated protein kinase (AMPK) is a key energy sensor regulating the cell metabolism in
14 response to energy supply and demand. The evolutionary adaptation of AMPK to different tissues is
15 accomplished through the expression of distinct isoforms that can form up to 12 complexes, which
16 exhibit notable differences in the sensitivity to direct activators. To comprehend the molecular factors
17 of the activation mechanism of AMPK, we have assessed the structural and dynamical properties of
18 β 1- and β 2-containing AMPK complexes towards the binding of a pan-activator PF-739. The
19 analysis let us to understand how the PF-739 can activate the AMPK through both β -isoforms and let
20 us to discern the slightly higher affinity upon β 1-isoform, due to the β 1-Asn111 to β 2-Asp111
21 substitution as a critical factor in modulating the dynamical sensitivity upon the β -isoforms. Our
22 work reveals which are the crucial structural factors that could help in the design of selective
23 activators to improve the therapeutic treatment focused on tissue specificity.

24 1 Introduction

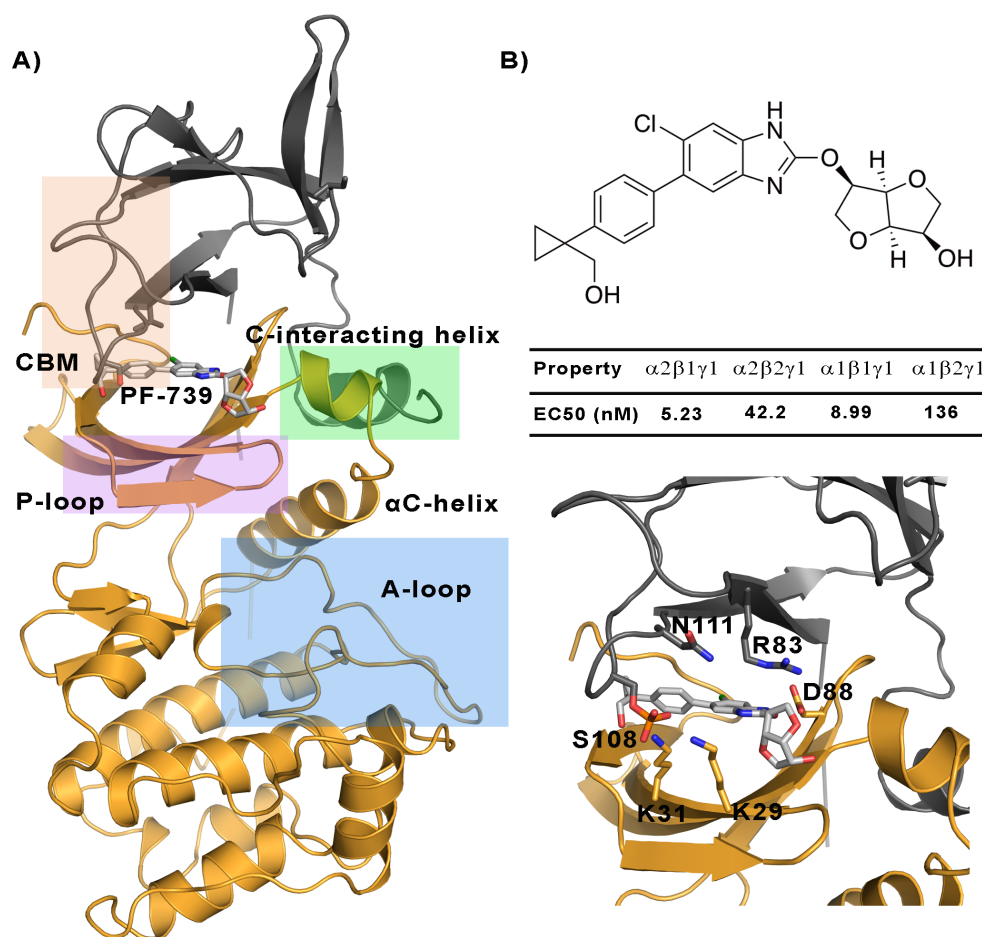
25 AMP-activated protein kinase (AMPK) is a Ser/Thr protein kinase with a key role as a sensor in
26 cellular energy homeostasis [1]. Once activated increase the levels of ATP, favoring the reduction of
27 anabolic pathways and up-regulation of catabolic pathways. Due to its critical role in cell
28 metabolism, AMPK is implicated in numerous metabolic disorders such as type 2 diabetes,
29 cardiovascular diseases, and obesity [2]. However, one of the most interesting aspects of this enzyme
30 comes from the different tissue distribution that is directly related to its structural complexity. AMPK
31 is a heterotrimeric complex consisting of a catalytic α -subunit and two regulatory subunits, namely β
32 and γ . Each subunit is encoded by multiple genes, involving two for α (α 1, α 2), two for β (β 1, β 2),
33 and three for γ (γ 1, γ 2, γ 3) isoforms [3,4]. The N-terminus of the α catalytic subunit contains a kinase
34 domain, while its C-terminus is needed for the formation of the complex with the other subunits. The
35 β -subunit has a central carbohydrate-binding module (CBM) that mediates AMPK interaction with

36 glycogen, and the C-terminal region acts as a scaffold for the heterotrimeric assembly. Finally, the γ -
37 subunit has four tandem repeats of the cystathionine β -synthase (CBS) domain, able to bind up to
38 four adenine nucleotides. AMPK is finely regulated by different mechanisms [5]. The allosteric
39 activation involves the phosphorylation of α 2-Thr172 in the activation loop of the kinase domain by
40 upstream kinases such as LKB1 and CaMKKb, together with the binding of AMP to one of the sites
41 of the BBS domain in the γ -subunit. This active AMPK complex could, thus, respond to subtle
42 fluctuations in the AMP/ATP ratio, and is several thousand-fold more active [6,7,8]. On the other
43 side, AMPK can also be indirectly activated by compounds such as metformin, phenformin and
44 oligomycin [9], which are able to increase the intracellular levels of AMP. However, we are currently
45 more interested in the understanding of the direct activation mechanism of AMPK, which can be
46 performed by small organic molecules. The first direct activator describe was the thienopyridone
47 drug A-769662 [10], which is bound to a cavity located at the interface between the CBM domain of
48 the β -subunit and the kinase domain of the α -subunit, called allosteric drug and metabolite (ADaM)
49 site [11,12]. One of the main features of the direct activation is that this kind of activation is
50 independent of the Thr172 phosphorylation, while it has been observed that the Ser108 located at the
51 CBM domain of the β -subunit is needed, and in presence of the A-769662 direct activator increases
52 the AMPK activity more than 90-fold [13].

53 Since then, a lot of efforts have been carried out obtaining several direct AMPK activators, which in
54 some cases show a marked isoform selectivity [14], while in other cases no significant specificity is
55 obtained towards a specific subunit isoform. This is the case of SC4 which can activate both β 1- and
56 β 2-isoforms, although a slightly higher activation is observed in case of the β 1-isoform. However,
57 this ligand was a priori designed to increase the selectivity upon the α -subunit, being more selective
58 for the α 2-isoform [15,16]. Another interesting example regarding pan-activators; i.e., activators that
59 can activate independently several isoforms, is the PF-739, which is able to activate both α 2 β 1 γ 1 and
60 α 2 β 2 γ 1 complexes (Figure 1). However, similarly to SC4 ligand, PF-739 still exhibits a larger
61 affinity for the β 1-containing isoforms [17]. In our previous works [18,19], we have studied the
62 molecular factors that determine that one ligand, like SC4, could be able to activate both β -isoforms,
63 although it seems to show higher affinity towards one specific isoform. We have hypothesized that
64 the change of β 1-Asn111 by β 2-Asp111 could be one of the responsible to show this preference.
65 Here, we would like to go a little bit deeper, and study if PF-739 accomplish similar structural
66 behaviors of those previously observed. We would like to emphasize that there is a critical difference
67 between the SC4 and the PF-739, since while the first one is negatively charged, the PF-739 is a
68 neutral compound. Could this change affect the structural behavior of the α 2 β 1 and α 2 β 2
69 complexes? Are we interested in obtaining a pan-activator regarding any of the subunits or it would
70 be preferably to have specific activators for the full set of 12 AMPK heterotrimers? Could we be able
71 to determine which are the molecular factors that could help us to design a specific pharmacophore
72 for each subunit and isoform? With these ideas in mind, we have carried out extensive molecular
73 dynamic simulations to shed some light over these questions that could help in the drug design of this
74 important therapeutic target.

75 2 Results and Discussion

76 MD simulations were run to assess the structural and dynamics of α 2 β 1 and α 2 β 2 apo systems and
77 complexed with PF-739 (holo), and in presence of ATP molecule (holo+ATP). Three independent
78 replicas (1 μ s/replica) were run for each system, leading to a total simulation time of 6 μ s for the apo
79 species and 24 μ s for the ligand-bound complexes.



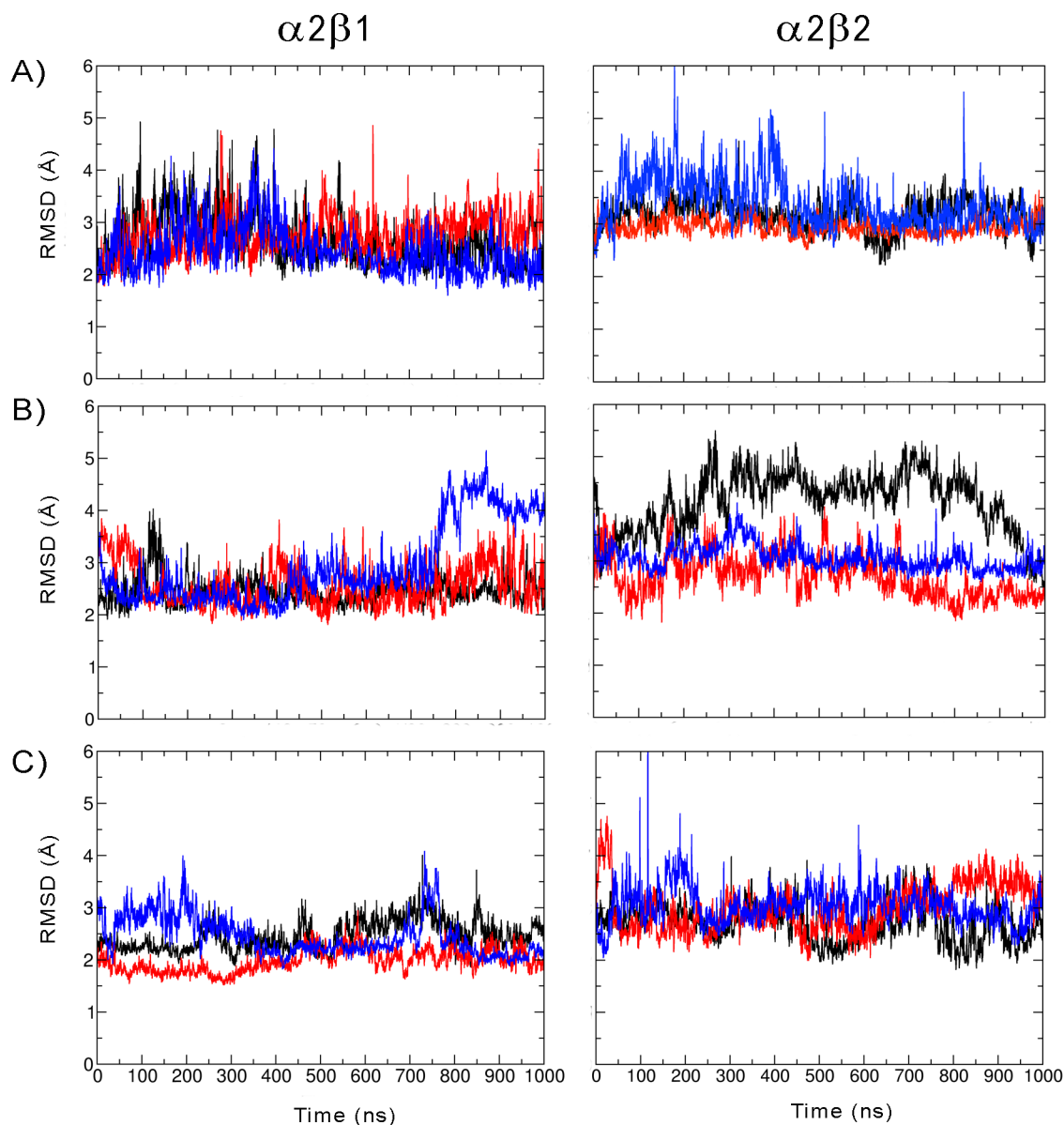
80

81 **Figure 1.** A) Representation of the simulation system selected for the study of the direct activation
 82 mechanism, which is composed of $\alpha 2$ (yellow cartoon) and $\beta 1$ -/ $\beta 2$ -subunits (grey cartoons). The
 83 most important regions of these subunits are highlighted; P-loop (purple), activation loop (A-loop,
 84 cyan), CBM domain (orange) and C-interacting helix (green). B) Chemical structure of PF-739,
 85 together with the experimental data obtained in Ref. 17 and the activator bound to the ADaM site.
 86 Relevant residues of the activator binding site are shown as white sticks.

87 2.1 Structural analysis of $\alpha\beta$ AMPK subunits complexed with ligand PF-739

88 We have examined the effect of the PF-739 ligand bound to the ADaM site (holo structures), and the
 89 binding of the ligand together with the ATP in the ATP-binding site (holo+ATP structures), over the
 90 global structural conformation of the apo $\alpha 2\beta 1$ and $\alpha 2\beta 2$ by means of the root mean square deviation
 91 (RMSD) of the protein backbone along the 1 μ s simulations (Figure 2). We have used as a reference
 92 to perform these analyses the average structure of the holo+ATP species sampled in the last 200 ns of
 93 the three independent replicas run for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ species. For the holo+ATP systems we can
 94 observe that there is a high structural resemblance for all the replicas as observed by small
 95 fluctuations in the RMSD profiles (Figure 2C), which agrees with the preservation of the overall fold
 96 of the protein upon binding of both PF-739 activator at ADaM site and ATP to the ATP-binding site.
 97 It is worth noting that this preservation is also noted in the RMSD values ranging from 2.0 to 2.5 for
 98 $\alpha 2\beta 1$ and from 2.7 to 3.0 for $\alpha 2\beta 2$ (Table 1). However, for the $\alpha 2\beta 1$ specie the RMSD of the
 99 simulated structures is not only not reduced from the apo (2.5 to 2.7 Å) to the holo (2.5 to 2.9 Å)

100 state, but also, in some replicas the RMSD slightly increase, and only in the presence of both the
 101 ligand and the ATP (holo+ATP) there is a proper reduction of the RMSD values. This effect is even
 102 more remarkable in the $\alpha 2\beta 2$ species, where in general in all the states (apo, holo and holo+ATP) the
 103 fluctuations of the overall structure seem to be larger than in $\alpha 2\beta 1$ (Figure 2 and Table 2). These
 104 results suggest that the PF-739 classified as a pan-activator exerts a weaker structural stabilization
 105 upon the binding of both $\alpha 2\beta 1$ and $\alpha 2\beta 2$ species.



106
 107 **Figure 2.** Root mean squared deviation (RMSD, Å) determined for the protein backbone along the
 108 three 1 μ s MD simulations run for the A) apo, B) holo and C) holo+ATP species of AMPK isoforms
 109 $\alpha 2\beta 1$ and $\alpha 2\beta 2$ bound to PF-739 (each replica is shown in black, blue and red, respectively). For each
 110 analysis the reference structure used corresponds to the energy-minimized average structure of the
 111 holo+ATP sampled in the last 200 ns of the three independent MD simulations.

112

113 **Table 1.** RMSD and standard deviation (Å) determined for the protein backbone of the snapshots
 114 sampled along the last 500 ns of MD simulations performed for all systems (apo, holo and holo+ATP
 115 states of AMPK isoforms $\alpha 2\beta 1$ and $\alpha 2\beta 2$. Values were determined using the energy-minimized
 116 holo+ATP species averaged for the last 200 ns of each simulation system as reference structure.

	System	Replica 1	Replica 2	Replica 3	Average
$\alpha 2\beta 1$	apo	2.6 ± 0.6	2.7 ± 0.4	2.5 ± 0.5	2.6
	holo	2.5 ± 0.3	2.6 ± 0.5	2.9 ± 0.8	2.6
	holo+ATP	2.5 ± 0.3	2.0 ± 0.2	2.4 ± 0.4	2.3
$\alpha 2\beta 2$	apo	3.2 ± 0.3	2.9 ± 0.2	3.4 ± 0.5	3.2
	holo	4.1 ± 0.6	2.7 ± 0.4	3.1 ± 0.3	3.3
	holo+ATP	2.7 ± 0.4	3.0 ± 0.5	3.0 ± 0.4	2.9

117

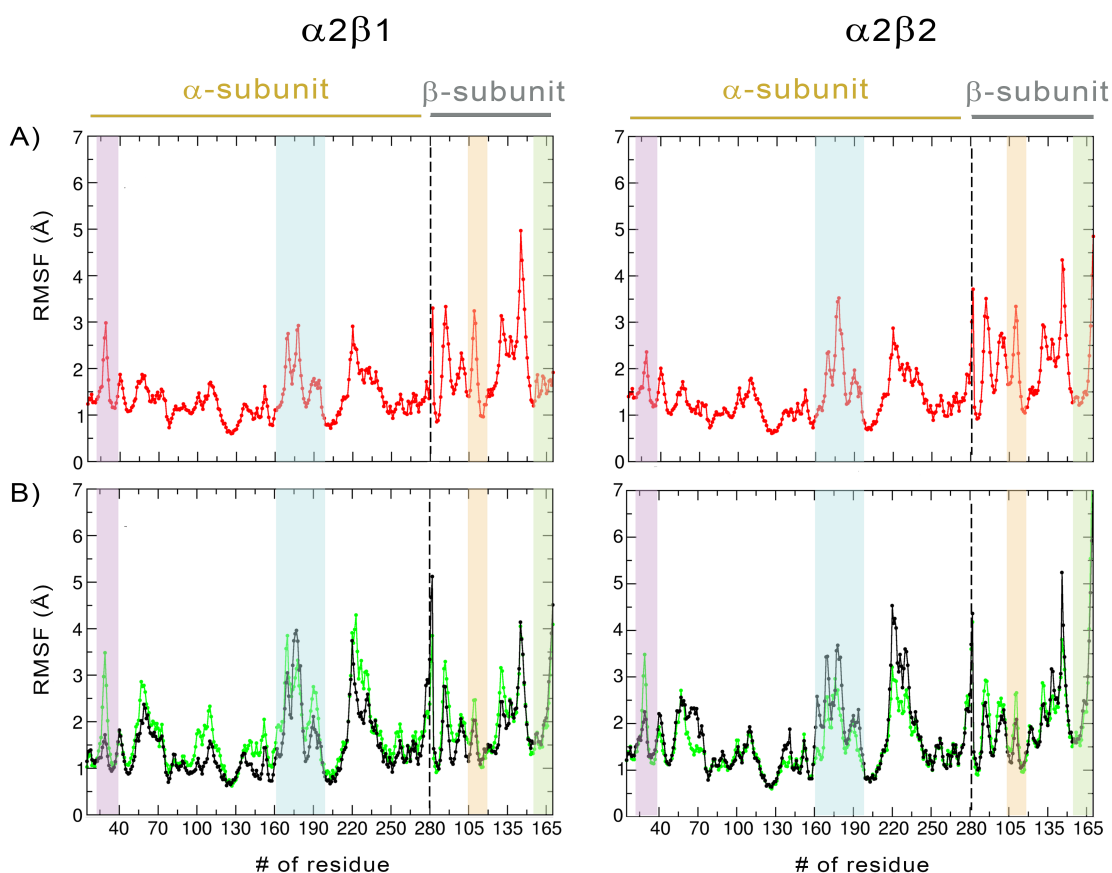
118 Regarding the per-residue mean square fluctuation (RMSF) profile, similar results are observed for
 119 both $\alpha 2\beta 1$ and $\alpha 2\beta 2$ species, as noted in the resemblance of the fluctuation pattern obtained by
 120 averaging the RMSF of the three replicas run for every system (Figure 3). The highest fluctuations in
 121 the α -subunit correspond to the residues in the activation loop (residues 165-185, highlighted in blue
 122 Figure 3) and the α -helix between residues 210-230. It is worth noting the higher fluctuation of the P-
 123 loop (residues 15-35, highlighted in purple in Figure 3) in the holo state in comparison to the
 124 holo+ATP, but also to the apo state, which is also in agreement with the higher RMSD observed in
 125 the holo species. We would like to remark that the P-loop moiety belongs to the kinase domain of the
 126 α -subunit, regulating the structural ensemble of both the ADaM and the ATP-binding site. Thus, the
 127 binding of the PF-739 in both $\alpha 2\beta 1$ and $\alpha 2\beta 2$ species, alters significantly its conformation. Regarding
 128 the fluctuations in the β -subunit, the largest oscillations regard on the CBM domain, which contains
 129 the Ser108 (highlighted in orange in Figure 3), phosphorylated in the holo and holo+ATP states, and
 130 the regions near to the C-interacting helix, at the terminal region (residues 162-172, highlighted in
 131 green in Figure 3). In general, can be observed that the binding of the PF-739 in holo states, as well
 132 as, ATP in holo+ATP states increase the fluctuations of the main moieties mentioned in the α -
 133 subunit, while reduce the fluctuations in the main regions of the β -subunit, independently of the β -
 134 isoform.

135 2.2 Dynamic properties of AMPK complexes

136 The protein dynamics is affected by the interaction of that protein with a small organic molecule, like
 137 the PF-739 activator. In order to examine the effect of the activator on the conformational behavior of
 138 the AMPK species, we have analyzed the essential dynamics (ED) of the protein backbone and the
 139 cross-correlation map of the residues.

140 The ED provide information about which are the essential motions of the protein, and which are the
 141 effects on these motions due to the ligand binding. The results for the first essential motion of apo
 142 states for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ species show a concerted bending that brings the α - and the β -subunits
 143 closer and then moves them apart (Figure 4). The most interesting feature is that the P-loop seems to
 144 act as a hinge, allowing this concerted bending between the subunits. Indeed, the first motion
 145 accounts on average for the 41% of the structural variance in $\alpha 2\beta 1$ from a total of 66% accounting
 146 from the first to the fourth essential motion, and the 31% in the $\alpha 2\beta 2$ from a total of 57% considering
 147 the first four essential motions (see Table 2). Therefore, this data emphasizes the importance of the
 148 first essential motion to the conformational flexibility of the AMPK complexes.

149



150

151 **Figure 3.** RMSF (Å) average of the residues determined along the last 500 ns of the three
 152 independent replicas run for A) apo and B) holo and holo+ATP (green and black lines, respectively).
 153 The highlighted bars denote the moieties corresponding to P-loop (purple), activation loop (cyan),
 154 CBM domain (orange), and C-interacting helix (green).

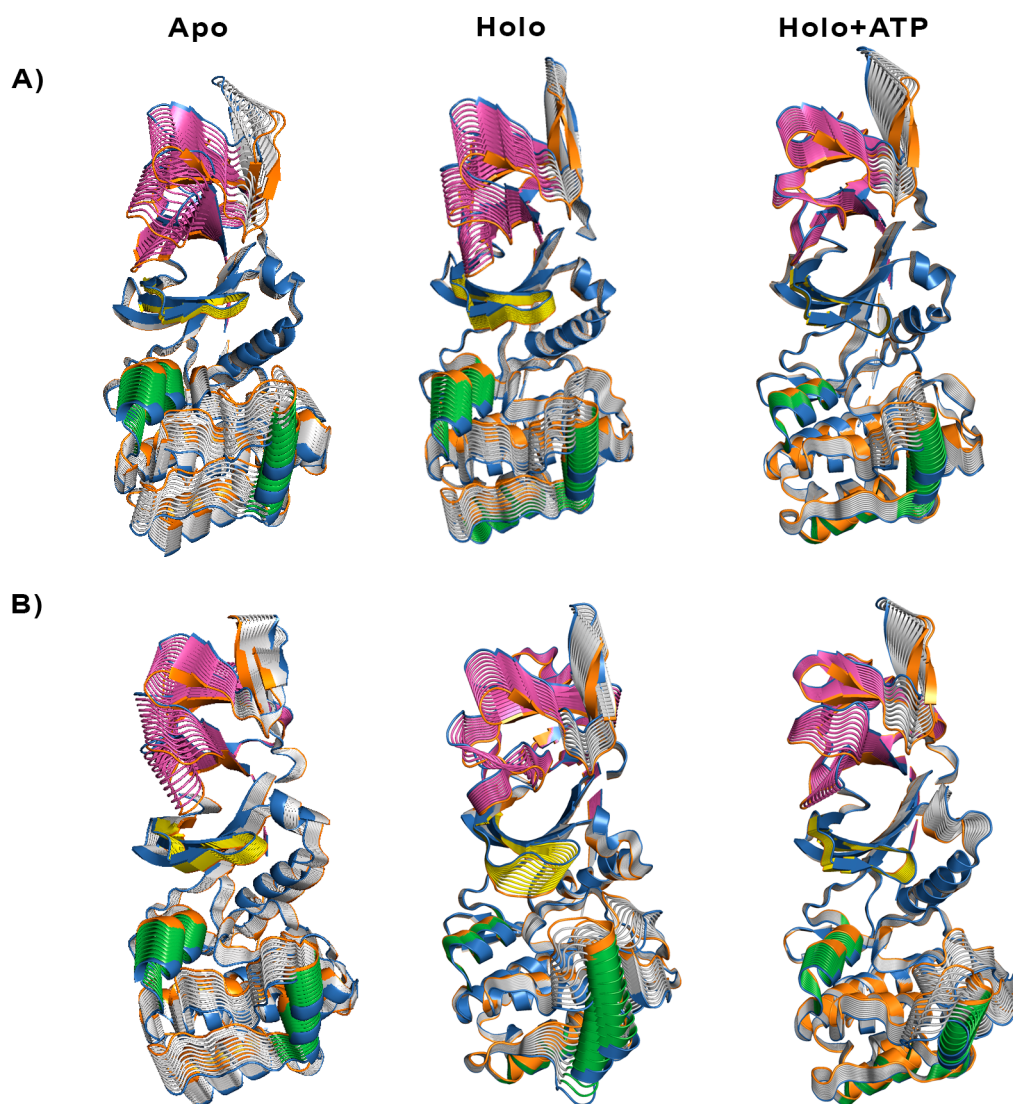
155 When move from apo to holo and holo+ATP states, it is observed that the binding of the activator
 156 reduces the conformational variance from 66% to 62% and 55%, respectively (see Table 2), for the
 157 $\alpha 2\beta 1$ species. However, for the $\alpha 2\beta 2$ species, the binding of the activator involves a slight increase in
 158 the conformational variance when move from apo to holo, at the contrary to the behavior observed in
 159 $\alpha 2\beta 1$ species, to then decrease in the holo+ATP states. These results are also reflected in the
 160 evolution of the conformational behavior of the projection 1 (Figure 4 and Table 2). The examination
 161 of the first essential motion of holo states presents a synchronous motion of the P-loop and the CBM
 162 domain. However, although the movements of the CBM domain are very similar between $\alpha 2\beta 1$ and
 163 $\alpha 2\beta 2$ species, the P-loop and the helices at the C-terminal region of the α -subunit suffer higher
 164 fluctuations in $\alpha 2\beta 2$ with respect to the $\alpha 2\beta 1$. When move to the holo+ATP state, the overall
 165 flexibility of the system is achieved, especially regarding the P-loop, and the kinase domain in the α -
 166 subunit, as well as the region of the CBM domain nearest to the ADaM site, reflecting the synergy
 167 between activator and ATP in increasing the stiffness. Additionally, we would like to remark that the
 168 stiffness achieved from the holo to holo+ATP states is also more remarkable in the case of $\alpha 2\beta 1$
 169 species. These conformational changes agree with the difference in the variability of the essential
 170 motions observed in Table 2.

171 To complement the previous results, dynamical cross-correlation (DCC) analysis was performed to
 172 examine the correlated motions of residues in AMPK complexes. Results for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ species
 173 interacting with PF-739 are shown in Figure 5. The apo systems (Figure 5A) reveal a correlation
 174 between the P-loop and the α C-helix from the α -subunit, and with the CBM and the C-interacting
 175 helix from the β -subunit (as noted by the yellow marks). The holo+ATP systems show lower
 176 dynamical correlation between residues, as observed by the progressive reduction in the number and
 177 intensity of the areas that exhibit a pronounced correlation (shown in yellow and blue for highly
 178 correlated and anticorrelated fluctuations between residues). On the contrary, the correlation between
 179 the motion of the P-loop and the CBM domain is reinforced in this state (black square in Figure 5).
 180 These effects are more noticeable when compare holo and holo+ATP systems in $\alpha 2\beta 1$ specie.
 181

182 **Table 2.** Contribution of the essential motion (%) to the structural variance of different AMPK
 183 systems and the total contribution of the first four projections.

Systems		Proj. 1	Proj. 2	Proj. 3	Proj.4	Total (P1-P4)
$\alpha 2\beta 1$	apo	41.2	12.0	8.1	4.6	66.0
	holo	38.6	12.1	7.6	4.0	62.3
	holo+ATP	30.7	12.6	7.0	5.1	55.4
$\alpha 2\beta 2$	apo	30.9	12.6	8.8	4.8	57.1
	holo	33.1	12.7	8.6	5.3	59.7
	holo+ATP	29.0	13.2	7.2	5.3	54.7

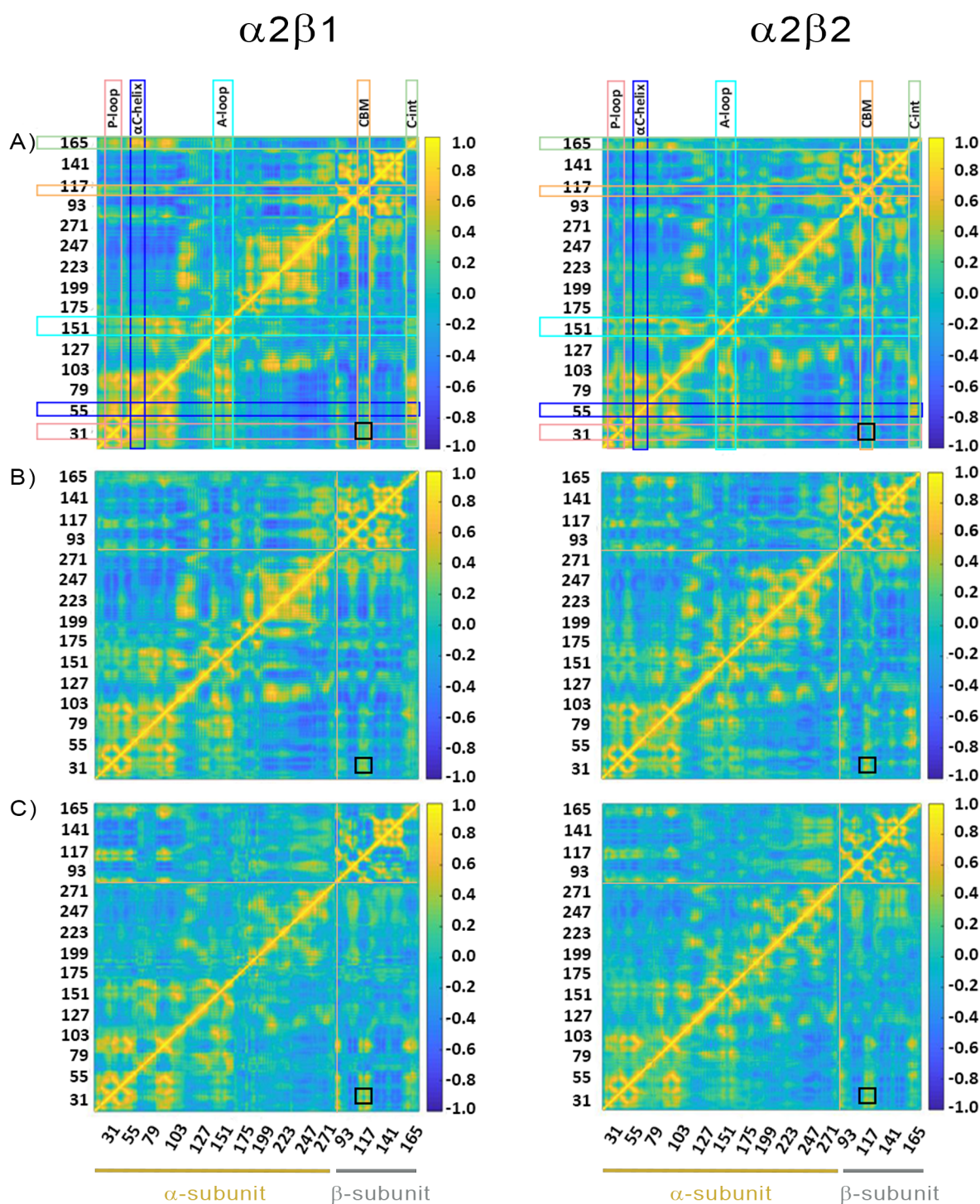
184 Although the preceding results show a high similarity in the dynamical behavior of both $\alpha 2\beta 1$ and
 185 $\alpha 2\beta 2$ species bound to PF-739 activator, which agrees with the definition of PF-739 as a pan-
 186 activator able to activate both $\beta 1$ - and $\beta 2$ -isoforms, still these global structural analyses reveal subtle
 187 differences between $\beta 1$ - and $\beta 2$ -containing AMPK complexes. In particular, the results suggest that
 188 the $\alpha 2\beta 2$ species have a slightly larger resilience to the structural modulation exerted by the activator,
 189 whereas the $\alpha 2\beta 1$ isoform is more sensitive to the conformational adaptation induced upon activator
 190 binding to the ADaM site, enhancing the stiffness of protein backbone for the $\beta 1$ -containing complex
 191 (see Figure 4A). These results are also in agreement with the experimental measurements that shows
 192 that, although PF-739 can activate both $\alpha 2\beta 1\gamma 1$ and $\alpha 2\beta 2\gamma 1$ complexes, still exhibits a larger affinity
 193 for the $\beta 1$ -isoform (see Figure 1) [17]. These results are also in agreement with our latest results
 194 published for SC4 direct activator, which is also able to activate both $\beta 1$ - and $\beta 2$ -containing AMPK
 195 complexes, although a higher affinity for the $\beta 1$ -isoform is also observed [19].
 196
 197



198

199 **Figure 4.** Representation of the first essential motion derived from the ED analysis of the protein
200 backbone for A) the $\alpha 2\beta 1$ and B) $\alpha 2\beta 2$ species, determined from the snapshots sampled along the last
201 500 ns of simulations. The P-loop is shown in yellow, and the helices formed by residues 100-110
202 and 220-229 in the α -subunit are shown in green. The CBM domain is colored in magenta.

203

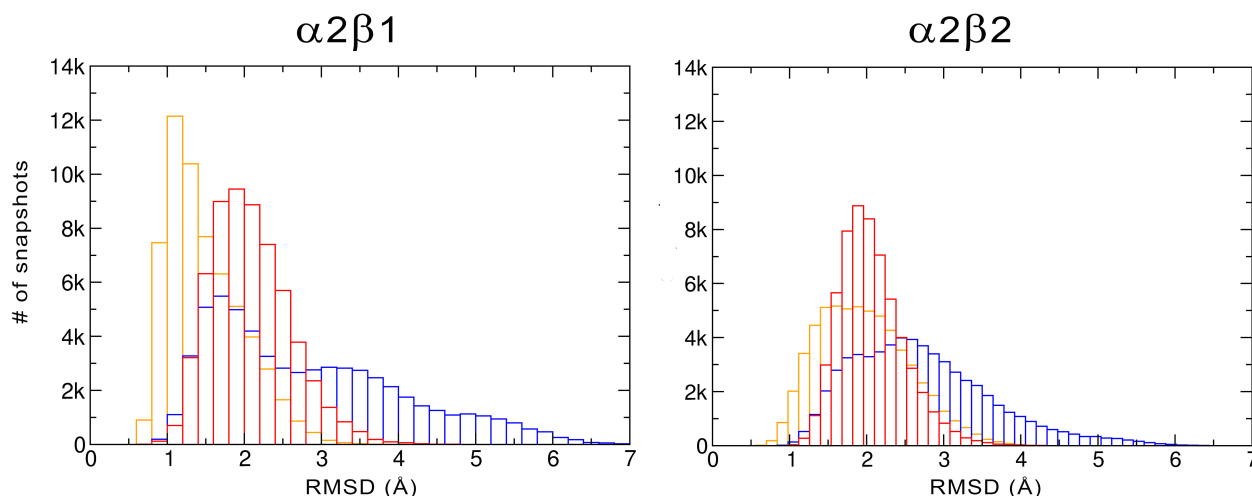


204

205 **Figure 5.** Dynamic cross-correlation (DCC) matrices for A) apo, B) holo and C) holo+ATP
 206 complexes of (left) $\alpha 2\beta 1$ and (right) $\alpha 2\beta 2$ with PF-739. The x- and y-axis denote the numbering of
 207 residue in the α - and β -subunit, indicated at the bottom of the plot (yellow and gray lines for α - and
 208 β - subunits, respectively). Specific regions of AMPK are highlighted with squares at the top of the plot: P-loop (pink), α C-helix (blue), activation loop (A-loop; cyan), CBM unit (orange) and C-
 209 interacting helix (C-int; green). Regions colored in yellow/blue show high correlated/anticorrelated
 210 fluctuations. The black square in the plots highlights the motion between the CBM domain and the P-
 211 loop.
 212

213 **2.3 Pre-organization of ATP-binding site**

214 With the aim to gain some new insight of how the binding of the PF-739 could influence the
 215 activation of the AMPK enzyme, and with the idea to discern the similar, but still the higher binding
 216 affinity of PF-739 upon the β 1-isoform, we have evaluated the dynamical response of the ATP-
 217 binding site due to the binding of the activator in the ADaM site. We have assessed the pre-
 218 organization of the ATP-binding site in the apo, holo and holo+ATP states, using as a reference the
 219 average structure of the holo+ATP complex.



220

221 **Figure 6.** Distribution of the positional deviation (RMSD; Å) of the structures sampled along the
 222 trajectories run for apo (red), holo (blue), and holo+ATP (orange) for the residues that shape the
 223 ATP-binding site (residues α 22– α 32, α 42– α 46, α 75– α 79, α 142– α 147, and α 153– α 157). A total of
 224 60,000 snapshots taken from the last 500 ns of MD simulations were considered for each system in
 225 the analysis.

226 For the holo+ATP states, the residues of the ATP-binding site sample a reduced conformational
 227 space, with a peak centered at a positional RMSD of 1.2 Å for α 2 β 1 and a little bit wider distribution
 228 with a peak centered at 1.9 Å for α 2 β 2 specie (Figure 6). Unexpectedly, the apo state shows a
 229 narrower distribution and with one main peak centered at 2.0 Å for both α 2 β 1 and α 2 β 2 species. This
 230 fact is completely different to our previous reports, where we studied the behavior of A-769662 or
 231 SC4 small activators, where the apo structure show larger distributions, or even a bimodal mode [19].
 232 In the case of complexes with PF-739, the conformations sampled by the apo state (without the
 233 presence of the ligand) are very similar to those conformations already sampled by the holo+ATP
 234 specie. This fact is even more remarkable in the case of the α 2 β 2 specie, where the distribution of the
 235 holo+ATP state is wider while the distribution of the apo state is narrower, and thus, both states
 236 comprise very similar profiles.

237 On the contrary, the holo state exhibits a wider distribution, showing a trimodal RMSD profile, with
 238 peak values at 1.7, 3.2 and 5.0 Å, for α 2 β 1, while for the α 2 β 2 species a bimodal distribution is
 239 observed with peaks values at 1.8 and 2.5 Å. In both α 2 β 1 and α 2 β 2 species, the first peak of the holo
 240 distribution corresponds with the holo+ATP resemblance, however these results suggest that the
 241 binding of the PF-739 to the ADaM site implies more fluctuations than rigidification of the structure,
 242 and therefore, due to this higher conformational flexibility, the ATP-binding site can adopt
 243 conformations close to those populated in the holo+ATP state, but also visit more dissimilar
 244 conformational regions (Figure 6).

245 2.4 Structural basis of the AMPK activation by pan-activator PF-739

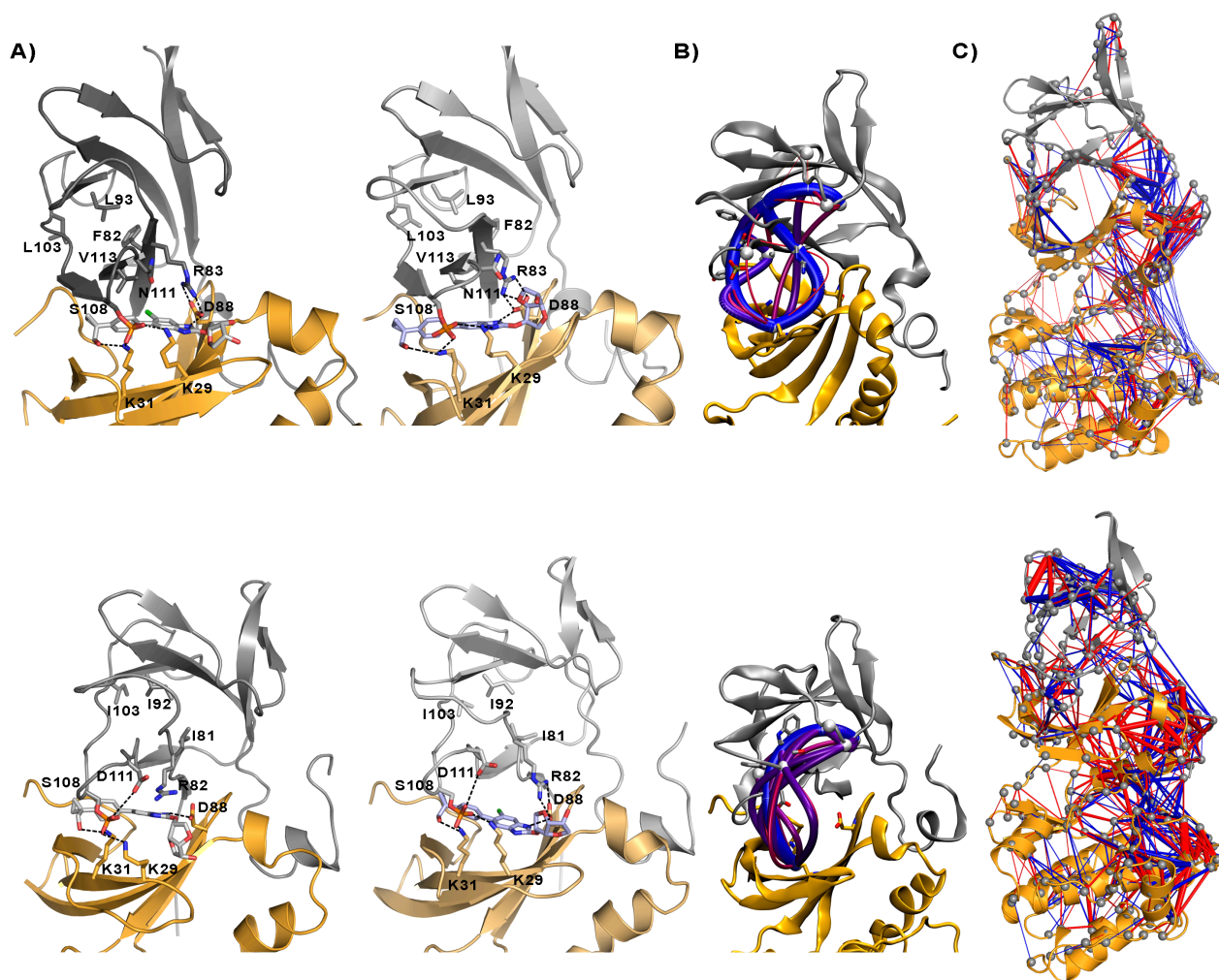
246 In order to understand the conformational behavior observed for the holo state in $\alpha 2\beta 1$ and $\alpha 2\beta 2$
247 species we have explored the structural moieties involved in the binding of the direct pan-activator
248 PF-739. We have examined the interaction network formed by the PF-739 ligand, the P-loop in the
249 kinase domain of the α -subunit and the CBM domain from the β -subunit.

250 We can divide the interaction network in two regions. The first one corresponds to the salt bridge
251 formed between the $\beta 1$ -Arg83 with the $\alpha 2$ -Asp88, which at the same time is making a hydrogen bond
252 interaction with the PF-739 ligand. This pattern is maintained for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ species along the
253 simulation (Figure 7A). The second region of interactions regards on the contacts performed by the
254 pSer108 located at the CBM domain of the β -subunit. The pSer108 performs salt bridges interactions
255 with the Lys29 and Lys31, both from the P-loop of the α -subunit. Moreover, an additional contact
256 conserved along the simulation and for all the species regards on the hydrogen bond between the
257 Lys31 and the hydroxymethyl-cyclopropyl group. In Figure 7A, we can observe a representative
258 snapshot of holo $\alpha 2\beta 1$ system after 500 ns simulation (top-left) and after 1 μ s of simulation (top-
259 right), showing that the motion of the sugar-like mannitol appendage involves an important change in
260 the P-loop, favoring a new interaction between the Lys29 of the P-loop and the ligand. Additionally,
261 this conformation, also lets the formation of a new interaction between the $\beta 1$ -Arg83 and the ligand,
262 maintaining the previous interactions described. These two conformations have been sample along
263 the simulations for this system. This orientational change of the sugar-like mannitol group of the
264 ligand has a higher critical effect on the $\alpha 2\beta 2$ species, due to the substitution of $\beta 1$ -Asn111 to $\beta 2$ -
265 Asp111. At the same time, this one-point change in the residues provokes two different
266 conformations of the $\beta 2$ -Arg82. In the case where the sugar-like mannitol group is facing towards the
267 P-loop (Figure 7A, bottom-left), the $\beta 2$ -Arg82 is making a clear interaction with the $\beta 2$ -Asp111,
268 losing all the contact with the $\alpha 2$ -Asp88. Additionally, the $\beta 2$ -Asp111 is also in contact with the β -
269 pSer108, which maintains its interactions with both α Lys29 and α Lys 31 from the P-loop. This chain
270 of interactions reinforced the connection between the $\alpha 2$ and the $\beta 2$ subunits basically through the
271 CBM domain and the P-loop. However, when the sugar-like mannitol group is facing up, the $\beta 2$ -
272 Arg82 recovers the salt bridge interaction with the $\alpha 2$ -Asp88, losing all the contact with the $\beta 2$ -
273 Asp111. Nevertheless, this conformation is less stable and less sampled during the simulation.

274 These observations are also reflected in the major interaction pathways identified from WISP
275 analysis for the two holo species. In the holo $\alpha 2\beta 1$ (Figure 7B, top) the CBM domain and the P-loop
276 are connected via two main pathways. One involves the pathway through the β -pSer108 and the
277 ligand with their connections with the Lys29 and Lys31 from the P-loop, while the second path
278 involves the triad $\beta 2$ -Arg82 – $\alpha 2$ -Asp88 – PF-739. However, in the holo $\alpha 2\beta 2$ system (Figure 7B,
279 top) only one main pathway is observed, which involves the connection between $\alpha 2$ - and $\beta 2$ -subunits
280 through the β -pSer108. These results agree with the interaction $\beta 2$ -Arg82 – $\beta 2$ -Asp111, which is
281 observed along the simulation, weakened considerably the interaction $\beta 2$ -Arg82 – $\alpha 2$ -Asp88, which
282 is rarely observed along the simulation. Indeed, these results could explain the higher affinity of PF-
283 739 towards the $\beta 1$ -isoform experimentally observed (Figure 1).

284 Finally, we have performed a dynamical contact network which provides information about which
285 contacts are lost (blue) and gained (red) when we compare two different states (apo – holo). In this
286 case, we have compared the difference dynamics between the apo and holo states for $\alpha 2\beta 1$ species
287 (Figure 7C, top) and $\alpha 2\beta 2$ (Figure 7C, bottom). Both species show similar dynamic contacts
288 regarding the CBM domain in the β -subunit. However, the $\alpha 2\beta 2$ species show larger changes in the
289 dynamic contacts between apo and holo systems with respect to the $\alpha 2\beta 1$ species. This result

290 suggests that the ligand does not involve a large conformational change of the whole $\alpha 2\beta 1$ system,
 291 and only the domains nearest to the ADaM and ATP-binding site, like the P-loop, αB - and αC -helix,
 292 and A-loop in the α -subunit, and the C-interacting helix and the CBM domain in the β -subunit are
 293 the most affected areas. In contrast in $\alpha 2\beta 2$ species, although the regions affected are the same than
 294 described before, the size of the edges that connects the dynamic residues are considerably larger,
 295 suggesting that the number of contacts are higher in this case, and therefore, implying a higher effect
 296 from the ligand. This dynamical behavior is also a sign of destabilization in the system due to the
 297 presence of the ligand, so, although the PF-739 is defined as a pan-activator since it can activate both
 298 β -isoforms, altogether, our results explain why there is a higher affinity upon the $\beta 1$ -isoform.



299

300 **Figure 7.** A) Representation of selected interactions between the CBM, P-loop and PF-739 in holo
 301 $\alpha 2\beta 1$ (top) and $\alpha 2\beta 2$ (bottom) species. The α -subunit is shown in orange cartoon, while the β -subunit
 302 is shown in grey cartoons. PF-739 is shown in grey and light blue sticks in the ADaM site. Selected
 303 polar interactions are highlighted as dashed lines. B) Major interaction networks obtained from WISP
 304 analysis. C) Substraction average of the dynamical contact network of apo – holo states for $\alpha 2\beta 1$
 305 (top) and $\alpha 2\beta 2$ (bottom) species. The edges colored in red show atomic couples of contacts formed in
 306 the holo state with respect to the apo state, while the blue edges represent contacts that are lost in the
 307 holo state.

308

309 3 Conclusions

310 Discern the molecular factors that regulates the structure-function relationships of AMPK isoforms is
311 of utmost importance to rationalize the tissue-dependent expression of AMPK complexes, and thus
312 enabling the design of specific compounds active against specific metabolic disorders. However, the
313 understanding of which are the differences between isoforms that allow a different ligand behavior
314 (i.e., selective activator, pan-activator or even inhibitor) is very challenging due to the high structural
315 complexity of the enzyme and the highly correlated dynamics observed for both $\alpha 2\beta 1$ and $\alpha 2\beta 2$
316 species.

317 However, our last results have confirmed that the subtle difference of $\beta 1$ -Asn111 to $\beta 2$ -Asp111,
318 could have larger implications in the dynamical response of AMPK to the binding of activators. This
319 single substitution can change the interaction networks formed surrounded the activator, thus
320 inducing a better mechanical response of the $\alpha 2\beta 1$ specie towards the interaction of PF-739, than in
321 the case of the $\alpha 2\beta 2$ species. So, even in case of a pan-activator like the PF-739, able to activate both
322 β -isoforms, still subtle residue substitutions in the ADaM site are responsible of difference in affinity
323 towards the isoform.

324 In summary, we were able to discern and confirm the key molecular features that mediate the
325 activation of pan-activator towards $\alpha 2\beta 1$ and $\alpha 2\beta 2$ species. All these findings shed light in the
326 comprehension of the role of specific residues in the ADaM site that can modulate or completely
327 change the direct activation mechanism of $\beta 1$ - and $\beta 2$ -containing AMPK complexes. Future studies
328 will be appreciated to distinguish the structural basis of the different sensitivity of AMPK complexes
329 formed by distinct α -subunits, and which is more important, the study of the full complex to
330 disentangle the full allosteric network connection. This understanding will really enable us the design
331 of tissue-selective modulators of this cellular energy sensor.

332 4 Materials and Methods

333 4.1 Molecular dynamics simulations.

334 Extended molecular dynamics (MD) simulations were utilized to analyze the structural and
335 dynamical characteristics of the simulated system. For this purpose, the $\alpha 2\beta 1\gamma 1$ systems were built
336 up using the complexes with A-769662 (PDB entry 4CFF) [20]. On the other hand, the system
337 related to the complex of $\alpha 2\beta 2\gamma 1$ bound to SC4 (PDB entry 6B2E) [15] was also used as a template
338 to model the complexes with PF-739. Following our previous studies [18,19], the γ -subunit was not
339 considered in MD simulations due to several reasons. First, the ADaM site is shaped only by α - and
340 β - isoforms. Furthermore, the lack of precise structural information about stretches of both α - and β -
341 subunits, particularly regarding the C-terminal regions, which are located close to the γ -subunit,
342 would introduce an additional level of uncertainty, opening the way to potential artefacts in the
343 simulations. Finally, inclusion of the γ -subunit would have required a larger computational cost in
344 order to guarantee a proper sampling of the dynamical motions of the three isoforms. Accordingly,
345 following the 'divide-and-conquer' strategy outlined above, the simulated systems comprise only α -
346 and β -subunits. Specifically, simulations were performed for residues 8-278 of the $\alpha 2$ isoform, and
347 residues 78-173 and 77-171 of the $\beta 1$ - and $\beta 2$ -isoforms, which were solved without disruptions in the
348 X-ray structures. Finally, these structures were used to model the apo protein, the complexes of the
349 activators bound to the phosphorylated Ser108 (pSer108)-containing isoforms (holo), and the
350 corresponding holo+ATP complexes with both activator in the ADaM site and ATP in the ATP-
351 binding site. Simulations were performed using the AMBER18 package [21] and the Amber
352 ff99SBILDN force field [22] for the protein, whereas the ligand (PF-739) were parameterized using

353 the GAFF [23] force field in conjunction with restrained electrostatic potential-fitted (RESP) partial
354 atomic charges derived from B3LYP/6-31G(d) calculations [24]. The standard protonation state at
355 physiological pH was assigned to ionisable residues, and a capping group (N-methyl) was added to
356 the C-terminus of the α -subunit. The simulated systems were immersed in an octahedral box of
357 TIP3P water molecules [25]. Counterions atoms were added to maintain the neutrality of the
358 simulated systems [26]. The final systems included the AMPK protein, around 25,500–26,700 water
359 molecules, and a variable number of Na⁺ and Cl⁻ ions, leading to simulated systems that contain
360 between 83000-86000 atoms.

361 Simulations were performed in the NPT ensemble for equilibration and NVT for MD productions
362 using periodic boundary conditions and Ewald sums (grid spacing 1 Å) for treating long-range
363 electrostatic interactions. Apo, holo and holo+ATP systems were simulated in triplicate. The
364 minimization of the systems was performed refining the position of hydrogen atoms in the protein
365 (2000 cycles of steepest descent algorithm followed by 8000 cycles of conjugate gradient),
366 subsequently minimizing the position of water molecules (using again the previous scheme), and
367 finally minimization of the whole system (4000 cycles for steepest descent and 1000 cycles of
368 conjugate gradient). Later, the temperature of the system was gradually raised from 100 to 300 K in 5
369 steps, 50 ps each using the NVT ensemble. In this process, suitable restraints (5 kcal mol⁻¹ Å⁻²) were
370 imposed to keep the ligand (activator, ATP) in the binding pocket and prevent artefactual
371 rearrangements along the equilibration stage. In order to equilibrate the density of the system an
372 additional 5 ns step performed in the NPT ensemble. In addition, the restraints were progressively
373 eliminated in this later step. Production MD simulations were run for 1 μ s per replica, leading to a
374 total simulation time of 12 μ s for the ligand-bound AMPK complexes, and 6 μ s for the two apo
375 species of AMPK.

376 4.2 Essential Dynamics.

377 This method was utilized to specify the most important motions from the structural variance sampled
378 in MD simulations. In essential dynamics (ED) [27], the dynamics along the individual modes can be
379 studied and visualized separately, so we can filter the main collective motions during our simulations.
380 Therefore, the positional covariance matrix is created and diagonalized in order to achieve the
381 collective deformation modes, i.e. the eigenvectors, while the eigenvalues account for the
382 contribution of each motion to the structural variance of the protein. ED analysis was done for 25000
383 snapshots from the last 500 ns of each simulation, taking into account only the backbone atoms and
384 the calculations were performed with PCAsuite program (available at
385 <http://www.mmb.irbbarcelona.org/software/pcasuite/pcasuite.html>), which is integrated in the
386 pyPCCazip program, a suite of tools for compression and analysis of molecular simulations [28].

387 4.3 Conformational entropy.

388 The quasi-harmonic approximation proposed by Schlitter [29] was used to calculate the
389 conformational entropy sampled for the different simulation systems. Due to the sensitivity of these
390 calculations to simulation time and sampling [30], the conformational entropy determined at regular
391 interval times ($S(t)$) along the MD simulation was used to estimate the conformational entropy at
392 infinite time (S^∞) [31]. In particular, S^∞ was determined by fitting the asymptotic behaviour of the
393 curve according to the expression $S(t) = S^\infty \frac{t}{t+\alpha}$,

394 where α is an adjustable parameter by curve fitting. In order to enable a direct comparison between
395 the conformational entropies of the different $\alpha 2\beta 1$ and $\alpha 2\beta 2$ systems, only the set of atoms of the

396 protein backbone were considered in entropy calculations, thus avoiding a bias originated from the
 397 residue substitutions between β 1- and β 2-isoforms in the calculated entropies (note that a different
 398 choice of atoms, including both backbone and side chains, was used in the comparison of the
 399 conformational entropies reported for α 2 β 1 bound to A- 769662 in a previous study) [18].

400 **4.4 Dynamical Cross-Correlation Analysis.**

401 To complement the information gained from the ED analysis, dynamical cross-correlation (DCC)
 402 was used to examine the correlation motion of residues along a given trajectory. To this end, all the
 403 snapshots were aligned by means of least-square fitting of $C\alpha$ atoms of the whole protein to the
 404 equilibrated starting configuration. Then, the DCC matrix was determined as noted in Eq. 1.

$$C_{ij} = \frac{c_{ij}}{c_{ii}^{1/2} c_{jj}^{1/2}} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{[(\langle r_i^2 \rangle - \langle r_i \rangle^2)(\langle r_j^2 \rangle - \langle r_j \rangle^2)]^{1/2}}$$

405 (1)

406 where the position vectors of two $C\alpha$ atoms i and j fitted in the structure at time t are denoted as $r_i(t)$
 407 and $r_j(t)$, respectively.

408 The cross-correlation coefficients range from -1 to +1, which represent anticorrelated and correlated
 409 motions, respectively, whereas values close to zero indicate the absence of correlated motions [32].
 410 This analysis was performed using the module available in AMBER package.

411 **4.5 Interaction Energy Network.**

412 Networks of local interactions are intrinsically linked to the structural response of proteins to external
 413 factors [33]. For our purposes, Weighted Implementation of Suboptimal Path (WISP) [34] was
 414 utilized to analyze the allosteric network. This method enabled us to perform a dynamic network
 415 analysis to understand how the binding of a ligand in an allosteric cavity could affect another binding
 416 site. In particular, WISP relies on the dynamical interdependence among the protein residues. To this
 417 end, each amino acid is treated as a node, which was located at the residue center-of-mass, and the
 418 interdependence among nodes is represented as a connecting edge with an associated numeric value
 419 that reflects its strength. The interdependence is determined from an $N \times N$ matrix C (N is the number
 420 of nodes) with values corresponding to the weights of each edge, reflecting the correlated motion
 421 among node-node pairs. Finally, the weight between the edge that connects nodes i and j is expressed
 422 as $w_{ij} = -\log(|C_{ij}|)$, so that highly correlated or anticorrelated motions are characterized by small
 423 values of w_{ij} . This analysis was performed for the last 500 ns of the MD simulations.

424 **5 Conflict of Interest**

425 *The authors declare that the research was conducted in the absence of any commercial or financial*
 426 *relationships that could be construed as a potential conflict of interest.*

427 **6 Author Contributions**

428 Elnaz Aledavood: Formal analysis, Investigation, Visualization, Writing - original draft. Aria
 429 Gheeraert: Formal analysis, Investigation, Visualization. Ivan Rivalta: Methodology, Investigation.
 430 Carolina Estarellas: Conceptualization, Methodology, Investigation, Supervision, Writing - review &

431 editing. F. Javier Luque: Conceptualization, Methodology, Investigation, Supervision, Writing -
432 review & editing, Funding acquisition

433 7 Acknowledgments

434 We thank the Spanish Ministerio de Economía y Competitividad (SAF2017-88107-R, and Maria de
435 Maetzu MDM-2017-0767, AEI/FEDER), and the Generalitat de Catalunya (2017SGR1746) for
436 financial support and the Barcelona Supercomputing Center (BCV-2014-3-0011 and BCV-2019-1-
437 0009) and the Consorci de Serveis Universitaris de Catalunya (CSUC) for computational resources.
438 E.A. thanks AGAUR (Generalitat of Catalunya; 2018FI-B1-00001) for a fellowship.

439 8 Reference styles

[1] Xiao, B.; Sanders, M. J.; Underwood, E.; Heath, R.; Mayer, F. V.; Carmena, D.; Jing, C.; Walker, P. A.; Eccleston, J. F.; Haire, L. F.; Saiu, P.; Howell, S. A.; Aasland, R.; Martin, S. R.; Carling, D.; Gamblin, S. J. Structure of Mammalian AMPK and its Regulation by ADP. *Nature* 2011, 472, 230–233.

[2] Carling, D. AMPK Signalling in Health and Disease. *Curr. Opin. Cell Biol.* 2017, 45, 31–37.

[3] Xiao, B.; Sanders, M. J.; Underwood, E.; Heath, R.; Carmena, D.; Jing, C.; Walker, P. A.; Eccleston, J. F.; Lesley, F.; Saiu, P.; et al. Structure of Mammalian AMPK and Its Regulation by ADP. *Nature* 2011, 472, 230–233.

[4] Calabrese, M. F.; Rajamohan, F.; Harris, M. S.; Caspers, N. L.; Magyar, R.; Withka, J. M.; Wang, H.; Borzilleri, K. A.; Sahasrabudhe, P. V.; Hoth, L. R.; et al. Structural Basis for AMPK Activation: Natural and Synthetic Ligands Regulate Kinase Activity from Opposite Poles by Different Molecular Mechanisms. *Structure* 2014, 22, 1161–1172.

[5] Mahlapuu, M.; Johansson, C.; Lindgren, K.; Hjlm, G.; Barnes, B. R.; Krook, A.; Zierath, J. R.; Andersson, L.; Marklund, S. Expression Profiling of the α -Subunits Isoforms of AMP-Activated Protein Kinase Suggests a Major Role for α 3 in White Skeletal Muscle. *Am. J. Physiol. Endocrinol. Metab.* 2004, 286, E194-200.

[6] Carling, D.; Thornton, C.; Woods, A.; Sanders, M. J. AMP-Activated Protein Kinase: New Regulation, New Roles? *Biochem. J.* 2012, 445, 11–27.

[7] Chen, L.; Wang, J.; Zhang, Y. Y.; Yan, S. F.; Neumann, D.; Schlattner, U.; Wang, Z. X.; Wu, J. W. AMP-Activated Protein Kinase Undergoes Nucleotide-Dependent Conformational Changes. *Nat. Struct. Mol. Biol.* 2012, 19, 716–718.

[8] Willows, R.; Sanders, M. J.; Xiao, B.; Patel, B. R.; Martin, S. R.; Read, J.; Wilson, J. R.; Hubbard, J.; Gamblin, S. J.; Carling, D. Phosphorylation of AMPK by Upstream Kinases Is Required for Activity in Mammalian Cells. *Biochem. J.* 2017, 474, 3059–3073.

[9] Vazquez-Martin, A.; Vellon, L.; Quirós, P. M.; Cufí, S.; De Galarreta, E. R.; Ferraros, C. O.; Martin, A. G.; Martin-Castillo, B.; Lopez-Otin, C.; Menendez, J. A. Activation of AMP-Activated

Protein Kinase (AMPK) Provides a Metabolic Barrier to Reprogramming Somatic Cells into Stem Cells. *Cell Cycle* 2012, 11, 974–989.

[10] Cool, B.; Zinker, B.; Chiou, W.; Kifle, L.; Cao, N.; Perham, M.; Dickinson, R.; Adler, A.; Gagne, G.; Iyengar, R.; Zhao, G.; Marsh, K.; Kym, P.; Jung, P.; Camp, H. S.; Frevert, E. Identification and Characterization of a Small Molecule AMPK Activator That Treats Key Components of Type 2 Diabetes and the Metabolic Syndrome. *Cell Metab.* 2006, 3, 403–416.

[11] Langendorf, C. G.; Kemp, B. E. Choreography of AMPK activation. *Cell Res.* 2015, 25, 5–6.

[12] Xiao, B.; Sanders, M. J.; Carmena, D.; Bright, N. J.; Haire, L. F.; Underwood, E.; Patel, B. R.; Heath, R. B.; Walker, P. A.; Hallen, S.; Giordanetto, F.; Martin, S. R.; Carling, D.; Gamblin, S. J. Structural Basis of AMPK Regulation by Small Molecule Activators. *Nat. Commun.* 2013, 4, 1–10.

[13] Hardie, D. G. AMPK - Sensing Energy While Talking to Other Signaling Pathways. *Cell Metab.* 2014, 20, 939–952.

[14] Olivier, S.; Foretz, M.; Viollet, B. Promise and Challenges for Direct Small Molecule AMPK Activators. *Biochem. Pharmacol.* 2018, 153, 147–158.

[15] Ngoei, K. R. W.; Langendorf, C. G.; Ling, N. X. Y.; Hoque, A.; Varghese, S.; Camerino, M. A.; Walker, S. R.; Bozikis, Y. E.; Dite, T. A.; Ovens, A. J.; Smiles, W. J.; Jacobs, R.; Huang, H.; Parker, M. W.; Scott, J. W.; Rider, M. H.; Foitzik, R. C.; Kemp, B. E.; Baell, J. B.; Oakhill, J. S. Structural Determinants for Small-Molecule Activation of Skeletal Muscle AMPK $\alpha 2\beta 2\gamma 1$ by the Glucose Importer GLUT4. *Cell Chem. Biol.* 2018, 25, 728–737.

[16] Cokorinos, E. C.; Delmore, J.; Reyes, A. R.; Albuquerque, B.; Kjbsted, R.; Jrgensen, N. O.; Tran, J. L.; Jatkar, A.; Cialdea, K.; Esquejo, R. M.; Meissen, J.; Calabrese, M. F.; Cordes, J.; Moccia, R.; Tess, D.; Salatto, C. T.; Coskran, T. M.; Opsahl, A. C.; Flynn, D.; Blatnik, M.; Li, W.; Kindt, E.; Foretz, M.; Viollet, B.; Ward, J.; Kurumbail, R. G.; Kalgutkar, A. S.; Wojtaszewski, J. F. P.; Cameron, K. O.; Miller, R. A. Activation of Skeletal Muscle AMPK Promotes Glucose Disposal and Glucose Lowering in Non-Human Primates and Mice. *Cell Metab.* 2017, 25, 1147–1159.

[17] Scott, J. W.; Galic, S.; Graham, K. L.; Foitzik, R.; Ling, N. X. Y.; Dite, T. A.; Issa, S. M. A.; Langendorf, C. G.; Weng, Q. P.; Thomas, H. E.; Kay, T. W.; Bimberg, N. C.; Steinberg, G. R.; Kemp, B. E.; Oakhill, J. S. Inhibition of AMP-Activated Protein Kinase at the Allosteric Drug-Binding Site Promotes Islet Insulin Release. *Chem. Biol.* 2015, 22, 705–711.

[18] Aledavood, E.; Moraes, G.; Lameira, J.; Castro, A.; Luque, F. J.; Estarellas, C. Understanding the Mechanism of Direct Activation of AMP-Kinase: Toward a Fine Allosteric Tuning of the Kinase Activity. *J. Chem. Inf. Model.* 2019, 59, 2859–2870.

[19] Aledavood, E.; Forte, A.; Estarellas, C.; Luque, F. J. Structural basis of the selective activation of enzyme isoforms: Allosteric response to activators of $\beta 1$ - and $\beta 2$ -containing AMPK complexes. *Comput. Struct. Biotechnol. J.* 2021, 19, 3394–3406.

-
- [20] Xiao B, Sanders MJ, Carmena D, Bright NJ, Haire LF, Underwood E, et al. Structural Basis of AMPK Regulation by Small Molecule Activators. *Nat Commun* 2013;4:1–10.
- [21] Case DA, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham TE, Cruzeiro VWD. AMBER 2018. San Francisco: University of California; 2018.
- [22] Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins* 2010;78:1950–8.
- [23] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general AMBER force field. *J Comput Chem* 2004;25:1157–74.
- [24] Bayly CI, Cieplak P, Cornell W, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* 1993;97:10269–80.
- [25] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;79:926–35.
- [26] Joung IS, Cheatham TE. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* 2008;112:9020–41.
- [27] Amadei A, Linssen ABM, Berendsen HJC. Essential dynamics of proteins. *Proteins* 1993;17:412–25.
- [28] Shkurti A, Goni R, Andrio P, Breitmoser E, Bethune I, Orozco M, et al. PyPcazip: A PCA-based toolkit for compression and analysis of molecular simulation data. *SoftwareX* 2015;5:44–50.
- [29] Schlitter J. Estimation of Absolute And Relative Entropies Of Macromolecules Using The Covariance Matrix. *Chem Phys Lett* 1993;215:617–21.
- [30] Baron R, Hünenberger PH, McCammon JA. Absolute single-molecule entropies from quasi-harmonic analysis of microsecond molecular dynamics: correction terms and convergence properties. *J Chem Theory Comput* 2009;5:3150–60.
- [31] Harris SA, Gavathiotis E, Searle MS, Orozco M, Laughton CA. Cooperativity in drug-DNA recognition: a molecular dynamics study. *J Am Chem Soc* 2001;123:12658–63.
- [32] Hunenberger P, Mark A, van Gunsteren W. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *J Mol Biol* 1995;252:492–503.
- [33] O’Rourke KF, Gorman SD, Boehr DD. Biophysical and computational methods to analyze amino acid interaction networks in proteins. *Comput Struct Biotechnol J* 2016;14:245–51.
- [34] Van Wart AT, Durrant J, Votapka L, Amaro RE. Weighted implementation of suboptimal paths (WISP): an optimized algorithm and tool for dynamical network analysis. *J Chem Theory Comput* 2014;10:511–7.

3.1.4. Paper 4: “Novel Indolic AMPK Modulators Increase Nitric Oxide Release in Human Endothelial Cells”

Manuscript submitted

Marta Sanz-Gómez^{1#}, Elnaz Aledavood^{2#}, Marina Beroiz¹, Laura Lagartera³, Elena Vega-Martín¹, Marta Gil-Ortega⁴, Jose Cumella³, Concepción Pérez³, Francisco Javier Luque², Carolina Estarellas^{2*}, María S Fernández-Alfonso^{1*}, Ana Castro^{3*}

¹Instituto Pluridisciplinar and Facultad de Farmacia, Universidad Complutense de Madrid, Spain

²Departamento de Nutrición, Ciencias de la Alimentación y Gastronomía, Facultad de Farmacia y Ciencias de la Alimentación, Instituto de Biomedicina (IBUB) e Instituto de Química Teórica y Computacional (ICTQBUB), Universidad de Barcelona, Campus Torribera, Santa Coloma de Gramenet, Spain

³Instituto de Química Médica, IQM-CSIC, Spain

⁴Departamento de Ciencias Farmacéuticas y de la Salud, Facultad de Farmacia, Universidad San Pablo-CEU, Madrid, Spain

These authors contributed equally to this work

Novel indolic AMPK modulators increase nitric oxide release in human endothelial cells

Marta Sanz-Gómez^{1, 2#}, Elnaz Aledavood^{3#}, Marina Beroiz¹, Laura Lagartera², Elena Vega-Martín¹, Marta Gil-Ortega⁴, Jose Cumella², Concepción Pérez², Francisco Javier Luque³, Carolina Estarellas^{3*}, María S Fernández-Alfonso^{1*}, Ana Castro^{2*}

¹Instituto Pluridisciplinar and Facultad de Farmacia, Universidad Complutense de Madrid, Spain

²Instituto de Química Médica, IQM-CSIC, Spain

³Departamento de Nutrición, Ciencias de la Alimentación y Gastronomía, Facultad de Farmacia y Ciencias de la Alimentación, Instituto de Biomedicina (IBUB) e Instituto de Química Teórica y Computacional (ICTQBUB), Universidad de Barcelona, Campus Torribera, Santa Coloma de Gramenet, Spain

⁴Departamento de Ciencias Farmacéuticas y de la Salud, Facultad de Farmacia, Universidad San Pablo-CEU, Madrid, Spain

These authors contributed equally to this work.

*Correspondence: acastro@iqm.csic.es (A.C.); marisolf@ucm.es (M.S.F-A.); cestarellas@ub.edu (C.E.)

Abstract:

Endothelial adenosine monophosphate-activated protein kinase (AMPK) plays a critical role in the regulation of vascular tone through stimulating nitric oxide (NO) release in endothelial cells. Since obesity leads to endothelial dysfunction and AMPK dysregulation, AMPK activation might be an important strategy to restore vascular function in cardiometabolic alterations. Here, we report the identification of a novel AMPK modulator, the indolic derivative IND6, which shows affinity for AMPK α 1 β 1 γ 1, the primary AMPK isoform in human EA.Hy926 endothelial cells. IND6 shows inhibitory action of the enzymatic activity *in vitro*, but increases the levels of p-Thr¹⁷⁴AMPK, p-Ser¹¹⁷⁷eNOS and p-Ser⁷⁹ACC in EA.Hy926. This paradoxical finding might be explained by the ability of IND6 to act as a mixed-type inhibitor, but also to promote the enzyme activation by adopting two distinct binding modes at the ADaM site. Moreover, functional assays reveal that IND6 increased NO levels demonstrating a functional activation of the AMPK-eNOS endothelial pathway. This kinase inhibition profile, combined with the paradoxical AMPK activation in cells, suggests that these new chemical entities may constitute a valuable starting point for the development of new AMPK modulators with therapeutic potential for the treatment of vascular complications associated with obesity.

Keywords: AMPK modulators; endothelial function; nitric oxide; obesity; mixed-type inhibition; protein dynamics.

1. Introduction

Obesity prevalence has increased over the past decades and is now a major public health problem worldwide. It is associated with an enhanced risk of developing cardiometabolic diseases such as

hypertension, insulin resistance, type 2 diabetes mellitus, coronary artery disease, myocardial infarction, heart failure, and stroke [1]. Obesity involves changes in body composition as a consequence of an energetic imbalance in which caloric intake is higher than energy expenditure [1]. The AMP-activated protein kinase (AMPK) is a sensor of cellular energy status that is responsible for maintaining the energy balance after depletion of energy stores, switching off ATP-consuming anabolic pathways [2,3].

Endothelial AMPK plays a key role in the regulation of vascular function through the activation of the PI3K-Akt-endothelial nitric oxide synthase (eNOS) pathway and stimulation of nitric oxide (NO) release in endothelial cells [4-6]. Obesity leads to AMPK dysregulation and endothelial dysfunction, which is the first step in the progression of cardiovascular disease [7]. We have shown that caloric restriction in young Zucker *fa/fa* rats has cardiovascular benefits by reducing endothelial dysfunction through AMPK-PI3K-Akt-eNOS activation associated to a reduction in blood pressure, plasma triglyceride levels, and cardiac hypertrophy [5]. AMPK activation might be thus an important strategy to restore vascular function in cardiometabolic alterations.

AMPK is a heterotrimeric Ser/Thr kinase of 1188 amino acids (~132KDa), which is ubiquitously distributed. It is formed by three subunits: α ($\alpha1$ and $\alpha2$), β ($\beta1$ and $\beta2$) and γ ($\gamma1$, $\gamma2$ and $\gamma3$), which combine to give 12 different isoforms [8]. AMPK α is the catalytic subunit and contains a conventional kinase domain (α KD) located at the N-terminus of the protein, and a C-terminal domain required for interaction with the AMPK β subunit. The C-terminal region of AMPK α subunit forms a globular domain around which the C-terminal region of the AMPK β subunit is wrapped. The extreme terminus of the AMPK β subunit then forms an interaction with the AMPK γ subunit, so that the AMPK β subunit acts as the scaffold that bridges α and γ subunits. A carbohydrate-binding module (CBM), located within the central region of the AMPK β subunit, forms a binding site for allosteric activators, termed the allosteric drug and metabolite (ADaM) binding pocket. The AMPK γ subunit contains four repeats in tandem of a structural module called cystathionine β -synthase (CBS) motif. Every pair of CBS repeats provides binding sites for the regulatory adenine nucleotides AXP (X = M, D, T) [9].

The range of AMPK modulators has been gradually expanded over the last years [10,11], covering from AMP mimetics, such as AICAR and C2, to ADP mimetics, such as O304, which is able to protect against pThr172 or pThr174 dephosphorylation in AMPK $\alpha2$ and AMPK $\alpha1$ isoforms, respectively, without allosteric activation of AMPK [12]. On top, AMPK activator drugs such as A769661, 991, PF-739, etc., directly activate AMPK targeting the ADaM site [13]. Strikingly, other AMPK modulators with unexpected mechanisms of action have been recently described. MT47-100 [14] is an allosteric AMPK $\beta2$ inhibitor that simultaneously activates AMPK $\beta1$, whereas SU6656 [15] paradoxically activates AMPK signaling by directly binding at the catalytic site. All these results reveal the complex regulation of this kinase, but at the same time offer the opportunity to be exploited in the search for drugs with novel mechanisms of action.

In this work, we describe novel indolic compounds as modulators of endothelial AMPK. For this purpose, the binding mode of these compounds has been assessed by combining molecular dynamics (MD) simulations, enzymatic and Surface Plasmon Resonance (SPR) assays, together with functional activation studies targeting AMPK, eNOS and Acetyl CoA Carboxylase (ACC) phosphorylation as well as NO production in human endothelial cells (EA. Hy926). Our findings offer new possibilities for

regulating the modulation of endothelial AMPK, as well as exploring the therapeutic implications of this novel mechanism of action.

2. Results

2.1. Identification of IND6 as AMPK α 1 β 1 γ 1 modulator

Within our drug discovery program focused on the search for novel AMPK modulators [16], we combined surface plasmon resonance (SPR) assays with enzymatic activity studies with the aim to evaluate the ability of the indole derivatives (**IND6**, **IND7**, **IND8**, **IND11**) to bind to AMPK α 1 β 1 γ 1 (Table 1).

The synthetic route employed to prepare **IND6**, **IND7**, **IND8**, and **IND11** is depicted in Scheme 1. The iodation of intermediate **2** was obtained by reaction of 1H-indole 2-ethyl carboxylate (**1**) with N-iodosuccinimide in dry DMF. From the key intermediate **2**, aryl substituted indole derivatives **3-5** and **9** were prepared by reaction with the corresponding boronic acid derivative, Pd(dppf)Cl₂ and K₂CO₃. Compound **10** was prepared by the reaction of ethyl 3-(4-hydroxyphenyl)-1H-indole-2 carboxylate **9** with 4-(3-chloropropyl)morpholine hydrochloride. Further, KOH was employed to obtain carboxylic acid derivatives **IND6**, **IND7**, **IND8**, and **IND11**.

Compounds **IND6**, **IND7**, **IND8**, and **IND11** were screened for binding to the AMPK α 1 β 1 γ 1 at a single concentration (100 μ M) using the SPR technique. Abbott's product A-769662 was used as a positive binding control, and β -cyclodextrin, which binds preferentially to the AMPK β 2 isoform, as a negative binding control. SPR results indicated that binding to AMPK α 1 β 1 γ 1 of compounds **IND6**, **IND7**, **IND8**, and **IND11** was slightly weaker relative to A-769662.

In parallel, the effect of compounds **IND6**, **IND7**, **IND8**, and **IND11** on the enzymatic activity of AMPK was examined through a luminescent assay with the recombinant AMPK isoform α 1 β 1 γ 1. This assay evaluates the enzymatic activity of AMPK to phosphorylate the SAMS peptide substrate, using A-769662 as positive control. All compounds in the series were found to reduce the activity of AMPK α 1 β 1 γ 1 at 30 μ M and therefore may be initially considered inhibitors (Table 1). Based on these results, **IND6** was selected as representative compound to carry out a detailed evaluation of the biological effect on the enzyme activity.

2.2. Competition assays of the indolic compound IND6

To investigate the inhibitory mechanism of **IND6**, dose-response assays were performed using different ATP concentrations (from 20 to 1000 μ M) and two inhibitor concentrations (10 and 20 μ M). The Lineweaver–Burk plot of enzyme kinetics is depicted in Figure 1. The results suggest that **IND6** acts as a mixed-type inhibitor, as noted by the increase in the K_s for ATP and the decrease in V_{max} with increasing concentration of **IND6**. Let us note that this modality of enzyme inhibition was also

reported previously for SBI-0206965 [17]. Representation of the relationship between K_m/V_m and the concentration of **IND6** led to inhibition constants K_i and K_i' of 6.9 and 27.1 μM , respectively. These values are 27- and 30-fold higher than the K_i and K_i' values determined for SBI-0206965, respectively ($K_i = 0.26 \mu\text{M}$; $K_i' = 0.89 \mu\text{M}$; [17]). Overall, although both **IND6** and SBI-0206965 exhibit a mixed-type inhibition, which conceptually combines both competitive and uncompetitive inhibition, these results suggest that the inhibitory activity may reflect different mechanisms of action.

To further explore the **IND6** binding mode, SPR studies were performed at increasing concentrations (from 10 to 100 μM) of either SBI-0206965 or **IND6** (Figure 2). The sensorgrams showed a progressive increase in the binding, which was much larger for SBI-0206965 in agreement with its stronger inhibitory potency (Figure 2A and B). In a separate assay, **IND6** was injected in the presence of SBI-0206965, both at a 100 μM concentration. Under these conditions, the sensorgrams showed an additive effect between SBI-0206965 and **IND6** (Figure 2C), suggesting that binding to AMPK might involve distinct binding sites. When the injection of SBI-0206965 and **IND6** was performed in the presence of a high ATP concentration (200 μM) (Figure 2D), SBI-0206965 binding was notably reduced, whereas **IND6** binding was less sensible to the presence of ATP.

2.3. *IND6 promotes the phosphorylation and activation of AMPK and downstream targets in a concentration-independent manner.*

In order to examine the effect of **IND6** in AMPK phosphorylation and some of its targets, such as eNOS and ACC, human endothelial cells of the EA Hy.926 line were treated with **IND6** at different concentrations (0.01, 0.1, 1, 10 and 100 μM). Untreated cells were used as control (CT). AICAR (at 5 mM), a known AMPK canonic activator, and 2-deoxyglucose (2-DG, at 1mM), a caloric restriction mimetic, were used as positive controls. AICAR, once inside the cell, is phosphorylated by adenosine kinases and is converted to ZMP, an AMP mimetic, which binds the CBS sites in γ -AMPK [18]. On the other hand, 2-DG is a competitive inhibitor of glucose metabolism [14] since it is phosphorylated by hexokinase to DG- PO_4 , which is trapped in the cell unable to undergo further metabolism [6, 19].

IND6 significantly increased the level of p-Thr¹⁷⁴AMPK at concentrations of 0.01, 0.1, 1 and 10 μM in a concentration independent manner (Figure 3A-3B). The effect was comparable to the enhancement elicited by AICAR at 5 mM and 2-DG at 1mM. Since the increased phosphorylation of AMPK is in contrast with the mixed-inhibitor features observed in the enzymatic assays, we determined levels of p-Ser⁷⁹ACC and p-Ser¹¹⁷⁷eNOS, which are well-known targets of AMPK in the endothelium [20, 21]. **IND6** significantly increased the levels of both p-Ser⁷⁹ACC/tubulin and p-Ser¹¹⁷⁷eNOS/tubulin at concentrations of 0.01, 0.1, 1 and 10 μM in a concentration

independent manner (Figure 3A, 3C, 3D), suggesting that **IND6** promotes a functional activation of AMPK.

2.4. Binding mode of **IND6** to AMPK

Due to the mixed-type inhibition of **IND6**, we investigated the binding to the ATP-binding site using Molecular Dynamics (MD) simulations. Four independent simulations were run for the AMPK α 1 β 1 γ 1 complexes with **IND6** and with SBI-020695, which was used as reference system. The X-ray structure of the AMPK–SBI-020695 complex (PDB entry 6BX6) revealed that SBI-020695 occupies a pocket located between the N- and C-lobes and the hinge region of the enzyme, overlapping with the binding site of compound C, which is a competitive inhibitor of AMPK [22]. The results obtained from the different MD simulations showed a consistent picture, where SBI-020695 remains stably bound in the ATP-binding pocket in all simulations (Figure 4). In particular, binding is assisted by two hydrogen bonds between SBI-020695 and the main chain of \checkmark Val98, with distances (averaged for the four MD simulations) of 3.2 ± 0.2 Å and 2.9 ± 0.1 Å between the pyrimidine nitrogen and exocyclic nitrogen of the inhibitor and the amide NH and carbonyl oxygen of α Val98, respectively. Furthermore, the ligand is enclosed in the hydrophobic pocket shaped by residues α Leu 24, α Val32, α Ile79, α Met95, α Leu148 and α Ala158.

The competitive binding mode of **IND6** to the ATP-binding pocket was guided by the superposition with both staurosporine, SBI-020695 and compound C, taking advantage of their X-ray structures (PDB entries 4ZHX, 4CFE, 4CFE, 6BX6 and 3AQV; see Supplementary Material), which revealed the formation of hydrogen bonds between these compounds with the hinge region of the kinase, particularly involving residues α Val96 and α Glu94 (α Val98 and α Glu96 in AMPK α 1).

The MD simulations performed for the AMPK–**IND6** complex revealed larger fluctuations of the ligand in the binding pocket compared to SBI-020695 (Figure 4A). This trait can be attributed to the flexibility of the benzyloxy moiety as well as to the non-planarity of the central benzene ring relative to the indole ring, enhancing also the fluctuations of the P loop (Figure 4B). This is also reflected in the hydrogen bond distances formed between the indole NH group of **IND6** with the carbonyl oxygen of \checkmark Val98 (average distance of 3.4 ± 1.1 Å), and the carboxylate oxygen of **IND6** with the NH group of \checkmark Glu96 (average distance of 3.7 ± 0.9 Å), which are larger than those formed by SBI-020695. Overall, these traits agree with the 27-fold lower potency of **IND6** relative to SBI-020695 (see above).

Additional MD simulations were also performed to examine the binding of **IND6** to the ADaM site, which mediates the activation effect played by several small molecules, such as A-769662 [23]. The ligand was oriented taking advantage of the close alignment exhibited by activators such as A-769662, 991 and SC4 (see Supplementary

Material). These studies showed that **IND6** may adopt two distinct binding modes (Figure 5). In one case, **IND6** is deeply bound into the hydrophobic cavity of the ADaM site, and the carboxylate group forms salt bridge interactions with the protonated amino groups of α Lys31 and α Lys33 (average distances of 3.0 ± 0.4 and 3.5 ± 0.9 Å). It is worth noting that the top of the P-loop points to the N-terminus of the α C-helix, leaving the ATP-binding site accessible for the binding of ATP. Indeed, a significant fraction of the conformations sampled by the P-loop superpose well with the conformations adopted in the ternary complex formed by AMPK bound to A-769662 and ATP [24] (Figure 6). This suggests that **IND6** might mimic the role of A-769662 in this binding mode.

In the other binding mode, **IND6** protrudes from the ADaM site toward the α C-helix, sitting on the top of the P-loop (Figure 5B). This binding mode is assisted by electrostatic interactions between the carboxylate group and the protonated residues α Lys31 and α Lys53 (average distances of 4.9 ± 0.9 and 5.3 ± 1.1 Å), and a cation- π interaction between β Arg83 and the benzyloxy ring of **IND6** (average distance of 4.0 ± 0.5 Å). Remarkably, this binding mode imposes a structural distortion of the P-loop, which occludes the ATP-binding site (Figures 5B and 6), making it unable to accommodate ATP. Therefore, this binding mode might explain the non-competitive mechanism of the mixed-type inhibition.

2.5. *IND6 promotes NO production in human endothelial cells with a major potency than AICAR does.*

In order to determine whether the increase in p-Thr¹⁷⁴AMPK and p-Ser¹¹⁷⁷eNOS elicited by **IND6** translates to [5, 21] an increment of NO production, NO levels were determined in human endothelial cells EA Hy.926 in presence of 1 and 5 μ M **IND6**. Untreated cells were used as a control (CT). Cells treated with DMSO 0.01% (maximum concentration of solvent for **IND6** (5 μ M)) were used to exclude cytotoxic effects. AICAR (at 5 mM) was used as a positive activation control (see point 2.3). AICAR induces an increase in fluorescence of DAF-2T, which indirectly represents the release of NO. **IND6** (at both 1 and 5 μ M) induces a similar increase in DAF-2T fluorescence as AICAR, but at 5,000 x lower concentrations (Figure 7).

3. Discussion

The results present in this study point out that the novel indolic derivatives appear to act as paradoxical activators of the endothelial AMPK α 1 β 1 γ 1 in EA.Hy926 cells, although they exhibit a mixed-type inhibition in the enzymatic assays.

Dysregulation in the AMPK signaling pathway in over-nutrition and obesity contributes to the development of metabolic disorders and endothelial dysfunction [1], which is considered the first step in the progression of cardiovascular disease [2]. Reduced endothelial AMPK phosphorylation leads to down-regulation of the PI3K-Akt-eNOS pathway together with low rates of NO synthesis [5, 7].

Contrarily, activation of endothelial AMPK restores impaired endothelial function and normalizes systolic blood pressure through the stimulation of the PI3K-Akt-eNOS pathway [5, 25]. In this context, identification of new chemical entities that can activate endothelial AMPK could be of significant interest for the treatment of obesity-related disorders. Starting with the use of SPR techniques, we selected **IND6**, which exhibits one of the best affinity values against recombinant AMPK α 1 β 1 γ 1 (RU 13.2, 100 μ M). In parallel, the AMPK α 1 β 1 γ 1 enzymatic activity was assessed by means of a luminescent assay, which revealed inhibitory activity values in the micromolar range against all tested compounds (Table 1). Hence, **IND6** was subjected to enzymatic kinetic analysis to examine its competition with ATP. We varied both ATP and **IND6** concentrations with a constant concentration of the peptide substrate used in the enzymatic reaction. The double reciprocal plot of data (Figure 1) indicated that **IND6** behaves as a mixed-type AMPK inhibitor. A similar inhibition mode has been recently reported for SBI-0206965, which was described as type IIb AMPK inhibitor [17]. Furthermore, SPR sensorgrams showed that **IND6** and SBI-0206965 presented an additive effect on the binding to AMPK, suggesting that they may bind at different binding sites. Moreover, when we performed the same experiments in the presence of a high ATP concentration (200 μ M), there was a significant reduction in the ability of SBI-020696 to bind to AMPK, while the affinity of **IND6** for AMPK remained at large extent unchanged. All these results suggest that **IND6** and SBI-020696 may bind the ATP-binding site, thus leading to competitive inhibition of the enzyme, but also suggest that **IND6** may regulate the AMPK activity through binding to an additional pocket.

The biological effect of **IND6** may be explained from the distinct binding modes observed for **IND6** in the ADaM site, and the drastic influence exerted on the structural conformation of the P-loop. Thus, the similar arrangements observed for the P-loop when **IND6** is deeply inserted into the ADaM site and in the X-ray structure of the AMPK bound to A-769662 suggest that **IND6** may mimic the activating role attributed to this latter compound. Nevertheless, the structural distortion of the P-loop caused by the alternative binding mode, where **IND6** protrudes from the ADaM site, might explain the non-competitive component of the mixed-type inhibition, in conjunction with the direct competition exerted by **IND6** upon binding to the ATP-binding site. At this point, let us remark that the adoption of two partially overlapping binding modes at the ADaM site may be facilitated by the lack of direct interactions between **IND6** and α -Asp90, in contrast to A-769662, which was found to form a hydrogen bond interaction with this residue in our previous studies of the AMPK α 2 β 1-A769662 complex [24]. In fact, previous experimental studies demonstrated that the interaction between β Arg83, α Asp90 and A-769662 is crucial for the enzyme activation [10].

Despite the mixed-type inhibition, cellular assays showed that **IND6** promotes AMPK phosphorylation (% p-AMPK/tubulin vs. CT) in the human endothelial cell line, EA Hy.926, which expresses the α 1 β 1 isoform. This suggests a paradoxical activation of AMPK similar to the one described for the indolic compound SU6656, which seems to promote AMPK's LKB1 dependent phosphorylation [26, 27]. In any case, the activation of AMPK by **IND6** is functional since it translates to ACC, an ubiquitous AMPK target [20], as well as to eNOS and NO release. This is in accordance to the AMPK-dependent eNOS activation described in endothelial cells [5, 21], although an attenuated NO production in response to AMPK activation has also been reported [28]. The efficacy of **IND6** to promote

AMPK stimulation is comparable to AICAR (5mM) or 2-DG (1mM). To note, the apparent discrepancy between the concentrations of **IND6** tested in the *in vitro* assays with recombinant proteins (SPR or inhibition studies) and the active concentrations observed in cell culture, which are much lower than in the former case. This could be explained by the synergic effect of the intracellular machinery, which leads to signal amplification [29].

In summary, **IND6** binding profile provides a basis to rationalize the activating behavior of **IND6** in EA Hy.926 cells by increasing AMPK activity in a functional manner, as demonstrated by the increment in both ACC and eNOS phosphorylation, two well-known targets of p-Thr172/174AMPK [20, 21]. Moreover, the final result of this study shows that **IND6** increases NO levels, demonstrating again a functional activation of AMPK. Our findings provide evidence that **IND6** holds potential as treatment of vascular complications associated with obesity, where intracellular ATP levels are high due to the energy surplus. Furthermore, since AMPK activation enhances the efficacy of rapamycin in suppressing mTORC2 and induces apoptosis in human cancer cells [30, 31], we might not exclude that **IND6** could be also a candidate for anticancer treatments.

4. Methods

4.1. Chemistry

All reagents were of commercial quality. Solvents were dried and purified by standard methods. Analytical TLC was performed on aluminum sheets coated with a 0.2 mm layer of silica gel 60 F254. Silica gel 60 (230–400 mesh) was used for flash chromatography. Analytical HPLC-MS was performed on Waters equipment coupled to a single quadrupole ESI-MS (Waters Micromass ZQ 2000) using a reverse-phase SunFire C18 4.6 × 50 mm column (3.5 μm) at a flow rate of 1 mL/min and by using a diode array UV detector. Mixtures of CH₃CN and H₂O were used as mobile phase (gradient of 15–95% of acetonitrile in water, as indicated in each case). HRMS (EI+) was carried out on Agilent 6520 Accurate-Mass Q-TOF LC/MS equipment. NMR spectra were recorded on a Bruker-AVANCE 300, a Varian-INOVA 400 and VARIAN SYSTEM-500 spectrometer. Melting points were determined on a Mettler MP70 apparatus and are uncorrected.

For experimental details, description of all synthetic intermediates, and characterization of final compounds see the Supplementary Material.

Synthesis of **IND6**: 3-(4-(Benzyloxy)phenyl)-1H-indole-2-carboxylic acid: To a solution of ethyl 3-(4-(benzyloxy)phenyl)-1H-indole-2-carboxylate (120 mg, 0.30 mmol) in EtOH (30 mL), KOH (30 mg, 1.10 mmol) dissolved in H₂O (2 mL) was added. After stirring at 100°C for 18 hours, HCl 1N (2 mL) was added. The resultant solid was filtrated and dried under reduced pressure to give 20 mg (18%) of **IND6**. HPLC (SunFire): t_R = 9.51 min (gradient: 15–95% of acetonitrile in water). MS (ES, positive mode): 344 (M + H)⁺. Mp 194°C–196°C. ¹H NMR (300 MHz, DMSO-d₆) δ 12.79 (s, 1H), 11.73 (s, 1H),

7.55 – 7.22 (m, 10H), 7.12 – 7.01 (m, 3H), 5.16 (s, 2H). ¹³C NMR (75 MHz, DMSO-d₆) δ 163.2; 157.7; 137.6; 136.3; 131.9; 128.8; 128.2; 128.1; 127.4; 126.6; 125.0; 123.6; 122.1; 120.9; 120.5; 114.4; 112.9; 69.6. HRMS (EI⁺) m/z ([M]⁺) calcd for C₂₂H₁₇NO₃ 343.12175; found 343.12179.

4.2. Identification of the AMPK isoform expressed by human endothelial cells EAHy.926 by proteomics

In order to choose the AMPK isoform to carry out the kinase assay and affinity studies, we determined the isoform mainly expressed in EA Hy.926 cells. Cell lysates were analyzed in the proteomics unit of the UCM (CAI Técnicas Biológicas). For this purpose, protein precipitation and digestion with trypsin gel was carried out. Then, LC-MS/MS (Q-EXACTIVE) analysis of tryptic peptides was performed (massive analysis: Shotgun and directed analysis: Parallel Reaction Monitoring, PRM), obtaining an m/z MS-MS spectrum that was used for the subsequent database search and identification based on the results. After comparing the information obtained with the data dumped in the databases, it was concluded that the most abundant isoform of AMPK in these cells is AMPK α 1 β 1 with 99.99% accuracy. Although the identification of the subunit γ was not conclusive, we chose the recombinant isoform AMPK α 1 β 1 γ 1 for the affinity and kinase assay studies.

4.3. Binding studies by surface plasmon resonance (SPR)

SPR experiments were performed at 25°C with a Biacore X-100 apparatus (Biacore, GE) in HBS-EP (10mM HEPES, 150mM NaCl, 3mM EDTA), with 2% de DMSO, 0,05% Tween 20 and 200 μ M ATP when was required, at 25°C. The protein AMPk was immobilized on a CM5 sensor chip (Biacore, GE) following standard amine coupling method [32]. The carboxymethyl dextran surface of the flow cell 2 was activated with a 7-min injection of a 1:1 ratio of 0.4M EDC and 0.1M NHS. The protein was coupled to the surface with a 300s injection at several dilutions at 40 μ g/mL in 10mM sodium acetate, pH 5.0. The unreacted N-hydroxysuccinimide esters were quenched by a 7-min injection of 0.1 M ethanolamine-HCl (pH 8.0). The levels of immobilization were around 1000 RUs. Flow cell 1 treated as a flow cell 2 (amine coupling procedure) without protein was used as a reference. Prior to use 10mM stock solutions of compounds were diluted several times until 100 μ M final concentration in the running buffer. Typically, a series of different compounds was injected onto the sensor chip a flow rate of 30 μ L/min for a period of 100s followed by a dissociation period of 200s. After the dissociation process an extrawash treatment was made over the flow cells with a 50% DMSO solution. No regeneration was needed. For competition measurements between AMPK and different compounds the concentrations used in the mixture were 100 μ M for each one. Sensorgrams data were double-referenced and solvent corrected using the Biaevaluation X-100 software (Biacore, GE).

4.4. Kinase Assay

The AMPK (α 1/ β 1/ γ 1) Kinase Enzyme System from Promega (Catalog number V1921) was used to screen AMPK inhibitors following the ADP-Glo™ Kinase Assay (Catalog number V9021). The assays were performed in 96-well plates (final volume 20 μ L), the assay buffer contains 40mM Tris, 7.5, 20mM MgCl₂, 0.1mg/mL BSA and 50 μ M DTT. 4 μ L of inhibitor was added to each well (final concentration of DMSO did not exceed 1%), followed by 8 μ L of enzyme (30 ng), after 5 min incubation at R/T, 8 μ L of ATP (150 μ M final concentration) and SAMStide (0.2 μ g/ μ L) were added and incubate 60 min at room temperature, then ADP-Glo™ reagent (20 μ L) was added allowing to incubate for 40 min at room temperature. Behind the incubation, the kinase detection agent (40 μ L) was added and incubated for 30 min at room temperature. Finally, the luminescence was recorded using a FLUO star Optima (BMG Labtechnologies GmbH, Offenburg, Germany) multimode reader. The inhibition activities were calculated based on the maximum activity, measured in the absence of inhibitor.

For competition assays, the experiments were carried out at four different concentrations of ATP (20-1000 μ M) in the absence or presence of the inhibitors, at two concentrations. The results were presented as double reciprocal Lineweaver–Burk plots (1/V vs 1/[ATP]).

4.5 Molecular modelling simulations

Molecular Dynamics (MD) simulations were used to examine the binding mode of SBI-020695 and IND6 to AMPK, considering the binding to both the ATP-binding site and the ADaM site, which is implicated in the enzyme activation by small molecules, such as A-769662. The X-ray structure of AMPK in the PDB entry 6C9J [33], which consists of chains α ₁, β ₁ and γ ₁, was utilized to build up the protein, and the position of SBI-020695 in the ATP-binding pocket was defined after superposition with the X-ray structure 6BX6 [34], which contains SBI-020695 bound to AMPK (isoform \checkmark 2). To generate the IND6-bound complexes, SBI-020695 and A-769662 (PDB entry 4CFF) were replaced by IND6 in order to build up the simulated systems with IND6 in the ATP-binding and ADaM sites, respectively. Following our previous studies [23], the γ -subunit was not considered in MD simulations because it does not participate in the inhibition process by SBI and IND6. Moreover, due to the lack of information of C-terminal tails of α and β subunits in the X-ray structure, these parts were not treated in our simulated systems. Finally, Thr174 was simulated in the phosphorylated form (pThr174).

Simulations were performed using the AMBER18 package [35] and the Amber ff99SB-ildn force field [36] for the protein, whereas the ligand (SBI, IND6) was parameterized using the GAFF force field in conjunction with restrained electrostatic

potential-fitted (RESP) partial atomic charges derived from B3LYP/6-31G(d) [37] calculations. The two simulated systems were immersed in an octahedral box of TIP3P water molecules [38]. The final systems contained around 370 residues, the ligand, around 26000 water molecules, and one/two Na⁺ atom for the complexes with SBI/IND6, which were added to maintain the neutrality of the simulated systems.

Simulations were done in the NPT ensemble for equilibration and NVT for MD productions. The simulations for SBI and IND6 were performed for 4 independent replicas. The minimization of the two systems was performed refining the position of hydrogen atoms in the protein (2000 cycles of steepest descent algorithm followed by 8000 cycles of conjugate gradient), and subsequently of the whole system (4000 cycles for steepest descent and 1000 cycles of conjugate gradient). Then, the temperature of the system was gradually increased from 100 to 300 K in 5 steps (50 ps each) using the NVT ensemble, followed by an additional 5 ns step performed in the NPT ensemble to equilibrate the density of the system. In this process, restraints were imposed to avoid artefactual changes in the hydrogen bonds between the ligand with Val98, as well as between pThr174 and Arg140. Production MD simulations were run for 250 ns per replica, leading to a total simulation time of 1.0 μ s per ligand. Restraints were gradually eliminated during the first 100 ns in order to avoid changes in the ligand binding mode due to structural fluctuations in the ATP-binding pocket, and the analysis of the trajectories was performed on the snapshots sampled in the last 150 ns unrestrained MD simulation.

4.6. Human endothelial cell cultures and treatments

Cell culture studies were performed on a human endothelium cell line (EA.Hy926). Briefly, cells were seeded in 75cm² flasks with Dulbecco's modification of Eagle's High Glucose Medium (DMEM, Biowest), 10% fetal bovine serum (FBS, Biowest), penicillin (100U/mL) and streptomycin (100 μ g/mL) (Biowest) and kept at 37°C in a humid atmosphere and 5% CO₂. Once subconfluent (70-80%) subcultures were carried out using trypsin-EDTA (0.25%, Sigma) and used between passages 3 and 7. Cells were incubated for 1h with either the AMPK activator AICAR (5mM), IND6 (0.01, 0.1, 1, 10 and 100 μ M); or 10 minutes with 2-DG (1mM). There were also control cells (CT) without treatment. All the treatments were kept at 37°C in a humid atmosphere and 5% CO₂.

4.7. Determination of the phospho-proteins by WB

The expression of the phosphorylated forms of: AMPK α in the residue Thr174 (pThr¹⁷⁴-AMPK α ; 62 KDa), ACC in the residue Ser79 (pSer⁷⁹-ACC; 265KDa), eNOS in the residue Ser 1177 (p-Ser¹¹⁷⁷eNOS; 133KDa) and tubulin (55 KDa) were determined by WB (as described in Supplementary Material) in EA.Hy926 cells lisates. Cells were seeded in 6-well plates (Sarstedt) with a density of 120,000 cells/well. After 48h, once

subconfluent, they were treated with modulating compounds (AICAR at 5 mM, 2-deoxyglucose at 1mM, and IND6 at 0.01, 0.1, 1, 10 and 100 μ M) and in control untreated cells (CT). All the treatments were kept at 37°C in a humid atmosphere and 5% CO₂ for 1 hour, except of the 2-deoxyglucose, which were maintained only during 10 minutes.

For the detection of all the proteins, acrylamide gels at 7% [H₂O (5.1 mL); 1.5M Tris-HCl pH=8.8 (2.5 mL); SDS at 20% (50 μ L) were used, acrylamide/bisacrylamide 30% (2.3mL); ammonium persulphate 10% (50 μ L), TEMED (5 μ L)]. Polyclonal rabbit antibodies against the pSer⁷⁹-ACC, pThr¹⁷⁴-AMPK (1:500, Cell Signaling Technology) and pSer¹¹⁷⁷-eNOS (1:500, EMD Millipore Corporation) were used as primary antibodies and an anti-rabbit antibody (IgG) marked with peroxidase (1:2000, Santa Cruz Biotechnology) was used as secondary antibody. Tubulin was detected with a monoclonal mouse antibody (1:5,000, Abcam) as primary antibody and an anti-mouse antibody (IgG) marked with peroxidase (1:10,000, GE Healthcare) as secondary. Quantification was carried out by establishing the relationship between the phosphorylated form of the different proteins and tubulin (p-protein/tubulin) based on the concentration of the modulating compounds administered, in addition to the negative control, on which no treatment was performed.

4.8. Detection of nitric oxide (NO) by fluorescence microscopy

EA. Hy926 cells were seeded on 8-well plates (Sarstedt) at a density of 6000 cells per well and allowed to grow in DMEM until a 60-70% confluence was reached. The medium was then aspirated and the cells for NO detection were incubated for 1h with no treatment (control, CT), with DMSO 0.01% (maximum solvent concentration achieved with IND6 5 μ M), AICAR 5mM as a control of activation of NO production, IND6 1 μ M and IND6 5 μ M. Once the incubation time had elapsed, the medium was aspirated and the cells were incubated with 4,5-diaminofluorescein diacetate (Molecular Probes) (DAF-2DA 10⁻⁵ M) for 30 min in the dark, at 37°C in a humid atmosphere and 5% CO₂, in DMEM. The DAF-2DA is a non-fluorescent permeable probe capable of diffusing through the cell membrane. Inside the cell it is degraded by esterases to 4,5-diaminofluorescein-2 (2-DAF), which when reacting with intracellular NO gives rise to thiazolofluorescein (DAF-2T), capable of emitting green fluorescence (excitation wavelength 488 nm and emission wavelength 530 nm), so that the intensity of fluorescence emitted will be proportional to the production of NO in the cells. The medium was then removed, the wells were washed twice with PBS for 15 min and then incubated with 4',6-diamino-2-phenylindol (DAPI, 1 μ g/mL; Molecular Probes) for 15 min at room temperature and in the dark. The DAPI (excitation wavelength 405 nm and emission wavelength 430 nm) stains the cell nuclei that are visualized by fluorescence microscopy. The cells were fixed in paraformaldehyde (PFA) for 1h and washed 4 times with PBS 5min. They were kept in PBS in darkness and at 4°C until they were used for

fluorescence microscopy (LeicaDM 2 000 Led). The quantification of the fluorescence intensity was performed with the ImageJ software and the fluorescence intensity at 530 nm is represented with respect to the control, which are the untreated cells in percentage form.

4.9. Preparation of drugs

AICAR (Toronto Chemical Research) was prepared at a concentration of $1.3 \times 10^{-1} \text{M}$ in distilled water and kept at -20°C until use. 2-deoxyglucose (Sigma) was prepared at a concentration of 10^{-1}M in distilled water and used immediately. SBI-0206965 (Sigma-Aldrich) and the different modulator candidates were prepared in DMSO (Sigma). A stock solution (10^{-2}M) was prepared for each of them and stored at -20°C until use (maximum 3 months).

4.10. Statistical analysis

The results obtained were expressed as the arithmetic mean \pm the standard error of the arithmetic mean (S.M.E.). Comparisons of the results obtained between individual groups were made using Saphiro Wilk's analysis of variance followed by a one-way (parametric) ANOVA and a Dunnett test or a Student's t-test to analyze differences between experimental groups with the untreated control. Significantly different groups were considered when $p < 0.05$. GraphPad Prism 8 software (San Diego) was used for statistical analysis. The free software ImageJ was used for image analysis.

Author Contributions: Idea, F.J.L.; C.E.; M.S.F.A. and A.C.; performed the experiments and the data analysis, M.S.G.; E.A.; M.B.; L.L.; E.V.M.; M.G.O.; J.C.; C.P.; F.J.L.; C.E.; M.S.F.A. and A.C.; writing—review and editing, M.S.G.; F.J.L.; C.E.; M.S.F.A. and A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fundación Eugenio Rodríguez Pascual, the Spanish Ministerio de Economía y Competitividad (MDM-2017-0767), the Spanish Ministerio de Ciencia Innovación y Universidades (RTI2018-095544-B-I00) and the Generalitat de Catalunya (2017SGR1746) for financial support. Computational resources from the Barcelona Supercomputing Center (BSC; BCV-2019-1-0009 and BCV-2019-2-0017) and the Consorci de Serveis Universitaris de Catalunya (CSUC; Molecular Recognition project) are acknowledged. E.A. thanks AGAUR (Generalitat of Catalunya; 2019FI_B2_00001) for her fellowship.

Acknowledgments: “The proteomic analysis was performed in the Proteomics Unit of Complutense University of Madrid, a member of ProteoRed and is supported by grant PT17/0019, of the PE I+D+i 2013- 2016, funded by ISCIII and ERDF”.

Competing interests

The authors declare no competing interests.

Abbreviations

2T-DAF: thiazolofluorescein (2T-DAF)
2-DAF: 4.5- diaminofluorescein-2

2-DG:	2-deoxyglucose
ACC:	Acetyl-CoA Carboxylase
ADaM:	allosteric drug and metabolite binding site
ADP:	adenosine diphosphate
AICAR:	5-aminoimidazole-4-carboxamide riboside
Akt:	protein kinase B
AMP:	5'-adenosine monophosphate
AMPK:	protein kinase activated by 5'-adenosine monophosphate
ANOVA:	analysis of variance
ATP:	5'-adenosine triphosphate
α AID:	self-inhibiting domain of the alpha subunit of AMPK α
α KD:	domain kinase of the alpha subunit of AMPK α
CBM:	carbohydrate binding site of AMPK β
CBS:	Cystathionine β -synthase
CM5:	carboxymethyl dextran 5
DAPI:	4',6-diamino-2-phenylindol
DG-PO4:	2-deoxyglucose phosphorylated by hexokinase
DMF:	Dimethylformamide
DMSO:	dimethyl sulfoxide
EA. Hy926:	human endothelial cell line
EDC:	1-ethyl-3-(3-dimethylpropanopropyl)-carbodiimide
eNOS:	Endothelial nitric oxide synthase
K _i :	inhibition constant
K _i ':	apparent inhibition constant
K _s :	dissociation constant
LC MS/MS:	proteomic technique that consists of performing a liquid chromatography followed by two consecutive mass spectrometry analyses
LKB1:	Liver Kinase B1
MD:	Molecular Dynamics
mTOR:	mamalian/mechanistic target of ramamycin
NHS:	N-hydroxysuccinamide
NO:	nitric oxide
p-ACC:	Acetyl-CoA Carboxylase phosphorylated in Ser79
p-AMPK:	AMPK phosphorylated in Thr174
p-eNOS:	Endothelial Nitric Oxide Synthase phosphorylated in Ser1177
PFA:	paraformaldehyde
PI3K:	phosphatidyl inositol 3 kinase
RU:	resonance units
SAMS:	Substrate for AMP-activated kinases
S.E.M.:	Standard Error of Mean

S.D.: Standard Deviation
SPR: Surface Plasmon Resonance
WB: Western Blot or Immune Detection.

References

1. Morris, M.J. Cardiovascular and metabolic effects of obesity. *Clin. Exp. Pharmacol. Physiol.* **2008**, *35*, 416–419. <https://doi.org/10.1111/j.1440-1681.2008.04912.x>
2. Hardie, D.G.; Carling, D.; Carlson, M. The AMP-activated/snfl protein kinase subfamily: metabolic sensors of the eukaryotic cell? *Annu. Rev. Biochem.* **1998**, *67*, 821–855. <https://doi.org/10.1146/annurev.biochem.67.1.821>
3. Kahn, B.B.; Alquier, T.; Carling, D.; Hardie, D.G.; AMP-activated protein kinase: Ancient energy gauge provides clues to modern understanding of metabolism. *Cell Metab.* **2005**. <https://doi.org/10.1016/j.cmet.2004.12.003>
4. García-Prieto, C.F.; Gil-Ortega, M.; Aránguez, I.; Ortiz-Besoain, M.; Somoza, B.; Fernández-Alfonso, M.S. Vascular AMPK as an attractive target in the treatment of vascular complications of obesity. *Vascul. Pharmacol.* **2015a**, <https://doi.org/10.1016/j.vph.2015.02.017>
5. García-Prieto, C.F.; Pulido-Olmo, H.; Ruiz-Hurtado, G.; Gil-Ortega, M.; Aranguez, I.; Rubio, M.A.; Ruiz-Gayo, M.; Somoza, B.; Fernández-Alfonso, M.S. Mild caloric restriction reduces blood pressure and activates endothelial AMPK-PI3K-Akt-eNOS pathway in obese Zucker rats. *Vascul. Pharmacol.* **2015b**. <https://doi.org/10.1016/j.vph.2014.12.001>
6. García-Prieto, C.F.; Gil-Ortega, M.; Plaza, A.; Javier Manzano-Lista, F.; González-Blázquez, R.; Alcalá, M.; Rodríguez-Rodríguez, P.; Viana, M.; Aránguez, I.; Gollasch, M.; Somoza, B.; Fernández-Alfonso, M.S. Caloric restriction induces H₂O₂ formation as a trigger of AMPK-eNOS-NO pathway in obese rats: Role for CAMKII. *Free Radic. Biol. Med.* **2019**, *139*, 35–45. <https://doi.org/10.1016/j.freeradbiomed.2019.05.016>
7. Prieto, D.; Contreras, C.; Sánchez, A. Endothelial dysfunction, obesity and insulin resistance. *Curr. Vasc. Pharmacol.* **2014**, *12*(3):412-26. <https://doi.org/10.2174/1570161112666140423221008>
8. Ross, F.A.; MacKintosh, C.; Hardie, D.G. AMP-activated protein kinase: a cellular energy sensor that comes in 12 flavours. *FEBS J.* **2016**, *283*, 2987–3001. <https://doi.org/10.1111/febs.13698>
9. Hardie, D. G. Regulation of AMP-activated protein kinase by natural and synthetic activators. *Acta Pharm. Sin. B* **2016**, *6*, 1–19. <https://doi.org/10.1016/j.apsb.2015.06.002>
10. Xiao, B.; Sanders, M.J.; Carmena, D.; Bright, N.J.; Haire, L.F.; Underwood, E.; Patel, B.R.; Heath, R.B.; Walker, P.A.; Hallen, S.; Giordanetto, F.; Martin, S.R.; Carling, D.; Gamblin, S.J. Structural basis of AMPK regulation by small molecule activators. *Nat. Commun.* **2013**, *4*, 1–10. <https://doi.org/10.1038/ncomms4017>
11. Steinberg, G.R.; Carling, D. AMP-activated protein kinase: the current landscape for drug development. *Nat. Rev. Drug Discov.* **2019**. <https://doi.org/10.1038/s41573-019-0019-2>
12. Steneberg, P.; Lindahl, E.; Dahl, U.; Lidh, E.; Straseviciene, J.; Backlund, F.; Kjellkvist, E.; Berggren, E.; Lundberg, I.; Bergqvist, I.; Ericsson, M.; Eriksson, B.; Linde, K.; Westman, J.; Edlund, T.; Edlund, H. PAN-AMPK activator O304 improves glucose homeostasis and microvascular perfusion in mice and type 2 diabetes patients. *JCI Insight* **2018**, *3*, 1–19. <https://doi.org/10.1172/jci.insight.99114>
13. Giordanetto, F.; Karis, D. Direct AMP-activated protein kinase activators: A review of evidence from the patent literature. *Expert Opin. Ther. Pat.* **2012**, *22*, 1467–1477. <https://doi.org/10.1517/13543776.2012.743994>
14. Scott, J.W.; Galic, S.; Graham, K.L.; Foitzik, R.; Ling, N.X.Y.; Dite, T.A.; Issa, S.M.A.; Langendorf, C.G.; Weng, Q.P.; Thomas, H.E.; Kay, T.W.; Birnberg, N.C.; Steinberg, G.R.; Kemp, B.E.; Oakhill, J.S. Inhibition of AMP-Activated Protein Kinase at the Allosteric Drug-Binding Site Promotes Islet Insulin Release. *Chem. Biol.* **2015**, *22*, 705–711. <https://doi.org/10.1016/J.CHEMBIOL.2015.05.011>
15. Ross, F.A.; Hawley, S.A.; Auciello, F.R.; Gowans, G.J.; Atrih, A.; Lamont, D.J.; Hardie, D.G. Mechanisms of

- Paradoxical Activation of AMPK by the Kinase Inhibitors SU6656 and Sorafenib. *Cell Chem. Biol.* **2017**, *24*, 813–824.e4. <https://doi.org/10.1016/j.chembiol.2017.05.02117>.
16. WO 2016042194 A1. US 20170283378 A1. PCT/ES2015/070677. Indole derivatives for the prevention and/or treatment of diabetes and associated metabolic disorders.
 17. Dite, T.A.; Langendorf, C.G.; Hoque, A.; Galic, S.; Rebello, R.J.; Ovens, A.J.; Lindqvist, L.M.; Ngoei, K.R.W.; Ling, N.X.Y.; Furic, L.; Kemp, B.E.; Scott, J.W.; Oakhill, J.S. AMP-activated protein kinase selectively inhibited by the type II inhibitor SBI-0206965. *J. Biol. Chem.* **2018**, *293*, 8874–8885. <https://doi.org/10.1074/jbc.RA118.003547>
 18. Vincent M.F.; Bontemps F.; Van den Berghe G. Inhibition of glycolysis by 5-amino-4-imidazolecarboxamide riboside in isolated rat hepatocytes. *Biochem J.* **1992**, *281*:267–272. <https://doi.org/10.1042/bj2810267>
 19. Zhu Z.; Jiang W.; McGinley J.N.; Thompson H.J. 2-Deoxyglucose as an Energy Restriction Mimetic Agent: Effects on Mammary Carcinogenesis and on Mammary Tumor Cell Growth In vitro. *Cancer Research.* **2005**, *65* (15). <https://doi.org/10.1158/0008-5472.CAN-05-0453>
 20. Stahmann N; Woods A.; Carling D.; Heller R. Thrombin Activates AMP-Activated Protein Kinase in Endothelial Cells via a Pathway Involving Ca²⁺/Calmodulin-Dependent Protein Kinase Kinase β . *Molecular and cellular biology.* **2006**, *26*(16):5933–5945 <https://doi.org/10.1128/MCB.00383-06>
 21. Morrow V.A.; Fougelle F.; Connell J.M.C.; Petrie J.R.; Gould G.W.; Salt I.P. Direct activation of AMP-activated protein kinase stimulates nitric-oxide synthesis in human aortic endothelial cells. *J Biol Chem.* **2003** *278*(34):31629–39. <https://doi.org/10.1074/jbc.M212831200>.
 22. Handa N.; Takagi T.; Saijo S.; Kishishita S.; Takaya D.; Toyama M.; Terada T.; Shirouzu M.; Suzuki A.; Lee S.; Yamauchi T.; Okada-Iwabu M.; Iwabu M.; Kadowaki T.; Minokoshi Y.; and Yokoyama S. Structural basis for compound C inhibition of the human AMP-activated protein kinase α 2 subunit kinase domain. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*:480–487 <https://doi.org/10.1107/S0907444911010201>
 23. Hardie D.G.; Schaffer B.E.; Brunet A. AMPK: An Energy-Sensing Pathway with Multiple Inputs and Outputs. *Trends Cell Biol.* **2016**, *26*(3):190–201. doi:10.1016/j.tcb.2015.10.013
 24. Aledavood E.; Moraes G.; Lameira J.; Castro A.; Luque F.J.; Estarellas C. Understanding the Mechanism of Direct Activation of AMP-Kinase: Toward a Fine Allosteric Tuning of the Kinase Activity. *J Chem Inf Model.* **2019**, *59*(6):2859–2870. <https://doi.org/10.1021/acs.jcim.8b00890>
 25. García-Prieto C.F.; Hernández-Nuño F.; Del Río D.; Ruiz-Hurtado G.; Aránguez I.; Ruiz-Gayo M.; Somoza B.; Fernández-Alfonso M.S. High fat diet induces endothelial dysfunction through a down-regulation of the endothelial AMPK-PI3K-Akt-eNOS pathway. *Mol Nutr Food Res.* **2015b**, *59*(3):520–532. <https://doi.org/10.1002/mnfr.201400539>
 26. Ross, F.A.; Hawley, S.A.; Auciello, F.R.; Gowans, G.J.; A Competing interests
- The author(s) declare no competing interests. trih, A.; Lamont, D.J.; Hardie, D.G. Mechanisms of Paradoxical Activation of AMPK by the Kinase Inhibitors SU6656 and Sorafenib. *Cell Chem. Biol.* **2017**, *24*, 813–824.e4. <https://doi.org/10.1016/j.chembiol.2017.05.021>
27. Langendorf C.G.; Scott J.W.; Kemp B.E. Fake Inhibitors: AMPK Activation Trumps Inhibition. *Cell Chemical Biology.* **2017**, *775*–777. <http://dx.doi.org/10.1016/j.chembiol.2017.07.003>
 28. Zippel, N.; Loot, A.E.; Stingl, H.; Randriamboavonjy, V.; Fleming, I; Fisslthaler, B. Endothelial AMP-Activated Kinase α 1 Phosphorylates eNOS on Thr495 and Decreases Endothelial NO Formation. *Int J Mol Sci.* 2018 Sep; *19*(9): 2753. doi: 10.3390/ijms19092753
 29. Raab, M.; Sanhaji, M; Pietsch, L; Béquignon, I.; Herbrand, A.K.; Süß, E.; Gande, S.L.; Caspar, B.; Kudlinzki, D; Saxena, K.; Sreeramulu, S.; Schwalbe, H.; Strebhardt, K.; Biondi, R.M. Modulation of the Allosteric Communication between the Polo-Box Domain and the Catalytic Domain in PLK1 by Small Compounds. *ACS*

Chem. Biol. **2018**, *13*, 1921–1931. <https://doi.org/10.1021/acscchembio.7b01078>

30. Mukhopadhyay, S.; Chatterjee, A.; Kogan, D.; Patel, D.; Foster, D.A. 5-aminoimidazole-4-carboxamide-1- β -D-ribofuranoside (AICAR) enhances the efficacy of rapamycin in human cancer cells. *Cell Cycle*. **2015a**, *14*, 3331–3339. <https://doi.org/10.1080/15384101.2015.1087623>
31. Mukhopadhyay, S.; Saqcena, M.; Chatterjee, A.; Garcia, A.; Frias, M.A.; Foster, D.A. Reciprocal regulation of AMP-activated protein kinase and phospholipase D. *J. Biol. Chem.* **2015b**, *290*, 6986–6993. <https://doi.org/10.1074/jbc.M114.622571>
32. Johnsson, B.; Löfås, S.; Lindquist, G. Immobilization of proteins to a carboxymethyl-dextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Anal. Biochem.* **1991**, *198*, 268–277. [https://doi.org/10.1016/0003-2697\(91\)90424-R](https://doi.org/10.1016/0003-2697(91)90424-R)
33. Yan, Y.; Edward Zhou, X.; Novick, S. J.; Shaw, S. J.; Li, Y.; Brunzelle, J. S.; Hitoshi, Y.; Griffin, P. R.; Eric Xu, H.; Melcher, K. Structures of AMP-Activated Protein Kinase Bound to Novel Pharmacological Activators in Phosphorylated, Non-Phosphorylated, and Nucleotide-Free States. *J. Biol. Chem.* **2019**, *294* (3), 953–967. <https://doi.org/10.1074/jbc.RA118.004883>.
34. Dite, T. A.; Langendorf, C. G.; Hoque, A.; Galic, S.; Rebello, R. J.; Ovens, A. J.; Lindqvist, L. M.; Ngoei, K. R. W.; Ling, N. X. Y.; Furic, L.; et al. AMP-Activated Protein Kinase Selectively Inhibited by the Type II Inhibitor SBI-0206965. *J. Biol. Chem.* **2018**, *293* (23), 8874–8885. <https://doi.org/10.1074/jbc.RA118.003547>.
35. Case, D. A. Amber 18. *Univ. California, San Fr.* **2018**.
36. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber Ff99SB Protein Force Field. *Proteins Struct. Funct. Bioinforma.* **2010**, *78* (8), 1950–1958. <https://doi.org/10.1002/prot.22711>.
37. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280. <https://doi.org/10.1021/j100142a004>.
38. 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.* **1983**, *79* (2), 926–935. <https://doi.org/10.1063/1.445869>.
39. Hawley S.A.; Fyffe F.A.; Russell F.M.; Gowans G.J.; Hardie D.G. Intact Cell Assays to Monitor AMPK and Determine the Contribution of the AMP-Binding or ADaM Sites to Activation. *Methods Mol Biol.* **2018**, 1732:239-253. https://doi.org/10.1007/978-1-4939-7598-3_16.

Scheme 1: (i) iodosuccinimide, DMF, 0°C, 1h; (ii) toluene:ethanol:water:1,4-dioxane (1:3:6:10), R-B(OH)₂, Pd (PPh₃)₄, K₂CO₃, 85°C, 18h; (iii) K₂CO₃, EtOH, 85°C, 18h (iv) KOH, EtOH, 100°C, 18h.

Table 1. AMPK α 1: β 1: γ 1 inhibitory activity and binding results measured by SPR (RU) of indole derivatives.


Figure 1. Lineweaver-Burk (double reciprocal) plot showing inhibition of AMPK(✓ 1 ) phosphorylation of SAMS peptide at two fixed IND6 concentrations (10 and 20 μ M). Assays were performed in the presence of varying ATP concentrations (20–1000 μ M) and the substrate concentration was kept constant at 0.2 μ g/ μ L. Each point is the mean of two different experiments, each one analyzed in triplicate.

Figure 2. Typical sets of SPR binding curves for the interaction of AMPK α 1 β 1 γ 1 with **A)** IND6, **B)** SBI-0206965, **C)** SBI-0206965 plus IND6, and **D)** SBI-0206965 plus IND6 in presence of ATP (200 μ M). Experiments were performed increasing the ligand concentration from 10 to 100 μ M (**A** and **B**) or at 100 μ M for competitive experiments (**C** and **D**). The concentration series of ligand solutions were injected over the surfaces in duplicate or quadruplicate.

Figure 3. IND6 increases p-AMPK, p-ACC, and p-eNOS levels in a concentration-independent manner. **A)** Representative immunoblots of p-Ser⁷⁹ACC, p-Ser¹¹⁷⁷eNOS, p-Thr¹⁷⁴AMPK and tubulin of EA Hy.926 human endothelial cells after the treatments. (p-Thr¹⁷⁴AMPK CT band was cropped from channel 9 of the gel to channel 1). **B)** Bar chart representation of the densitometry of the immunosensing bands (WB) expressed as the percentage of p-Thr¹⁷⁴AMPK/tubulin with respect to the control group. **C)** Bar chart representation of the densitometry of the immunosensing bands (WB) expressed as the percentage of p-Ser⁷⁹ACC/tubulin with respect to the control group. **D)** Bar chart representation of the densitometry of the immunosensing bands (WB) expressed as the percentage of p-Ser¹¹⁷⁷eNOS/tubulin with respect to the control group. *= $p < 0.05$; **= $p < 0.01$. $n = 4 - 5$. Data are presented as means \pm SEM.

Figure 4. Representation of the binding mode of SBI-020695 and IND6 to the ATP-binding site of AMPK. **A)** Superposition of the X-ray crystallographic structure of AMPK bound to SBI-020695 and a representative snapshot taken from the MD simulations. **B)** Representative snapshot of the MD simulation run for the AMPK-IND6 complex. The P-loop is highlighted in yellow. SBI-020695 and IND6 are shown in sticks. C atoms colored in cyan and orange shown the X-ray and final position of the ligand in MD simulation, respectively. Hydrogen atoms have been removed for the sake of clarity.

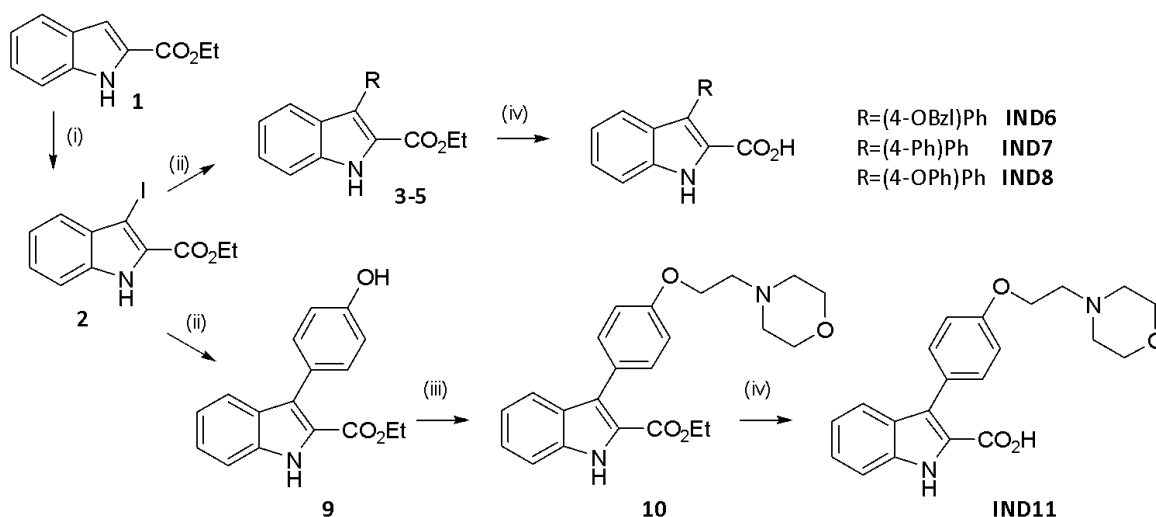
Figure 5. Superposition of the X-ray structure of the ternary complex formed by AMPK bound to A-769662 and ATP (C atoms in cyan and green, respectively; taken from [24]), and a representative snapshot of IND6 (C atoms in orange) bound to the ADaM site. **(A)** Activating binding mode. **(B)** Non-competitive binding mode. The P-loop is shown in yellow. Selected interactions with protein residues are indicated as dashed lines. Hydrogen atoms have been removed for the sake of clarity.

Figure 6. Distribution of the positional deviation (RMSD; \AA) of the backbone atoms that shape the ATP-binding along the trajectories run for the two binding modes of IND6 bound to AMPK \checkmark 1 $\text{\textcircled{1}}$ (shown in Figures 5A and 5B), and for the ternary complex between A-769662, ATP and AMPK \checkmark 2 $\text{\textcircled{1}}$ (data taken from [24]).

Figure 7. IND6 promotes an increase in DAF-2T fluorescence in human endothelial cells

(A) Representative fluorescence microscopy images of NO fluorescent indicator (DAF-2D incubation, first column, green) and DAPI (second column, blue) in cultured EA.Hy926 endothelial cells in basal conditions (CT) and after 1 hour treatment with DMSO 0.01%, AMPK activator AICAR (5 mM) and IND6 (1 and 5 μ M). (n=3). (B) Bar chart of the quantification of the intensity of green fluorescence (530 nm.) by densitometry represented as the percentage with respect to the control cells. *= $p < 0.05$, **= $p < 0.01$ and ***= $p < 0.001$; n=3. The data are presented as means \pm SEM. (C) Graphical abstract of the effect of **IND6** in the activation of AMPK and the subsequent increase of NO production because of eNOS activation by phosphorylation in Ser¹¹⁷⁷.

Scheme 1



Compound	% inhibition \pm SD at 30 μ M ^(a)	Binding 100 μ M (RU) ^(b)
AMPK α1:β1:γ1		

IND6	62.00±2.77	13.3
IND7	39.42±3.03	13.6
IND8	52.38±2.51	14.6
IND11	23.46±2.48	0.9

Table 1

(a) Data is the mean ± standard deviation (SD) of two independent experiments. Positive control: A-769662: % activation ± SD at 30 μM: 394±76
 (b) Binding (100 μM): Positive binding control: A-769662 (RU 31.7); Negative binding control: β-cyclodextrin (RU 1.6). Data is the mean ± SD of 1-6 independent experiments.

Figure 1

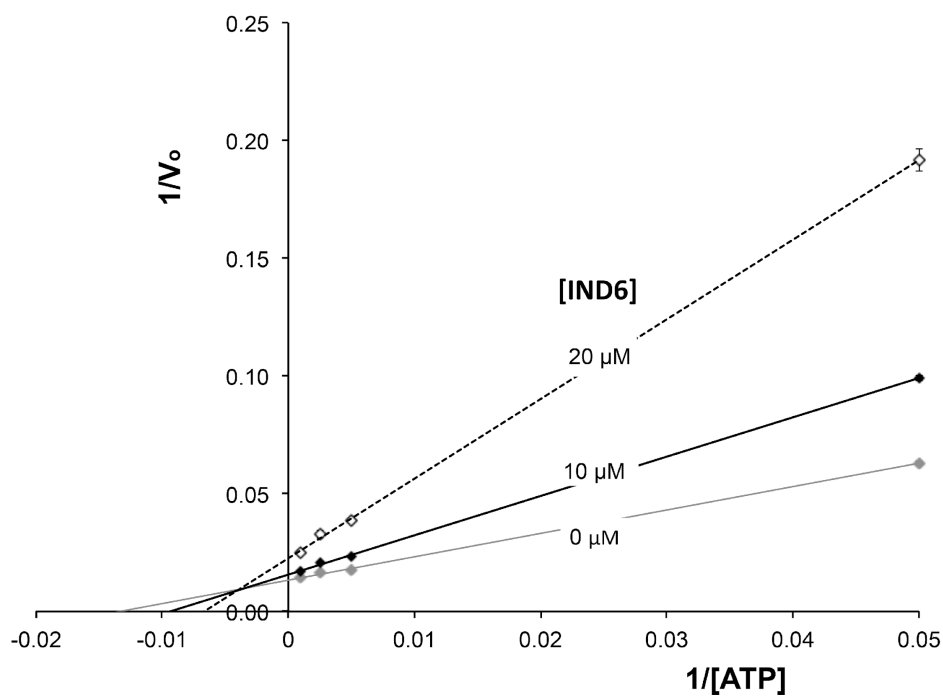


Figure 2

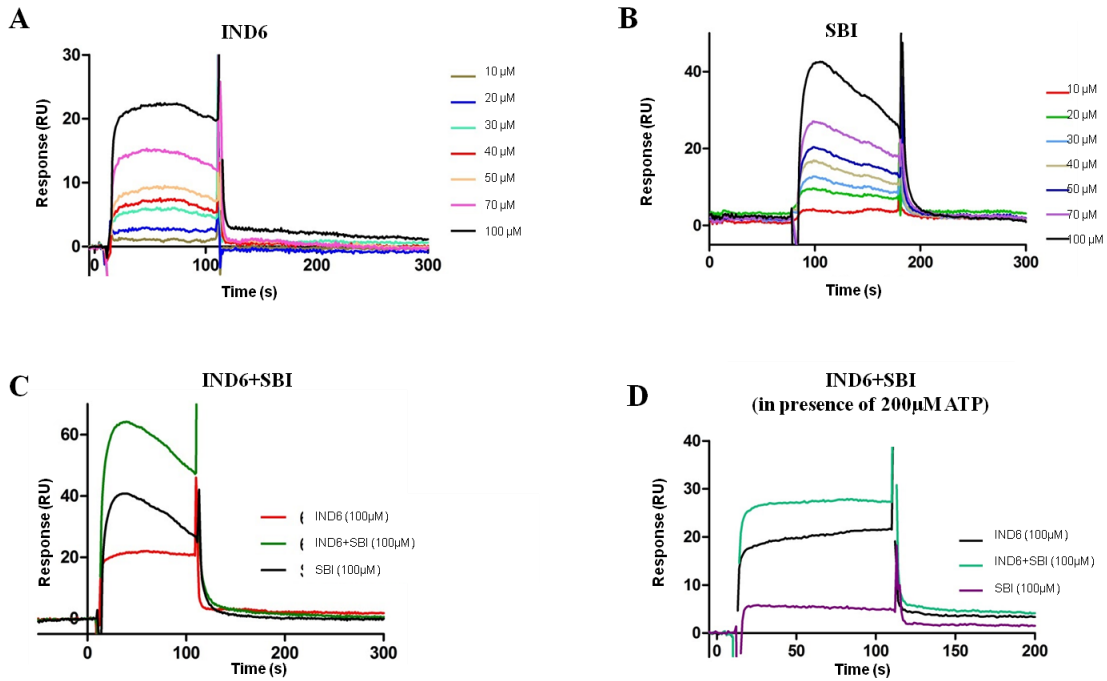


Figure 3

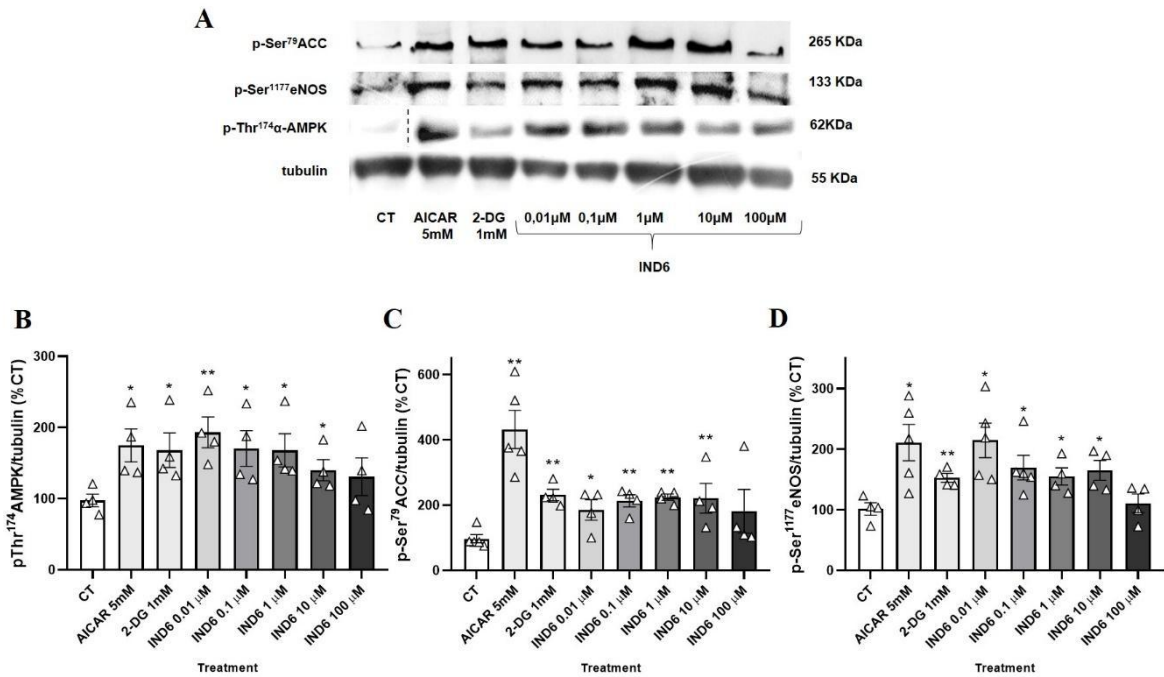


Figure 4

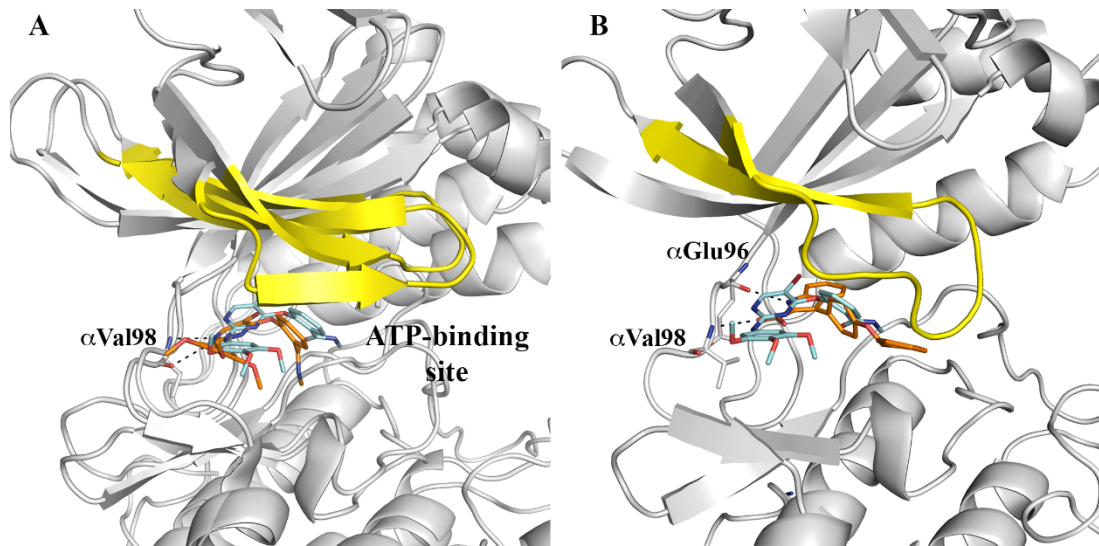


Figure 5

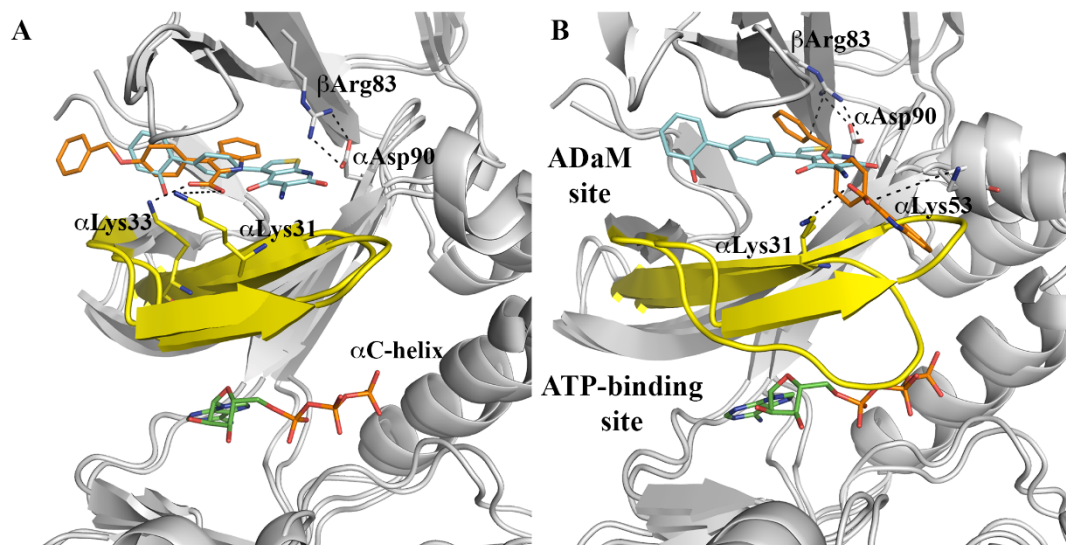


Figure 6

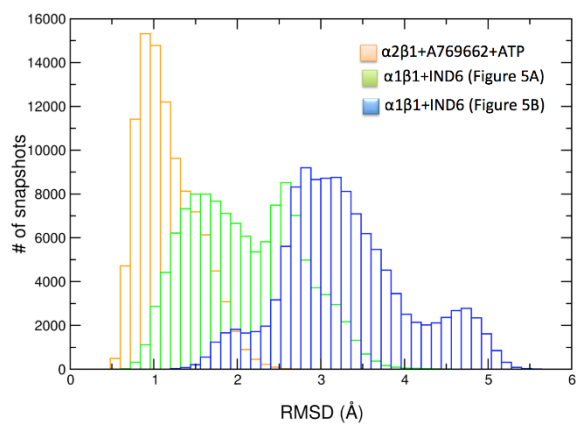
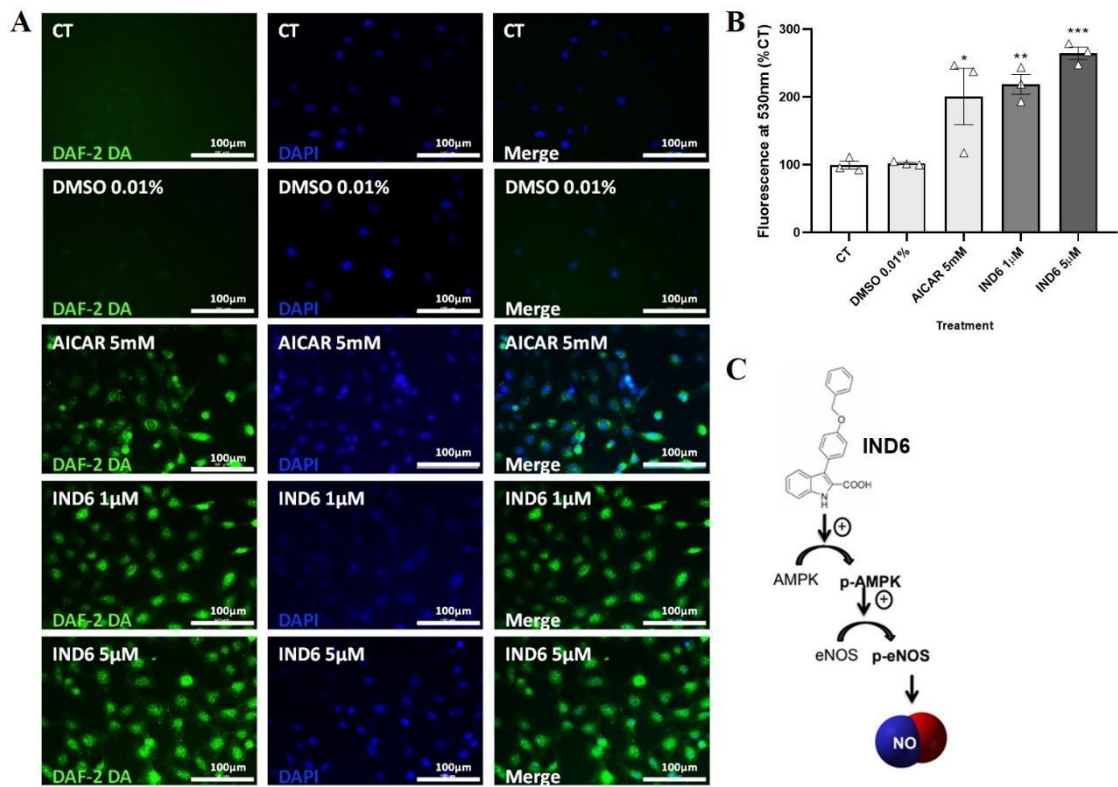


Figure 7



Supplementary information

Novel indolic AMPK modulators increase nitric oxide release in human endothelial cells

Marta Sanz-Gómez^{1, 2#}, Elnaz Aledavood^{3#}, Marina Beroiz¹, Laura Lagartera², Elena Vega-Martín¹, Marta Gil-Ortega⁴, Jose Cumella², Concepción Pérez², Francisco Javier Luque³, Carolina Estarellas^{3*}, María S Fernández-Alfonso^{1*}, Ana Castro^{2*}

¹Instituto Pluridisciplinar and Facultad de Farmacia, Universidad Complutense de Madrid, Spain

²Instituto de Química Médica, IQM-CSIC, Spain

³Departamento de Nutrición, Ciencias de la Alimentación y Gastronomía, Facultad de Farmacia y Ciencias de la Alimentación, Instituto de Biomedicina (IBUB) e Instituto de Química Teórica y Computacional (ICTQBUB), Universidad de Barcelona, Campus Torribera, Santa Coloma de Gramenet, Spain

⁴Departamento de Ciencias Farmacéuticas y de la Salud, Facultad de Farmacia, Universidad San Pablo-CEU, Madrid, Spain

These authors contributed equally to this work.

*Correspondence: acastro@iqm.csic.es (A.C.); marisolff@ucm.es (M.S.F-A.); cestarellas@ub.edu (C.E.)

Table of contents

Synthesis of synthetic intermediates and final compounds

- Synthesis of 3-Iodo-1H-indole-2-ethyl carboxylate (**2**)
- General synthesis for ethyl 3-aryl-1H-indole-2-carboxylates (**3-5** and **9**)
- Synthesis of ethyl 3-(4-(2-morpholinoethoxy)phenyl)-1H-indole-2-carboxylate (**10**)
- General procedure for the synthesis of the corresponding acids (**IND7**, **IND8**, **IND11**)

Western Blot detailed complementary methods

Original unedited Western Blot gels

All reagents were of commercial quality. Solvents were dried and purified by standard methods. Analytical TLC was performed on aluminum sheets coated with a 0.2 mm layer of silica gel 60 F254. Silica gel 60 (230–400 mesh) was used for flash chromatography. Analytical HPLC-MS was performed on Waters equipment coupled to a single quadrupole ESI-MS (Waters Micromass ZQ 2000) using a reverse-phase SunFire C18 4.6 × 50 mm column (3.5 μm) at a flow rate of 1 mL/min and by using a diode array UV detector. Mixtures of CH₃CN and H₂O were used as mobile phase (gradient of 15–95% of acetonitrile in water, as indicated in each case). HRMS (EI+) was carried out on Agilent 6520 Accurate-Mass Q-TOF LC/MS equipment. NMR spectra were recorded on a Bruker-AVANCE 300, a Varian-INOVA 400 and VARIAN SYSTEM-500 spectrometer. Melting points were determined on a Mettler MP70 apparatus and are uncorrected.

3-Iodo-1H-indole-2-ethyl carboxylate (2)

To a solution of NIS (2.25 g, 12 mmol) in DMF (20 mL) at 0°C, a solution of 1*H*-indole 2-ethyl carboxylate (1.89 g, 10 mmol) in DMF (15 mL) was slowly added. The mixture was stirred at room temperature for 1 hour. Then, a solution of sodium thiosulfate at 10% (5 mL) and water (10 mL) were added. The mixture was stirred at room temperature for another hour and the precipitate was collected by filtration to afford 2.9 g (96%) of **2** as a white solid. MS (ES, positive mode): 315 (M+H)⁺. ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.25 (s, 1H), 7.78 – 6.88 (m, 4H), 4.38 (q, *J* = 7.1 Hz, 2H), 1.39 (t, *J* = 7.1 Hz, 3H).

General synthesis for ethyl 3-aryl-1H-indole-2-carboxylates

To a solution of ethyl 3-iodo-1*H*-indole-2-carboxylate (**2**) (1 eq.), tetrakis (triphenylphosphine) palladium (0) (0.05 eq.) and the corresponding boronic acid (1.2 eq.) in a solvent mixture of toluene:ethanol:water:1,4-dioxane 1:3:6:10 (40 mL), a solution of K₂CO₃ (4 eq.) in water (5 mmol/mL) was slowly added while stirring under argon. After 18 hours at 85°C, the mixture was concentrated in vacuo, and water (20 mL) was added. The mixture was extracted with AcOEt (3 x 20 mL). The organic phase was dried over anhydrous sodium sulfate, filtered and concentrated in vacuo. The crude was purified by chromatographic column using Hex/AcOEt (4:1) as eluent to afford, after concentration and high vacuum-drying, the corresponding aryl products.

Ethyl 3-([1,1'-biphenyl]-4-yl)-1H-indole-2-carboxylate (3)

Prepared from **2** (300 mg, 0.95 mmol), tetrakis (triphenylphosphine) palladium (0) (54 mg, 0.47 mmol), [1,1'-biphenyl]-4-ylboronic acid (226 mg, 1.10 mmol). Column chromatography on silica gel (Hex/AcOEt 4:1) by following the general procedure described for ethyl 3-aryl-1*H*-indole-2-carboxylates. Yield 170 mg (53 %). HPLC (SunFire): *t*_R = 9.33 min (gradient: 15–95% of acetonitrile in water); MS (ES, positive mode): 342 (M + H)⁺. ¹H NMR (400 MHz, CDCl₃) δ 8.97 (s, 1H), 7.65 – 7.55 (m, 7H), 7.43 – 7.35 (m, 3H), 7.33 – 7.26 (m, 2H), 7.09 (ddd, *J* = 8.1, 6.9, 1.1 Hz, 1H), 4.25 (q, *J* = 7.1 Hz, 2H), 1.19 (t, *J* = 7.1 Hz, 3H).

Ethyl 3-(4-(benzyloxy)phenyl)-1H-indole-2-carboxylate (4)

Prepared from **2** (291 mg, 0.90 mmol), tetrakis (triphenylphosphine) palladium (0) (42 mg, 0.03 mmol) and 4-(benzyloxy)phenylboronic acid (250 mg, 1.10 mmol). Column chromatography on silica gel (Hex/AcOEt 4:1) by following the general procedure described for ethyl 3-aryl-1*H*-indole-2-carboxylates. Yield 315 mg (92 %). HPLC (SunFire): *t*_R = 11.0 min (gradient: 15–95% of acetonitrile in water); MS (ES, positive mode): 372 (M + H)⁺. ¹H NMR (300 MHz, CDCl₃) δ 9.23 (s, 1H), 7.64 – 6.80 (m, 13H), 4.99 (s, 2H), 4.19 (q, *J* = 7.1 Hz, 2H), 1.11 (t, *J* = 7.1 Hz, 3H).

Ethyl 3-(4-phenoxyphenyl)-1H-indole-2-carboxylate (5)

Prepared from **2** (1.0 g, 3.17 mmol), tetrakis (triphenylphosphine) palladium (0) (179 mg, 0.15 mmol), (4-phenoxyphenyl)boronic acid (814 mg, 3.80 mmol). Column chromatography on silica gel (Hex/AcOEt 4:1) by following the general procedure described for ethyl 3-aryl-1*H*-indole-2-carboxylates. Yield 891 mg (78 %). HPLC (SunFire): *t*_R = 9.30 min (gradient: 15–95% of acetonitrile in water). MS (ES, positive mode): 358 (M + H)⁺. ¹H NMR (300 MHz, DMSO-*d*₆) δ 11.90 (s, 1H), 9.34 (s, 1H), 7.55 – 7.49 (m, 4H), 7.47 – 7.39 (m, 2H), 7.35 – 7.28 (m, 1H), 7.21 – 7.05 (m, 4H), 6.91 – 6.74 (m, 1H), 4.24 (q, *J* = 7.1 Hz, 2H), 1.21 (t, *J* = 7.1 Hz, 3H).

Ethyl 3-(4-hydroxyphenyl)-1*H*-indole-2-carboxylate (9)

Prepared from **2** (1.0 g, 3.10 mmol), tetrakis (triphenylphosphine) palladium (0) (179 mg, 0.15 mmol), 4-hydroxyphenylboronic acid (524 mg, 0.38 mmol). Column chromatography on silica gel (Hex/AcOEt 4:1) by following the general procedure described for ethyl 3-aryl-1*H*-indole-2-carboxylates. Yield 350 mg (40 %). HPLC (SunFire): t_R = 7.83 min (gradient: 15–95% of acetonitrile in water). MS (ES, positive mode): 282 (M + H)⁺. ¹H NMR (300 MHz, DMSO-*d*₆) δ 11.75 (s, 1H), 9.46 (s, 1H), 7.52 – 7.45 (m, 2H), 7.34 – 7.25 (m, 3H), 7.10 – 7.04 (m, 1H), 6.87 – 6.81 (m, 1H), 4.23 (q, *J* = 7.1 Hz, 2H), 1.21 (t, *J* = 7.1 Hz, 3H).

Synthesis of ethyl 3-(4-(2-morpholinoethoxy)phenyl)-1*H*-indole-2-carboxylate (10)

To a solution of ethyl 3-(4-hydroxyphenyl)-1*H*-indole-2-carboxylate (**9**) (50 mg, 0.17 mmol.) in EtOH (10 mL) at 0°C, K₂CO₃ (73 mg, 0.51 mmol.) and 4-(3-chloropropyl)morpholine hydrochloride (42 mg, 0.23 mmol) were added. After 18 hours at reflux, water was added and the organic solvent was removed at reduced pressure. The organic phase was extracted with AcOEt (3 x 20 mL), and the organic extracts were washed with H₂O. After drying over MgSO₄, the solvent was removed to dryness and the resulting residue was purified by column chromatography, using AcOEt/Hex (1:4) as eluent, to afford 12 mg (18%) of the title compound. HPLC (SunFire): t_R = 5.18 min (gradient: 15–95% of acetonitrile in water). MS (ES, positive mode): 394 (M + H)⁺. ¹H NMR (300 MHz, CD₃OD) δ 8.56 (s, 1H), 7.54 – 7.37 (m, 4H), 7.28 (ddd, *J* = 8.2, 6.9, 1.1 Hz, 1H), 7.11 – 6.93 (m, 3H), 4.32 – 4.14 (m, 4H), 3.77 – 3.69 (m, 4H), 2.83 (t, *J* = 5.5 Hz, 2H), 2.68 – 2.57 (m, 4H), 1.25 (t, *J* = 7.1 Hz, 3H).

General procedure for the synthesis of the corresponding acids

To a solution of the corresponding ester (1 eq) in EtOH (30 mL), KOH (4 eq.) dissolved in H₂O (5 mmol/mL) was added. After stirring at 100°C for 18 hours, HCl 1N (2 mL) was added. The resultant solid was filtrated and dried under reduced pressure to give the corresponding acids.

3-([1,1'-Biphenyl]-4-yl)-1*H*-indole-2-carboxylic acid (IND7)

Prepared from **3** (100 mg, 0.30 mmol) and KOH (70 mg, 1.20 mmol) by following the general procedure described for 3-aryl-1*H*-indole-2-carboxylic acid. Yield 18 mg (18%). HPLC (SunFire): t_R = 9.71 min (gradient: 15–95% of acetonitrile in water). MS (ES, positive mode): 314 (M + H)⁺ Mp 219°C–222°C. ¹H NMR (300 MHz, DMSO-*d*₆) δ 11.24 (s, 1H), 7.80 – 7.26 (m, 11H), 7.11 (ddd, *J* = 8.1, 6.9, 1.2 Hz, 1H), 6.98 (ddd, *J* = 8.1, 6.9, 1.1 Hz, 1H). ¹³C NMR (75 MHz, DMSO-*d*₆) δ 140.8, 137.4, 135.4, 134.9, 131.5, 129.3, 127.9, 127.4, 126.79, 125.9, 122.8, 119.9, 119.6, 112.6. HRMS (EI+) *m/z* ([M]⁺) calcd for C₂₁H₁₅NO₃; 313.11053; found 313.11028.

3-(4-Phenoxyphenyl)-1*H*-indole-2-carboxylic acid (IND8)

Prepared from **7** (100 mg, 0.28 mmol) and KOH (62 mg, 1.10 mmol) by following the general procedure described for 3-aryl-1*H*-indole-2-carboxylic acid. Yield 30 mg (32%). HPLC (SunFire): t_R = 9.69 min (gradient: 15–95% of acetonitrile in water). MS (ES, positive mode): 330 (M + H)⁺. Mp 191°C–193°C. ¹H NMR (300 MHz, DMSO-*d*₆) δ 12.87 (s, 1H), 11.80 (s, 1H), 7.59 – 6.91 (m, 13H). ¹³C NMR (75 MHz, DMSO-*d*₆) δ 161.7, 157.08, 156.0, 136.5, 132.4, 130.4, 129.1, 127.1, 125.4, 123.8, 123.0, 122.2, 120.9, 120.8, 119.0, 118.3, 113.0, 60.6, 14.3. HRMS (EI+) *m/z* ([M]⁺) calcd for C₂₁H₁₅NO₃; 329.10656 found 329.10619

3-(4-(2-Morpholinoethoxy)phenyl)-1*H*-indole-2-carboxylic acid (IND 11)

Prepared from **10** (120 mg, 0.30 mmol) and KOH (67 mg, 1.20 mmol) by following the general procedure described for 3-aryl-1*H*-indole-2-carboxylic acid. Yield 32 mg (29%). HPLC (SunFire): t_R = 4.46 min (gradient: 15–95% of acetonitrile in water). MS (ES, positive mode): 367 (M + H)⁺. Mp 206°C-208 °C. ¹H NMR (500 MHz, DMSO-*d*₆) δ 11.75 (s, 1H), 7.54 – 7.38 (m, 4H), 7.26 (ddd, *J* = 8.2, 6.9, 1.2 Hz, 1H), 7.11 – 6.92 (m, 3H), 4.52 (s, 2H), 4.03 – 3.81 (m, 4H), 3.65 – 3.45 (m, 4H), 3.23 (s, 3H). ¹³C NMR (125 MHz, DMSO-*d*₆) δ 163.3, 136.4, 132.0, 127.5, 125.1, 123.8, 122.0, 120.9, 120.6, 114.5, 113.0, 63.6, 52.1. HRMS (EI⁺) *m/z* ([M]⁺) calcd for C₂₁H₂₂N₂O₄ 366.15957; found 366.15976.

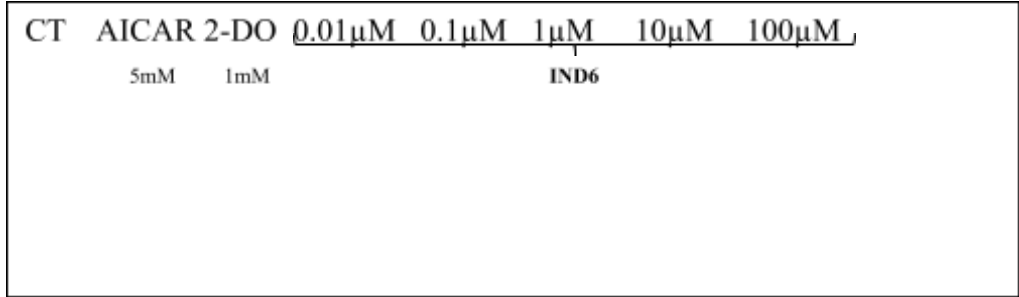
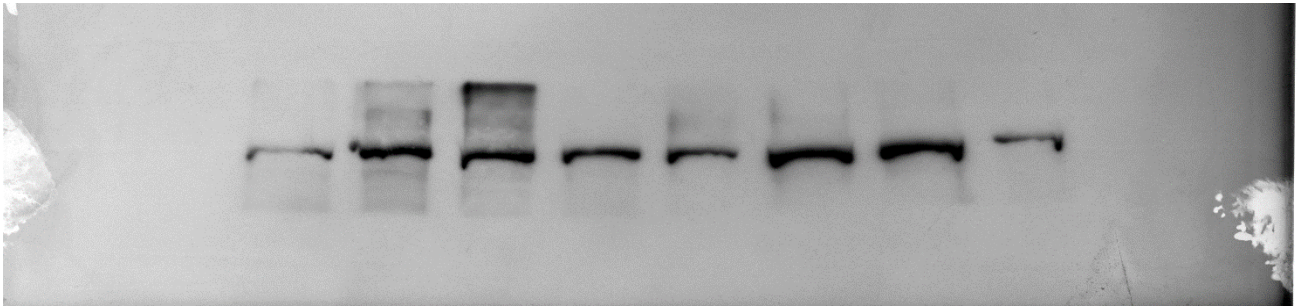
Western blot analysis

Adapted from Hawley *et al.* [39]. Western Blot (WB) studies were conducted in treated and CT EA. Hy926 cells dissolved in ice cold lysis buffer (50mM Na₄P₂O₇, 50mM NaF, 5mM EDTA, 5mM NaCl, 5mM EGTA, 10mM HEPES, 0.5% Triton, 1μg/mL leupeptin, 1μg/mL aprotinin, 0.5μL/mL N- α -p- tosyl-l-lysine in chloromethyl ketone and 20μM phenylmethylsulfonyl fluoride) and frozen immediately at -80°C until use. Just before gel running, they were centrifuged at 13,000 rpm at 4°C for 20min (MiniSpin Plus, Eppendorf) and the supernatant was collected. Protein concentration was determined by colorimetry using Protein Assay (Bio-Rad) in a 96-well plate. After an incubation of 5min at 25°C, absorbance was determined at a wavelength of 595nm. A standard albumin curve (SAB) was used to quantify the protein concentration. Protein concentration was adjusted to 1 mg/mL protein with Laemli [50mM Tris, (pH=6.8), 10% glycerol, 10% sodium dodecyl sulfate (SDS), 5% β -mercaptoethanol, 2mg/mL bromophenol blue] and separated by electrophoresis, using a MiniProtean3 (Bio-Rad) system. The samples were boiled in a bath at 100°C for 5min and loaded (17μg protein/well) into a 7% acrylamide gel. PrecisionPlusProteinTM All blue standard (Bio-Rad) was used as a molecular weight marker. Electrophoresis was performed under denaturing conditions at 150V for 70min at 25°C and in an electrode buffer (0.2 M glycine; 0.025M Tris and 0.1% SDS). Then, the gel proteins were transferred to a nitrocellulose membrane with a pore size of 0.2μm (Bio-Rad) through a transfer system at 25V for 10min for high molecular weight proteins (ACC) and for 7min for medium-high molecular weight proteins (<250KDa), such as eNOS, AMPK or tubulin (Trans-Blot Turbo Transfer, Bio-Rad). The membranes were washed twice for 5min in agitation with a washing solution [0.1% skimmed milk powder in phosphate buffer saline with 0.5% Tween 20 (PBS-T)]. In order to avoid non-specific binding, the membranes were incubated for 1h in agitation with a blocking solution (5% skimmed milk powder in PBS-T). Subsequently, they were washed 3 times for 5min and incubated at 4°C for 24h with the corresponding primary antibody at the concentration of 1:500 for p-ACC, p-eNOS and p-AMPK and 1:5,000 for tubulin. Primary antibodies were prepared in a specific solution (0.5% BSA; PBS-T and 0.05% azide). The following day, the membranes were washed 3 times for 5min and incubated for 1h at 25°C in agitation with the corresponding secondary antibody marked with peroxidase at the concentration of 1:4,000 for p-ACC and p-eNOS; 1:6,000 for p-AMPK and 1:10,000 for tubulin. This was prepared in a solution of 1% skimmed milk powder in PBS-T. Then, the membranes were washed again 3 times for 5min to remove the remains of the secondary antibody. The detection of the different proteins was carried out by chemiluminescence, using a detection kit (Enhanced Chemiluminescence ECL Prime, Amersham Bioscience) and subsequent exposure in a dark room for 1-2 minutes

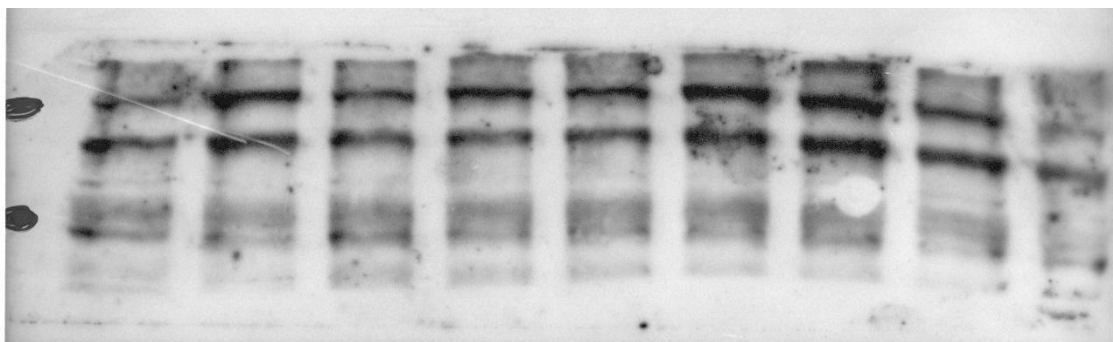
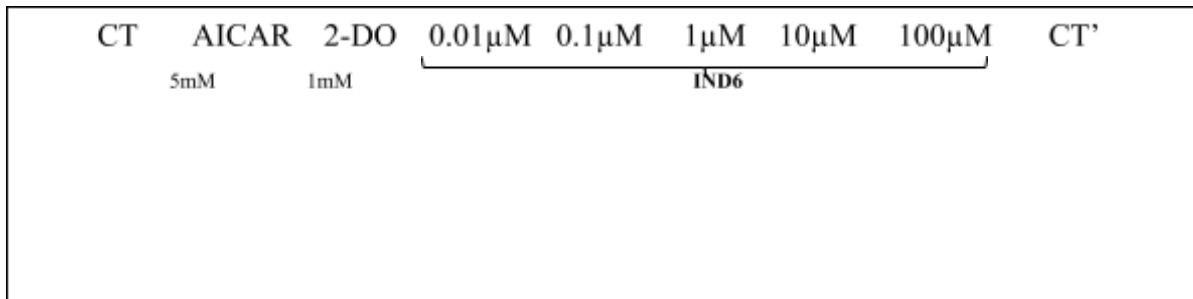
on autoradiographic film (GE Healthcare) for subsequent development with developer and fixer (Sigma). The quantification of the bands obtained by densitometry was carried out using ImageJ (Java Community Process) software of the scanner-digitalized films. Once the final image was merged, same contrast and bright corrections were performed in all the films in order to improve their interpretation.

Original unedited Western Blot gels

p-Ser⁷⁹ACC

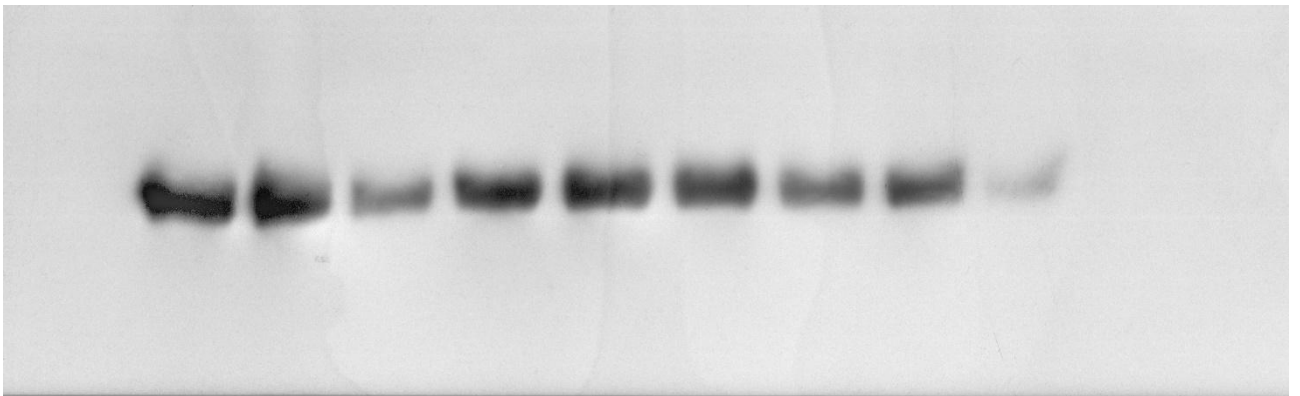


p-Ser¹¹⁷⁷eNOS



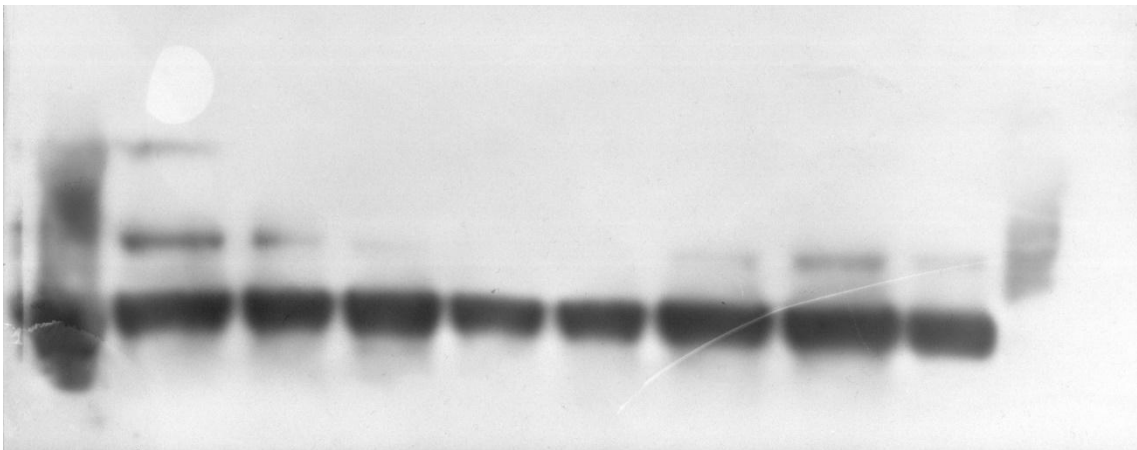
p-Thr¹⁷⁴AMPK

AICAR	AICAR'	2-DO	0.01μM	0.1μM	1μM	10μM	100μM	CT
5mM	5mM	1mM	IND6					



Tubulin

MW marker	CT	AICAR	2-DO	0.01 μ M	0.1 μ M	1 μ M	10 μ M	100 μ M
		5mM	1mM	IND6				



3.2. TrHbN

The fifth paper, entitled “*Searching for Novel Targets in Tuberculosis: Potential Reductase Partner in the Nitric Oxide Detoxification Activity by Mycobacterium tuberculosis Truncated Hemoglobin N*” led to the identification of a probable reductase partner for trHbN able to accomplish the redox reaction (section 4.2).

3.2.1. Paper 5: “Searching for Novel Targets in Tuberculosis: Potential Reductase Partner in the Nitric Oxide Detoxification Activity by Mycobacterium tuberculosis Truncated Hemoglobin N”

Manuscript under preparation

Elnaz Aledavood,[†] Carolina Estarellas,[†] and F. Javier Luque^{*,†}

[†]Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicina (IBUB) and Institute of Theoretical and Computational Chemistry (IQTCUB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain

Searching for novel targets in tuberculosis: Potential reductase partner in the nitric oxide detoxification activity by *Mycobacterium Tuberculosis* truncated hemoglobin N

Elnaz Aledavood,[§] Carolina Estarellas,[§] F. Javier Luque^{§,*}

[§]Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicine (IBUB) and Institute of Theoretical and Computational Chemistry (IQTUB), University of Barcelona, Santa Coloma de Gramenet 08921, Spain

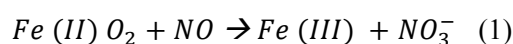
Abstract: Tuberculosis is one of the major worldwide health threats. Multidrug resistance, drug side effects and co-infections make it necessary to find novel antimycobacterial agents. In this study, we have turned our attention to one of the defense mechanisms used by *Mycobacterium tuberculosis* (Mtb), which involves the truncated hemoglobin N (trHbN), a protein that converts nitric oxide (NO) produced by macrophages into the harmless nitrate anion. For the turnover of this process Mtb requires a reductase partner involved in restoring the ferrous state of trHbN after conversion of NO to NO₃⁻, which to the best of our knowledge has not been identified yet. Our studies have led to three reductase candidates, and three-dimensional models for their interaction with trHbN. The critical role played by one of these reductases, *ndh*, which is essential for the respiratory chain, has been highlighted by experimental studies. Accordingly, we suggest that *ndh* appears to be a promising target for drug discovery targeting the nitrosative defense mechanisms of Mtb.

Key words: *Mycobacterium tuberculosis*, nitrosative stress, truncated hemoglobin N

Introduction

Mycobacterium tuberculosis (Mtb) is the causative agent of human tuberculosis (TB),¹ a disease that affects around ten million people each year and it is the ninth cause of death worldwide.² Despite the discovery of novel compounds with inhibitory properties against different targets, TB is still among the most challenging diseases around the world.^{3,4} The pathogenicity of Mtb is related to the remarkable adaptability to scavenge reactive nitrogen and oxygen species within the intracellular environment.⁵ In particular, nitric oxide (NO) produced by macrophages in early infections

would limit the survival of Mtb, but the toxic effect of NO can be reduced or even eliminated by an efficient defense mechanism carried out by the truncated hemoglobin N (trHbN),^{6,7} which converts NO to harmless nitrate anion (NO₃⁻). The NO dioxygenase (NOD) activity of trHbN is crucial in relieving the nitrosative stress (eq 1).⁸



The trHbN is a small heme protein (128 residues) encoded by the *glbN* gene.⁹ Its tertiary structure consists of a two-over-two α -helical fold.¹⁰ A one-turn A-helix that locates at the N-terminal region to the protein core, the short C-helix, the E-helix hosting residue LeuE7 (54) and the trHbN PheE15 (62) are among the functionally

relevant regions of the protein. Conserved Gly-based motifs located at AB, E-pre-F and pre-F-F regions facilitate the adoption of the trHbN fold (Figure 1).¹¹ HisF8 (81) is bound to the heme group and it is the only residue that is conserved between the Hb and trHb families.¹² The heme stabilization is achieved through hydrogen bonds to ThrCD4 (49) and Ala (75), and salt bridges to ArgE6 (53) and ArgFG3 (84).¹³

Mechanistic studies suggested that trHbN follows a dual-path mechanism to drive migration of O₂ and NO to the heme cavity via short and long branches of an inner tunnel.⁶ The O₂ reaches to the heme cavity through the tunnel short branch (~10 Å), which is shaped by residues in helices G and H. Binding to the heme facilitates opening of the long branch (~20 Å), which is formed by helices B and E. In this process, PheE15 plays the role of a gate. The opening of PheE15 involves O₂-induced conformational changes and also is affected by the movement of the N-terminal of the protein (pre-A motif).¹⁴ Excision of pre-A motif leads to distinctive changes in the protein dynamics, which cause the gate of the tunnel long branch to be trapped into a close conformation that impedes migration of NO toward the heme active site.¹⁴ As the NOD function of trHbN relies on the migration of NO to the O₂-bound heme via the long branch, PheE15 is a fundamental residue for the NO scavenging.^{15,16} Although trHbN is essential for NO resistance of Mtb, this protein does not contain typical druggability hallmarks, which limits its suitability for drug discovery.

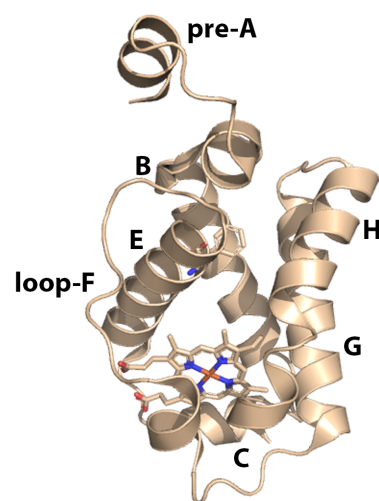


Figure 1. The location of the structural regions of trHbN (PDB: 1IDR). The heme and Phe15 are shown as sticks.

The NOD function also depends on the efficiency of the redox recycling needed to recover the ferrous state from the ferric species generated in the NO dioxygenase activity.^{17,18} Accordingly, we hypothesize that the reduction of HbN-Fe(III) to HbN-Fe(II) must be performed very efficiently to guarantee the survival of the bacillus in the hazardous environment of macrophages. Otherwise, it has no sense that trHbN may accomplish the chemical conversion of NO to nitrate anion if the recovery of the ferrous step would limit the overall rate of the turnover cycle. In fact, despite lacking a reductase domain, the NO dioxygenation efficiency of trHbN is high ($K_{\text{NOD}} \sim 745 \mu\text{M}^{-1} \text{S}^{-1}$) and it is comparable with the flavoHbs that contain a reductase domain. At present, however, there is no information about the reductase partner that can assist the electron recycling of trHbN.

To see if the survival of Mtb requires the expression of a specific reductase partner, Singh,

et al. carried out a research to confirm this hypothesis.¹ Since trHbN is functionally active in *Escherichia Coli* (E.coli), it should be able to use reductases to carry out NO dioxygenation. In this context, they examined the efficiency of different reductases in assisting the electron transfer process. The results pointed out a marked dependence of the efficiency of this process on the nature of the reductase. Specifically, this study showed that NADH-ferredoxin reductase (FdR) could perform efficiently the reduction of ferric to ferrous trHbN. This interaction is mediated by the flexible CD loop and the dynamical motion of the protein backbone is crucial for facilitating the interaction which is modulated by the pre-A region of trHbN.

In this context, this study was conducted with the aim to identify putative reductases able to accomplish the redox reaction required for efficient turnover of trHbN in Mtb and examine the suitability as potential targets to interfere with NOD scavenging ability of Mtb.

Computational Methods

Searching in databases. Since trHbN is found in the membrane and the periplasmic space,¹⁹ we performed a search for putative reductase attached to the membrane using the Tuberculist database²⁰ and an integrative approach that combined i) the structural and functional information available for trHbN, ii) a comprehensive bioinformatics analyses of the Mtb genome searching for reductase and electron transfer (ET) proteins, and iii) a detailed search of the data available regarding transcriptional changes in early phases of the infection. The

results of these analyses have led to the identification of three putative reductases, as detailed below.

Homology Modeling. Homology modeling predicts the three-dimensional structure of a query protein via the sequence alignment of template proteins. For our purposes here, two homology modeling servers were used to build the structure of our reductase candidates: SWISS-MODEL,²¹ which works only with the amino acid sequence of a target protein, and I-TASSER,²² which performs the threading by identifying template proteins from solved structure databases that have a similar structure or similar structural motifs as the target sequence. We used these predictors, because both of them performed well in the last Critical Assessment of Protein Structure (CASP) challenge and I-TASSER has been consistently ranked amongst the top methods in the community-wide CASP experiments for accurate 3D structure prediction.²³ The three-dimensional structures of three candidates were built using their sequence excluding the residues related to the TMH.

Protein-Protein interaction. Protein-protein interactions are important to understand the function and organization of proteins in the cell. For this purpose, ClusPro server²⁴ was used. This web-based server follows three steps: i) rigid-body docking by sampling of billions of conformations using the PIPER program, which is based on the Fast Fourier Transform (FFT) correlation approach, ii) a root mean square deviation-based clustering of the 1000 structures with the lowest energy, iii) and refinement of the

selected structures by energy minimization. The docking was performed by introducing trHbN as the receptor and the reductase partner as the ligand. The best protein complexes were further refined with MD simulations.

Molecular Dynamics (MD) Simulations.

Simulations were performed using the AMBER16 package²⁵ to refine the reductase-trHbN complexes. The Amber ff99SB-ildn²⁶ force field was used for the protein, whereas cofactors (heme, FAD and NADH) were parameterized using the GAFF force field in conjunction with restrained electrostatic potential-fitted (RESP) partial atomic charges derived from B3LYP/6-31G(d)²⁷ calculations. The simulated systems were immersed in an octahedral box of TIP3P²⁸ water molecules. Na⁺ atoms were added to maintain the neutrality of the system, which comprised the trHbN-reductase complex, cofactors, around 24000 water molecules and Na⁺, leading to systems containing ~80000 atoms.

Simulations were performed in the NPT ensemble for equilibration and NVT for MD productions. The minimization of the systems was performed refining the position of hydrogen atoms (2000 cycles of steepest descent algorithm followed by 8000 cycles of conjugate gradient), and subsequently the whole system (4000 cycles for steepest descent and 1000 cycles of conjugate gradient). Later, the temperature of the system was gradually increased from 100 to 300 K in 5 steps, 50 ps each using the NVT ensemble. In order to equilibrate the density of the system, an additional 5 ns step was performed in the NPT ensemble. First, production MD simulations were

run for 50 ns imposing restraints to the FAD cofactor with a force constant of 10 kcal mol⁻¹ Å⁻¹ to avoid artefactual displacements (Figure 2). Then, three replicas were created and the simulations extended for 500 ns with restraints (10 kcal mol⁻¹ Å⁻¹). On the other hand, another three replicas were simulated with restraints, but they were gradually reduced every 10ns (restraints reduced by 1 kcal mol⁻¹ Å⁻¹ at each step) and finally 500ns were launched without any restraints for each replica. After finishing these calculations, the NADH cofactor was inserted in the most stable simulated system and we extended the simulation of the full system for 500 ns.

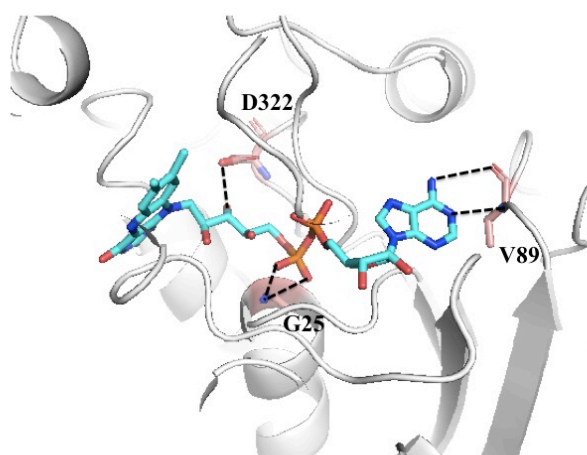


Figure 2. Representation of the FAD and the restraints applied to the position of FAD. The FAD and the important residues of reductase partner are indicated in cyan and pink sticks, respectively. Hydrogen atoms have been removed for the sake of clarity.

Electron transfer. To identify the most probable ET pathways between the heme group and the FAD isoalloxazine ring, the semiempirical pathway method developed by Beratan et al.^{29,30} was used. This method searches for the path that shows the highest donor to acceptor electron

coupling (H_{AB}). For a given path, H_{AB} is computed as the product of a number of steps, each with a given coupling value that defines the pathway. There are two types of steps for ET including transfer through atoms, when two orbitals share an atom, or through space, when orbitals are separated by empty space. For the first type, the coupling value is assumed to be 0.6 and for the other one, it depends on the orbital-to-orbital distances, which degrade exponentially by using a decay factor of -1.7. The decay factor for the best ET path is multiplied by 1 eV (which is the default value in the algorithm) to compute H_{AB} . Minor adjustments of parameters used in the pathway algorithm are needed to achieve a proper description of the system. These adjustments follow two considerations: first, all orbitals of the heme iron were considered to be equivalent and second, to reproduce the resonant character of all porphyrin and FAD isoalloxazine ring located orbitals, coupling between them were set to one. ET calculations were done for snapshots taken at 10 ps intervals during the last 200 ns of MD simulations.

Results and discussion

Finding the reductase candidates. The search in Tuberculist database²⁰ was performed to retrieve reductases and ET proteins containing at least a single transmembrane helix (TMH), which would enabled the insertion in the internal membrane. Three putative reductases were identified: two type II NAD dehydrogenases, denoted as *ndh* and *ndhA*, and a ferredoxin reductase, *fdxB*. The number of predicted TMH in these three proteins is shown in Figure 3; *ndh*

and *ndhA* have 1 TMH needed for insertion into the membrane while *fdxB* has 4 TMH.

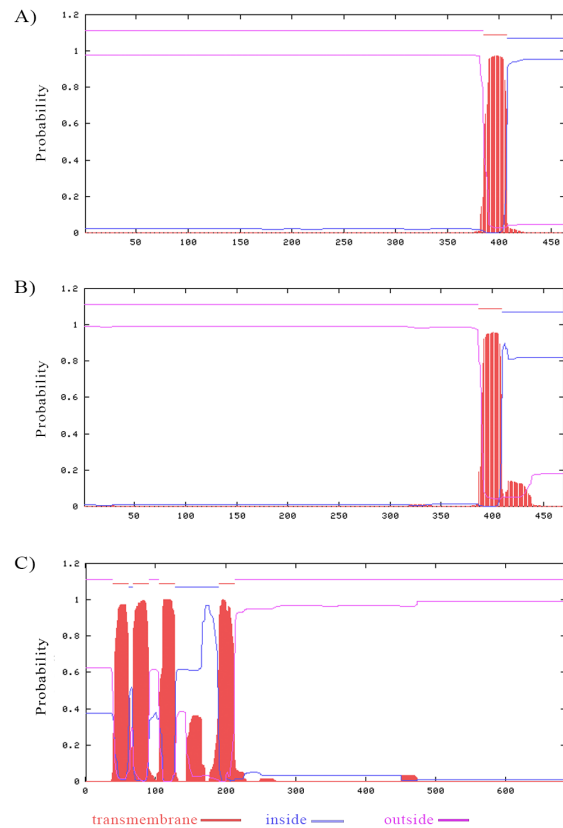


Figure 3. Predicted TMH regions for three candidates, A) *ndh*, B) *ndhA* and C) *fdxB*.²⁰

On the other hand, in a study performed by D. Schnappinger *et al.* in order to identify the genes that are differentially expressed by intraphagosomal *Mtb* in naïve and interferon- γ macrophages, it was demonstrated that the expression level of *ndh* and *ndhA* genes are induced (color yellow, Figure 4) with increased in the expression level of activated *trHbN* gene in γ -stimulated macrophages (color green, Figure 4) after 24h³¹. They concluded that the induction of *ndh* gene could signal an increased need for NAD^+ regeneration, the second major function of the respiratory chain besides energy production. So, considering all this information, we decided

to build the three-dimensional structures of these candidates in order to find the best candidates for electron transfer.

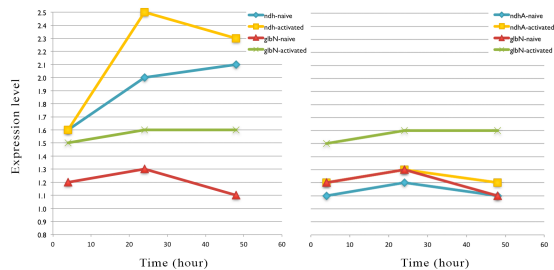


Figure 4. Comparison of gene regulation (ndh and ndhA) in IFN-gamma-stimulated macrophages with gene regulation in naive macrophages.³¹

3D structures of the selected candidates. The 3D structural models were built using both I-TASSER²² and SWISS-MODEL²¹ in order to check for consistency of the predicted structures. The selected templates used by SWISS-MODEL and I-TASSER included 4NWZ (ndh), 4XDB (ndhA) and 2PIA (fdxB).

The results in Figure 5 and the calculated RMSD between these two structures (around 2.0Å for the backbone atoms) indicate that there are no significant structural differences between the results obtained from the homology modelling algorithms. However, since the I-Tasser results are based on threading by scanning protein motif databases and was ranked as the best server for full-length 3D protein structural predictors, we

selected these models to conduct the protein-protein docking.

Protein-protein docking. Docking was used to identify the interaction between trHbN and the three reductase candidates, which should lead to models keeping heme and FAD cofactors at a distance adequate to accomplish the ET. Between the various protein-protein poses, the most probable complexes were chosen on the basis of the largest number of the members in the cluster of docked solutions, the shortest distance from the geometrical centers of heme and FAD, and the lowest energy as specified by using the balanced ClusPro algorithm. The results show that the interaction of the two proteins include the CD loop of trHbN (Figure 6), which locates close to the C-terminus of reductase partner. The interaction is assisted by two π -stacking interactions between His of trHbN and Phe residues of ndh. Also, there is a non-covalent interaction formed by Phe and Pro residues. Moreover, two salt bridges can be found through two Asp of trHbN and Lys residues of ndh in the interface of two proteins. The docked poses are superposed to the crystal structure of trHbN (1IDR) and the template (which were used to build the 3D structures of the candidates by homology modeling) with respect to the placement of two cofactor-binding regions.

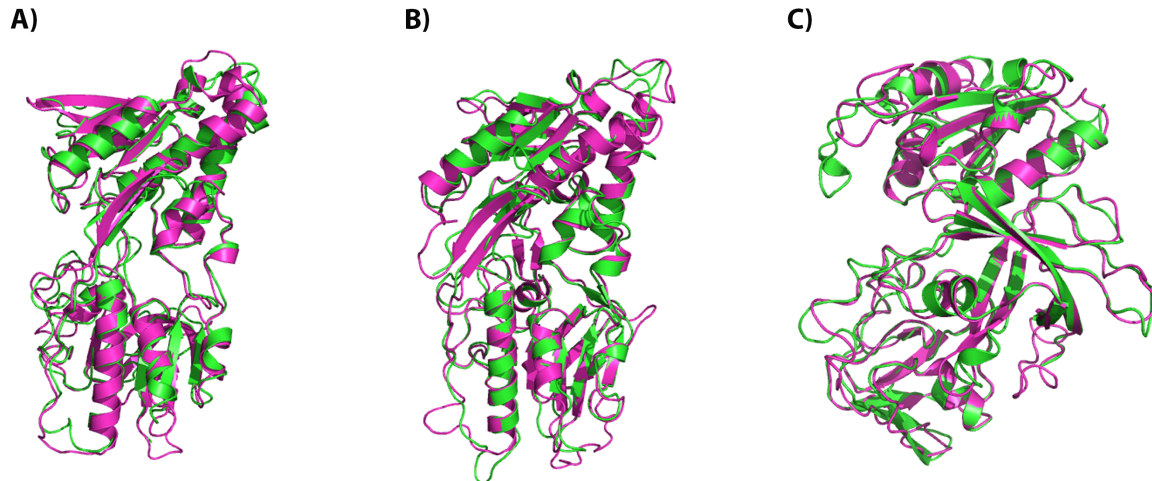


Figure 5. The three-dimensional structures of A) *ndh*, B) *ndhA* and C) *fdxB* reductases built by SWISS-MODEL (green) and I-TASSER (magenta).

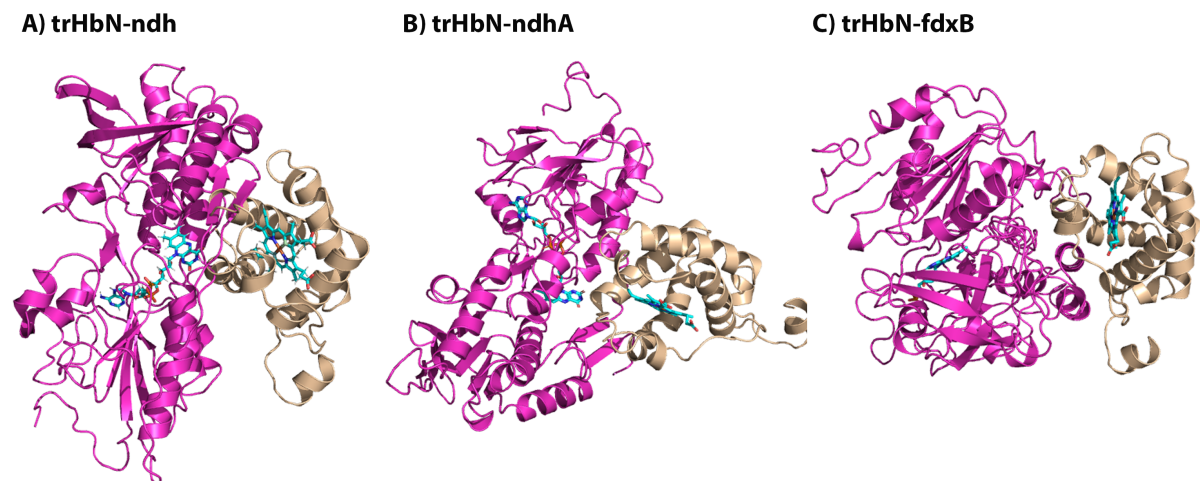


Figure 6. The representation of the dimeric complex trHbN (wheat) – reductase (magenta) for A) *ndh*, B) *ndhA* and C) *fdxB* obtained after protein-protein docking. The FAD cofactor and heme are indicated with sticks.

In the case of *ndh* and *ndhA*, the distance between the heme and FAD cofactor are around 11 Å, but it is enlarged to 25 Å for *fdxB* (Figure 6). This large distance might be affected by the predicted 3D structure, especially regarding the arrangement of loops in the interface between trHbN and reductase, suggesting that the 3D model obtained for the trHbN-*fdxB* complex is the less favorable. But, to relieve this distance, we divided the sequence of *fdxB* into two parts and we built one 3D structure with I-TASSER

for each part and then, we docked these two structures by the use of Cluspro webserver to make the complete protein. After accomplishing this step, we again performed the protein-protein interaction between trHbN and *fdxB* and the distance between cofactors was reduced to 15Å. *Ndh* and *ndhA* share a high degree of amino acid similarity (67% sequence homology). Gene knockout of *ndh* and *ndhA* genes in wild type and merodiploid *Mtb* showed that *ndh* gene is essential in a wild type strain, though it can be

inactivated when an additional copy of *ndh* was provided.³² In contrast, the *ndhA* gene could be deleted in *Mtb* without causing adverse effects *in vitro*. Clearly, these results point out that *ndh* and *ndhA*, despite the biochemical similarity, play different roles in *Mtb*, and confirm the essential character of *ndh*. Additionally, in line with these studies, a double (*ndh+ndhA*) knockout could not be obtained, suggesting that at least one type II NADH dehydrogenase is required for *Mtb* growth.³³ Further, it was concluded that *ndh* is the main NADH dehydrogenase in the ET chain, confirming its essentiality for *Mtb*.³⁴ Finally, it has been demonstrated by the recent identification of two compounds that *ndh* is druggable,³⁵ which could solve the druggability problem of trHbN. Hence, the design of *ndh* inhibitors could benefit from blockage of ATP synthesis and weakening of NO scavenging. Furthermore, *ndh* is absent from the mammalian genome, which reinforces its relevance as a potential target. So, based on all this information, we selected the *ndh* reductase candidate to refine the trHbN-*ndh* complex.

Molecular Dynamics (MD) Simulations. The trHbN-*ndh* complex was simulated using six distinct replicas containing FAD and heme

cofactors. We performed three MD simulations (500 ns each) applying restraints to keep the arrangement of the FAD cofactor. In these simulations the RMSD results show that both proteins fluctuate in the first 200ns of the calculations, but remain stable until the end of the simulations, leading to an average RMSD of 2.8 ± 0.4 Å (Table 1). Three additional MD simulations were performed eliminating the restraints imposed to the FAD cofactor in 10 steps of 10 ns each, followed by a production simulation of 500 ns without restraints.

Table 1. RMSD and standard deviation determined for simulated systems with and without restraints.

Systems		Rep 1	Rep 2	Rep 3
With restraint	trHbN+ <i>ndh</i>	3.1 (0.4)	2.8 (0.5)	2.7 (0.3)
	FAD	1.8 (0.2)	1.6 (0.3)	1.5 (0.2)
Without restraint	trHbN+ <i>ndh</i>	3.0 (0.2)	3.8 (0.5)	3.0 (0.3)
	FAD	1.9 (0.8)	2.3 (0.3)	1.8 (0.2)

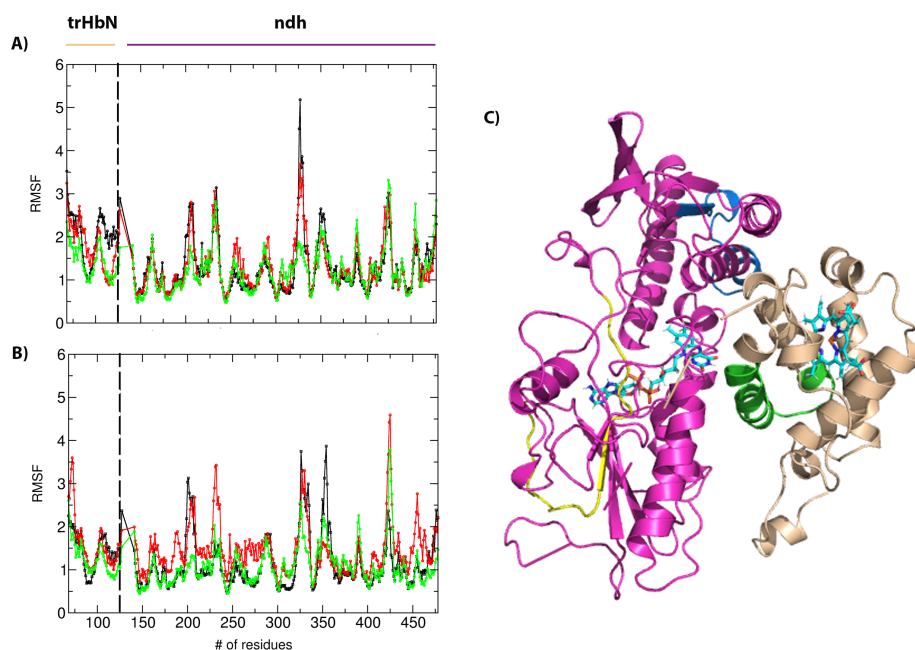


Figure 7. RMSF (Å) for the three replicas A) with and B) without restraints. The representation of the dimeric complex trHbN (wheat) – reductase (Magenta) indicating the regions with the highest fluctuations in trHbN (green) and ndh (yellow and blue). Residues 1-17 in trHbN were excluded from the analysis.

Two of the three replicas are stabilized after the first 50 ns, while the third replica is stabilized after 300 ns. Moreover, the RMSD of the FAD cofactor show the stable behaviour in all the replicas during the MD simulations (note that the highly flexible pre-A and loop-F regions of trHbN were excluded from the RMSD analysis).

The root-mean square fluctuation (RMSF) is similar for the simulations with and without restraints. For the trHbN the region with the highest flexibility corresponds to residues 100-121 (green in Figure 7C), which corresponds to the interface of the trHbN with the ndh reductase. Comparatively, the structural fluctuations in ndh are larger than the values determined for trHbN. Figure 7C highlights in blue and yellow the residues related to the loops that shape the FAD binding pocket.

The final snapshot of one of the replicas without restraints was selected to add the NADH cofactor

based on the X-ray structure of *S. cerevisiae* ndh (PDB code 4G6H, Figure 8).

The RMSD value of the trHbN-ndh complex in presence of all cofactors is larger than the previous RMSD calculated when also heme and FAD cofactors were present in the trHbN and ndh, respectively (Figure 9). This increase is basically due to the structural accommodation of the NADH cofactor by ndh. The RMSD value for trHbN is very stable along the whole simulation. The RMSF pattern is comparable with the previous calculations and the largest fluctuations correspond to the reductase residues (Figure 10).

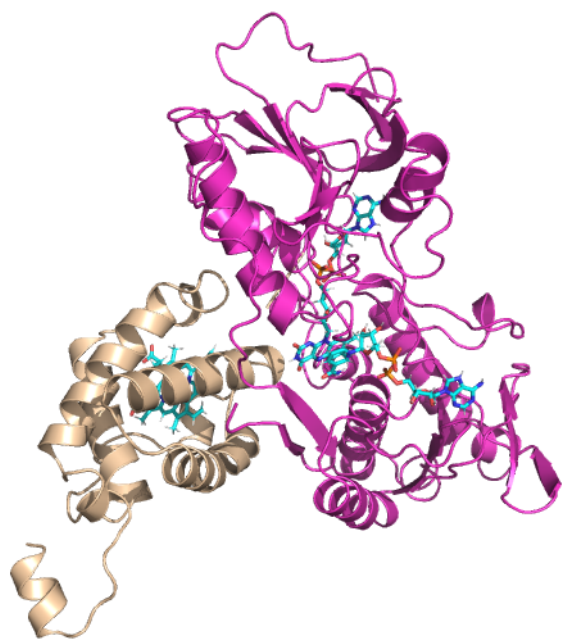


Figure 8. The representation of the dimeric complex trHbN (wheat) – reductase (magenta). The cofactors heme, FAD and NADH are shown as cyan sticks.

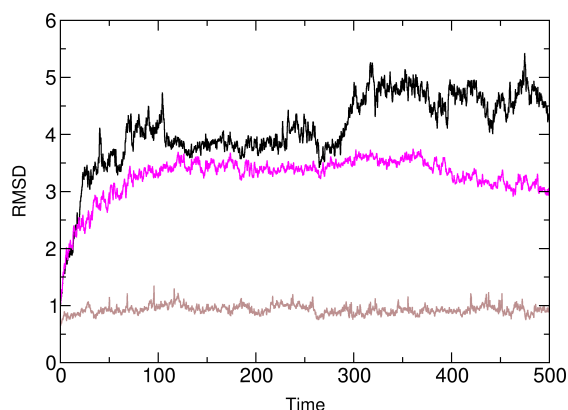


Figure 9. RMSD plot of the trHbN:ndh dimeric complex (black) in presence of all cofactors. The separate RMSD of trHbN and ndh reductase are indicated in brown and magenta, respectively.

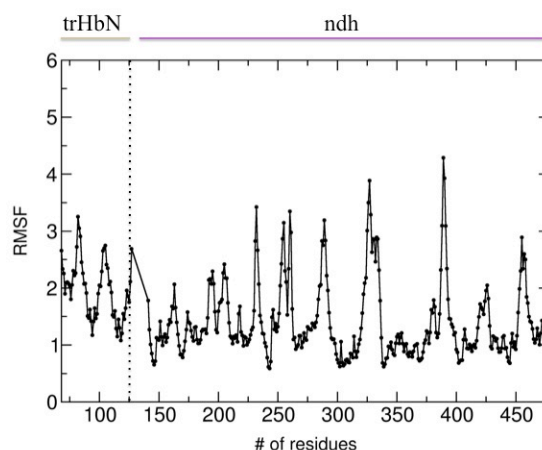


Figure 10. RMSF Plot of the dimeric complex. The most flexible regions correspond to the reductase structure.

Electron transfer. To check the efficiency of the ET for the refined structure of the trHbN-ndh complex, the ET rate determined from the 4000 frames along the whole trajectory. The results indicate an electron transfer rate of 3.7×10^{-5} eV, 2.5×10^{-5} and 4×10^{-5} for replicas 1, 2 and 3, respectively, which are comparable with the electron coupling reported for *E.coli* flavoHb (7.1×10^{-5} eV), a system that is known to cycle electrons very efficiently.²⁹

There are two main ET pathways that link the isoalloxazine ring of FAD in ndh to the heme cofactor of trHbN. One of the pathways involves the residues Met87 and Phe90 of the trHbN (Figure 11). The second path involves the direct contribution of the water molecules of the interface and the Phe90 of the trHbN. So, Met87, Phe90 and water molecules have a major contribution to the ET pathways. Water molecules play an important role in mediating the transition between FAD and Phe90, however in two replicas where Met87 participates in the path, there are usually no water molecules involved.

We can conclude that Met87 play an important role in the ET, so blocking the path of this residue can significantly affect the potential of the reductase partner.

Binding site analysis. A potential binding pocket has been identified in the interface between both proteins, near to the position of FAD and NADH. Key residues at this interface and directly related to the electron transfer path together with potential binding pocket could be used for the drug design strategy (Figure 12). Moreover, based on the studies by M. B. Harbut *et al.*³⁴ two molecules named CBR-4032 (tetrahydroindazole series) and CBR-1922 (thioquinazoline scaffold) can occupy the quinone and NADH binding sites, respectively, given their structural resemblance to the quinone and adenine molecules that normally reside in those positions. In our docking analysis, we could confirm that the CBR-4032 occupied the NADH position in

our 3D structure of ndh.

Conclusion

TrHbN has a significant role on the pathophysiology of MtB because of its high oxygen affinity and NO detoxification ability. However, to ensure this efficient ability a reductase partner, which has not been identified yet, is needed to recover the trHbN-Fe(III) to trHbN-Fe(II). In this manuscript we show a compatible redox partner to maintain the trHbN's function, presenting the ndh as the potential reductase partner for trHbN. We have modeled the 3D structure of the reductase for ndh sequence of Mtb, and the best candidates of the protein-protein docking were refined by MD simulations. Finally, the electron transfer efficiency was assessed by the trHbN-ndh complex, confirming its ability for an efficient reaction.

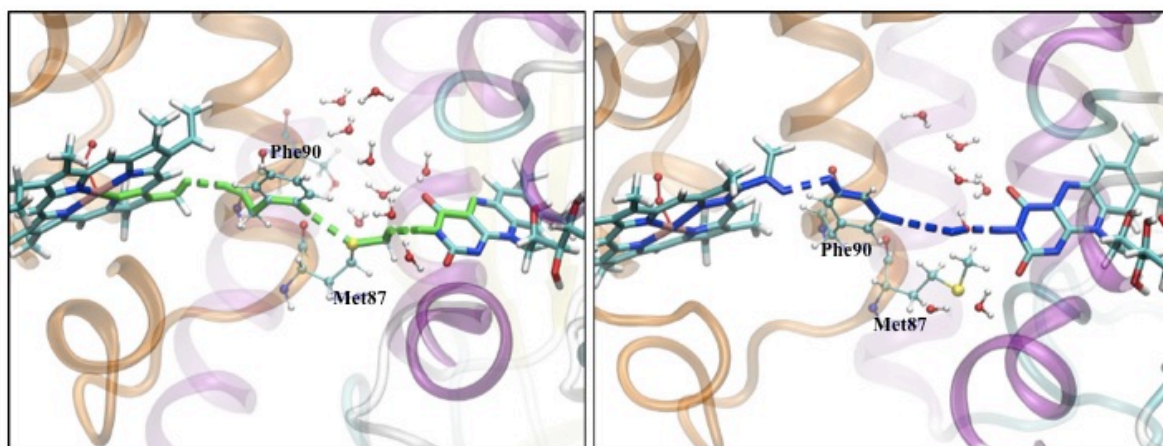


Figure 11. Representation of the major ET pathways. Heme, FAD, Met87 and Phe90 are shown as cyan sticks.

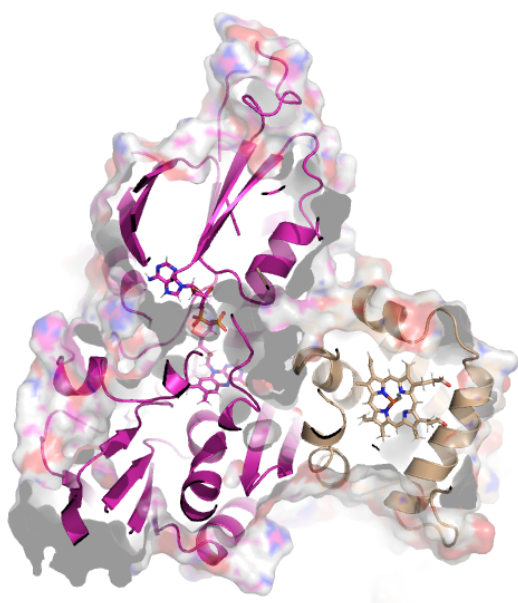


Figure 12. Representation of the binding pocket identified at the interface between both proteins.

AUTHOR INFORMATION

Corresponding Author

*E-mail: fjluque@ub.edu

ORCID

F. Javier Luque: 0000-0002-8049-3567

Carolina Estarellas: 0000-0002-0944-9053

Elnaz Aledavood: 0000-0002-4837-7849

Notes

The authors declare no competing financial interest

Acknowledgments

We thank the Spanish Ministerio de Economía y Competitividad (SAF2017-88107-R, and Maria de Maetzu MDM-2017-0767), and the Generalitat de Catalunya (2017SGR1746), for financial support and the Barcelona Supercomputing Center (BSC-CNS: BCV-2014-

3-0011 and BCV-2019-1-0009) and the Consorci de Serveis Universitaris de Catalunya (CSUC) for providing access to computational resources. E.A. thanks AGAUR (Generalitat of Catalunya; 2018FI_B1_00001) for a fellowship.

References

- (1) Russell, D. G. Mycobacterium Tuberculosis: Here Today, and Here Tomorrow. *Nat. Rev. Mol. Cell Biol.* **2001**, 2 (8), 569–577. <https://doi.org/10.1038/35085034>.
- (2) World Health Organization. The top 10 causes of death <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death> (accessed Mar 17, 2020).
- (3) Schön, T.; Miotto, P.; Köser, C. U.; Viveiros, M.; Böttger, E.; Cambau, E. Mycobacterium Tuberculosis Drug-Resistance Testing: Challenges, Recent Developments and Perspectives. *Clin. Microbiol. Infect.* **2017**, 23 (3), 154–160. <https://doi.org/10.1016/j.cmi.2016.10.022>.
- (4) Singh, V.; Mizrahi, V. Identification and Validation of Novel Drug Targets in Mycobacterium Tuberculosis. *Drug Discov. Today* **2017**, 22 (3), 503–509. <https://doi.org/10.1016/j.drudis.2016.09.010>.
- (5) Oliveira, A.; Singh, S.; Bidon-Chanal, A.; Forti, F.; Martí, M. A.; Boechi, L.; Estrin, D. A.; Dikshit, K. L.; Luque, F. J. Role of PheE15 Gate in Ligand Entry and Nitric Oxide Detoxification Function of Mycobacterium Tuberculosis Truncated Hemoglobin N. *PLoS One* **2012**, 7 (11). <https://doi.org/10.1371/journal.pone.0049291>.
- (6) Bidon-Chanal, A.; Martí, M. A.; Crespo, A.; Milani, M.; Orozco, M.; Bolognesi, M.; Luque, F. J.; Estrin, D. A. Ligand-Induced

- Dynamical Regulation of NO Conversion in Mycobacterium Tuberculosis Truncated Hemoglobin-N. *Proteins Struct. Funct. Genet.* **2006**, *64* (2), 457–464.
<https://doi.org/10.1002/prot.21004>.
- (7) Crespo, A.; Martí, M. A.; Kalko, S. G.; Morreale, A.; Orozco, M.; Gelpi, J. L.; Luque, F. J. Theoretical Study of the Truncated Hemoglobin HbN : Exploring the Molecular Basis of the NO Detoxification Mechanism. **2005**, No. 7, 4433–4444.
<https://doi.org/10.1021/ja0450004>.
- (8) Pawaria, S.; Lama, A.; Raje, M.; Dikshit, K. L. Responses of Mycobacterium Tuberculosis Hemoglobin Promoters to In Vitro and In Vivo Growth Conditions □. **2008**, *74* (11), 3512–3522.
<https://doi.org/10.1128/AEM.02663-07>.
- (9) Heroux, M. S.; Mohan, A. D.; Olsen, K. W. Ligand Migration in the Truncated Hemoglobin of Mycobacterium Tuberculosis. **2011**, *63* (March), 214–220.
<https://doi.org/10.1002/iub.438>.
- (10) Pesce, A.; Couture, M.; Dewilde, S.; Guertin, M.; Yamauchi, K.; Ascenzi, P.; Moens, L.; Bolognesi, M. A Novel Two-over-Two a - Helical Sandwich Fold Is Characteristic of the Truncated Hemoglobin Family. **2000**, *19* (11).
- (11) Milani, M.; Pesce, A.; Ouellet, Y.; Ascenzi, P.; Guertin, M.; Bolognesi, M. Mycobacterium Tuberculosis Hemoglobin N Displays a Protein Tunnel Suited for O₂ Diffusion to the Heme. *EMBO J.* **2001**, *20* (15), 3902–3909.
<https://doi.org/10.1093/emboj/20.15.3902>.
- (12) Perutz, M. F. AFFINITY OF HEMOGLOBIN : Influence of Structure of the Globin on the Heme Iron. **1979**.
- (13) Milani, M.; Pesce, A.; Ouellet, Y.; Ascenzi, P.; Guertin, M.; Bolognesi, M. Mycobacterium Tuberculosis Hemoglobin N Displays a Protein Tunnel Suited for O₂ Diffusion to the Heme. **2001**, *20* (15).
- (14) Lama, A.; Pawaria, S.; Bidon-chanal, A.; Anand, A.; Gelpi, L.; Arya, S.; Martí, M.; Estrin, D. A.; Luque, F. J.; Dikshit, K. L. Role of Pre-A Motif in Nitric Oxide Scavenging by Truncated Hemoglobin , HbN , of Mycobacterium Tuberculosis * □. **2009**, *284* (21), 14457–14468.
<https://doi.org/10.1074/jbc.M807436200>.
- (15) Milani, M.; Pesce, A.; Nardini, M.; Ouellet, H.; Ouellet, Y.; Dewilde, S.; Bocedi, A.; Ascenzi, P.; Guertin, M.; Moens, L.; et al. Structural Bases for Heme Binding and Diatomic Ligand Recognition in Truncated Hemoglobins. **2005**, *99*, 97–109.
<https://doi.org/10.1016/j.jinorgbio.2004.10.035>.
- (16) Ascenzi, P.; Bolognesi, M.; Milani, M.; Guertin, M.; Visca, P. Mycobacterial Truncated Hemoglobins : From Genes to Functions. **2007**, *398*, 42–51.
<https://doi.org/10.1016/j.gene.2007.02.043>.
- (17) Gardner, P. R.; Gardner, A. M.; Martin, L. A.; Dou, Y.; Li, T.; Olson, J. S.; Zhu, H.; Riggs, A. F. Nitric-Oxide Dioxygenase Activity and Function of Flavohemoglobins. Sensitivity to Nitric Oxide and Carbon Monoxide Inhibition. *J. Biol. Chem.* **2000**, *275* (41), 31581–31587.
<https://doi.org/10.1074/jbc.M004141200>.
- (18) Eich, R. F.; Li, T.; Lemon, D. D.; Doherty, D. H.; Curry, S. R.; Aitken, J. F.; Mathews, A. J.; Johnson, K. A.; Smith, R. D.; Phillips, G. N.; et al. Mechanism of NO-Induced Oxidation of Myoglobin and Hemoglobin. *Biochemistry* **1996**, *35* (22), 6976–6983.
<https://doi.org/10.1021/bi960442g>.
- (19) Pathania, R.; Navani, N. K.; Rajamohan, G.;

- Dikshit, K. L. Mycobacterium Tuberculosis Hemoglobin Hbo Associates with Membranes and Stimulates Cellular Respiration of Recombinant Escherichia Coli. *J. Biol. Chem.* **2002**, *277* (18), 15293–15302. <https://doi.org/10.1074/jbc.M111478200>.
- (20) No Title <https://mycobrowser.epfl.ch> (accessed Nov 20, 2020).
- (21) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. SWISS-MODEL: An Automated Protein Homology-Modeling Server. *Nucleic Acids Res.* **2003**, *31* (13), 3381–3385. <https://doi.org/10.1093/nar/gkg520>.
- (22) Yang, J.; Zhang, Y.; Rohini, K.; Srikumar, P. S.; Anbarasu, K.; Structural, L.; Suite, B.; Ec, A.; Security, D. C.; Roy, A.; et al. Function Prediction. *Curr. Protoc. Bioinforma.* **2011**, *9* (December), 5–9. <https://doi.org/10.1038/nprot.2010.5.I-TASSER>.
- (23) Moul, J.; Fidelis, K.; Kryshchak, A.; Schwede, T.; Tramontano, A. Critical Assessment of Methods of Protein Structure Prediction (CASP)—Round XII. *Proteins Struct. Funct. Bioinforma.* **2018**, *86* (August 2017), 7–15. <https://doi.org/10.1002/prot.25415>.
- (24) Kozakov, D.; Hall, D. R.; Xia, B.; Porter, K. A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S. The ClusPro Web Server for Protein-Protein Docking. *Nat. Protoc.* **2017**, *12* (2), 255–278. <https://doi.org/10.1038/nprot.2016.169>.
- (25) Case, D. A.; Betz, R. .; Botello-Smith, W.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; TGiese, T. J.; Gohlke, H.; Goetz, A. W.; et al. Amber 2016. *Univ. California, San Fr.* **2016**.
- (26) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber Ff99SB Protein Force Field. *Proteins Struct. Funct. Bioinforma.* **2010**, *78* (8), 1950–1958. <https://doi.org/10.1002/prot.22711>.
- (27) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280. <https://doi.org/10.1021/j100142a004>.
- (28) 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.* **1983**, *79* (2), 926–935. <https://doi.org/10.1063/1.445869>.
- (29) Beratan, D.; Onuchic, J.; Winkler, J.; Gray, H. Electron-Tunneling Pathways in Proteins. *Science (80-.)*. **1992**, *258* (5089), 1740–1741.
- (30) Beratan, D.; Betts, J.; Onuchic, J. Protein Electron Transfer Rates Set by the Bridging Secondary and Tertiary Structure. *Science (80-.)*. **1991**, *252*, 1285–1288.
- (31) Schnappinger, D.; Ehrt, S.; Voskuil, M. I.; Liu, Y.; Mangan, J. A.; Monahan, I. M.; Dolganov, G.; Efron, B.; Butcher, P. D.; Nathan, C.; et al. Transcriptional Adaptation of Mycobacterium Tuberculosis within Macrophages: Insights into the Phagosomal Environment. *J. Exp. Med.* **2003**, *198* (5), 693–704. <https://doi.org/10.1084/jem.20030846>.
- (32) Pathania, R.; Navani, N. K.; Gardner, A. M.; Gardner, P. R.; Dikshit, K. L. Nitric Oxide Scavenging and Detoxification by the Mycobacterium Tuberculosis Haemoglobin, HbN in Escherichia Coli. *Mol. Microbiol.* **2002**, *45* (5), 1303–1314. <https://doi.org/10.1046/j.1365->

- 2958.2002.03095.x.
- (33) Sethi, D.; Mahajan, S.; Singh, C.; Lama, A.; Hade, M. D.; Gupta, P.; Dikshit, K. L. Lipoprotein LprI of Mycobacterium Tuberculosis Acts as a Lysozyme Inhibitor. *J. Biol. Chem.* **2016**, *291* (6), 2938–2953. <https://doi.org/10.1074/jbc.M115.662593>.
- (34) Harbut, M. B.; Yang, B.; Liu, R.; Yano, T.; Vilchèze, C.; Cheng, B.; Lockner, J.; Guo, H.; Yu, C.; Franzblau, S. G.; et al. Small Molecules Targeting Mycobacterium Tuberculosis Type II NADH Dehydrogenase Exhibit Antimycobacterial Activity. *Angew. Chemie - Int. Ed.* **2018**, *57* (13), 3478–3482. <https://doi.org/10.1002/anie.201800260>.
- (35) Crespo, A.; Martí, M. A.; Kalko, S. G.; Morreale, A.; Orozco, M.; Gelpi, J. L.; Luque, F. J.; Estrin, D. A. Theoretical Study of the Truncated Hemoglobin HbN: Exploring the Molecular Basis of the NO Detoxification Mechanism. *J. Am. Chem. Soc.* **2005**, *127* (12), 4433–4444. <https://doi.org/10.1021/ja0450004>.

CHAPTER



***RESULTS
SUMMARY***

4. Results summary

4.1. Study of the direct activation mechanism of AMPK

Over the last decades since formally naming AMPK¹⁰⁵, interest in this enzyme as a target has continued to grow due to the fact that AMPK is activated by physiological regulators.^{106,107} AMPK has a crucial role in cellular energy homeostasis and it is regulated by different mechanisms including the indirect and allosteric activation,¹⁰⁸ and direct activation mechanism reported more recently (see section 1.2.2). In this thesis we have centered our efforts in the understanding of molecular factors that govern the direct activation mechanism of AMPK.

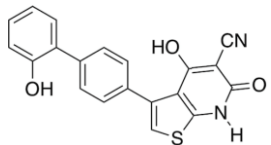
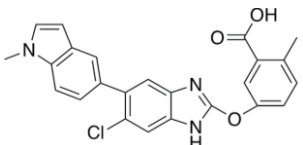
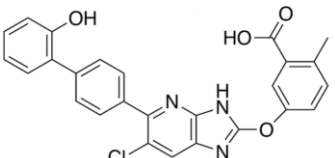
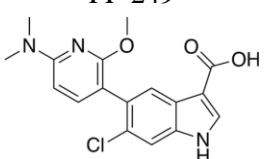
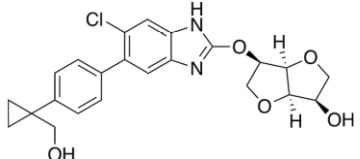
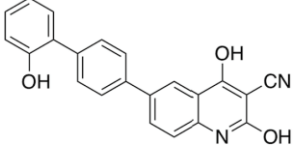
The first direct activator was A-769662, which binds to the ADaM site formed by α and β subunits, and it needs the phosphorylation of β Ser108 instead of α Thr172 for its activation. For several years, A-769662³⁶ remained the only direct activator for its type and after that, several direct activators were reported. In 2012, a paper reviewed the existing patent databases for AMPK activators reporting 26 patents that disclosed 10 classes of direct AMPK activators.¹⁰⁹ Despite the discovery of the direct activators of AMPK, their activation mechanism remains only partially understood. On the other hand, some of them are selective for specific AMPK complex isoforms, while others can activate both β 1 and β 2 containing isoforms (see section 1.2.2.1).

Regarding the different subunits of AMPK, the γ -subunit does not participate in the direct activation mechanism, so it was not considered in our studies. Furthermore, as it is shown in Table 1, exchange of α 1 and α 2 subunits does not have an important impact on the direct activation of this enzyme, whereas substitution of β 1 subunit is a key feature that mediates the selective activation of α 2 β 1 and α 2 β 2 isoforms, explaining the different selectivity of A-769662 and SC4. On the basis of this information, we focused on the role played by β 1 and β 2 subunits to understand the molecular mechanism that underlies the regulatory effect of direct activators and to gain insight into the puzzling tissue-dependent modulation of AMPK.

To this end, we explored the direct activation mechanism of A-769662 in α 2 β 1 isoform in the first paper, and the molecular basis of the selective isoform activation of AMPK specially targeting the role played by β 1 and β 2 subunits in the presence of A-769662 and SC4 was examined in the second paper. In the third paper, which was carried out in collaboration with Profs. A. Castro and M. S. Fernández, we studied a novel AMPK modulator named IND6, which shows affinity for isoform α 1 β 1 γ 1 and can act as a mixed-type inhibitor but also, can promote the enzymatic activation by adopting two distinct binding modes at the ADaM site.

Results summary

Table 1. Activation of AMPK isoforms by selected direct activators.

Activator	Property	Isoform			
		$\alpha 2\beta 1\gamma 1$	$\alpha 2\beta 2\gamma 1$	$\alpha 1\beta 1\gamma 1$	$\alpha 1\beta 2\gamma 1$
A-769662 	Activation fold ^a	14.3	NA ^b	2.0	NA ^b
	K_d (μM) ^a	0.40	---	0.51	14.5
991 	Activation fold ^a	12.2 7.7 ^{c,d}	5.4 5.7 ^c	4.8	2.4
	K_d (μM) ^a	0.06 ^e 0.085	---	0.06 0.078	0.51 1.18
SC4 	Activation fold ^f	5.2	2.9	2.4	1.2
	EC_{50} (nM) ^f	---	17.2	5.1	---
PF-249 	EC_{50} (nM) ^g	8.04	> 40 μM	9.37	> 40 μM
PF-739 	EC_{50} (nM) ^g	5.23	42.4	8.99	136
MT47-100 ¹¹⁰ 	Activation fold ^h	2.3	0.5	2.5	0.4

^a Ref 36. ^b No activation detected up to a concentration of 10 μM No ^c Ref. 35. ^d Determined for $\alpha 2\beta 1\gamma 2$. ^e Values obtained by using either biolayer interferometry or circular dichroism. ^f Ref. 39. ^g Ref. 40. ^h Ref. 133.

The conformational rearrangement triggered upon binding of the compounds was studied computationally by means of unbiased MD simulations (Table 2), and the generated ensembles were analysed by different techniques such as essential dynamics (ED), dynamical cross-correlation matrix (DCCM) analysis and interaction energy network. Our simulated systems consist of 368 residues in the case of $\alpha 2\beta 1$, 367 for $\alpha 2\beta 2$ and 370 residues for $\alpha 1\beta 1$, and

simulations were performed in triplicate for apo, holo (pAMPK+activator) and holo+ATP (pAMPK+activator+ATP).

Table 2. Total simulation time for each simulated system

System	Simulation time
apo	6 μ s
pAMPK+activator (A-769662, SC4 and PF-739)	18 μ s
pAMPK+activator+ATP	18 μ s
pAMPK+modulator (IND6)	1 μ s
pAMPK+modulator+ATP	1 μ s

4.1.1. Structural Analysis

Analysis of the trajectories helps us to explore the changes induced by the binding of activators and ATP on the transition from the basal state of AMPK to the activated form. The root mean square deviation (RMSD) points out that the presence of activator and ATP leads to an overall reduction in the enzyme flexibility in all the $\alpha 2\beta 1$ replicas. Nevertheless, this effect is less visible in the case of $\alpha 2\beta 2$, since the RMSD values of holo systems are similar or even higher than the values for the apo forms (Table 3). It is worth noting that in the case of PF-739, increase in the RMSD of the holo respect to the apo species does not only happen in $\alpha 2\beta 2$ isoform, even in some replicas of $\alpha 2\beta 1$ AMPK complexes, the RMSD of the simulated structures is not reduced from apo to the holo state.

Moreover, the per-residue mean square fluctuation (RMSF) results show a similar pattern for both apo forms. The regions with the highest flexibility in the α subunit correspond to the activation loop and the α helix between the residues 210 and 230 (Figure 10). Binding of the activator leads to subtle changes in the RMSF profile, including enhanced fluctuations in the activation loop and the α C-helix and reduction in the fluctuation of residues in the β subunit. Binding of ATP has no apparent effect on this profile. In general, binding of the activators in the holo states, increase the fluctuations of the main moieties in the α -subunit while reduce the fluctuations in the main regions of the β -subunit, independently of the β -isoform.

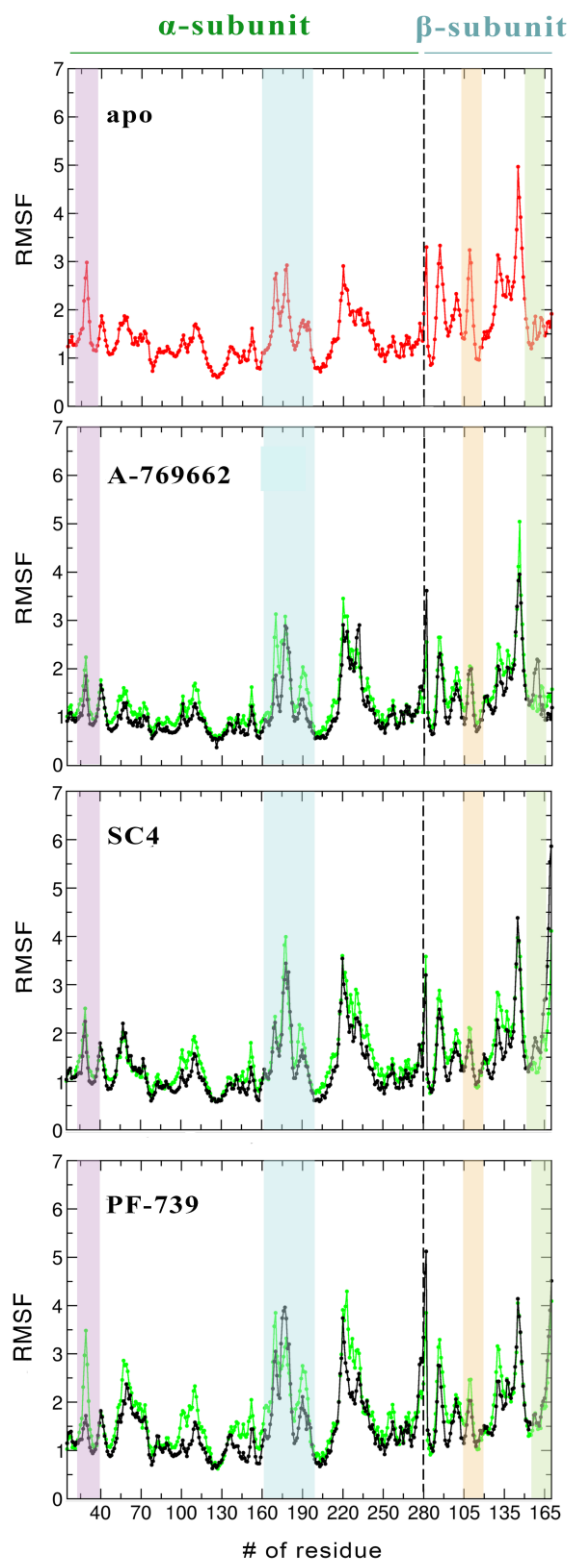


Figure 10. Averaged RMSF (\AA) of the residues along the last 500 ns of the three independent replicas run for apo (red), holo and holo+ATP (green and black, respectively). Complexes bound to A-769662, SC4 and PF-739 are shown in different panels. The highlighted bars denote regions corresponding to P-loop (purple), activation loop (cyan), CBM domain (orange), and C-interacting helix (green).

Table 3. RMSD and standard deviation (Å) determined for the protein backbone in apo, holo and holo+ATP states of AMPK isoforms $\alpha 2\beta 1$ and $\alpha 2\beta 2$. Values determined using the energy-minimized holo+ATP species averaged for the last 200 ns of each simulation system as reference structure.

Activator	System	Replica 1	Replica 2	Replica 3	Average
$\alpha 2\beta 1$					
A-769662	apo	2.5 ± 0.5	3.1 ± 0.6	2.6 ± 0.3	2.7
	holo	2.3 ± 0.4	1.9 ± 0.3	2.2 ± 0.3	2.1
	holo+ATP	1.9 ± 0.2	2.0 ± 0.3	1.9 ± 0.2	1.9
SC4	apo	3.2 ± 0.5	2.9 ± 0.5	3.1 ± 0.4	3.1
	holo	2.9 ± 0.4	2.2 ± 0.3	2.3 ± 0.6	2.5
	holo+ATP	2.0 ± 0.2	2.3 ± 0.2	2.2 ± 0.2	2.2
PF-739	apo	2.6 ± 0.6	2.7 ± 0.4	2.5 ± 0.5	2.6
	holo	2.5 ± 0.3	2.6 ± 0.5	2.9 ± 0.8	2.6
	holo+ATP	2.5 ± 0.3	2.0 ± 0.2	2.4 ± 0.4	2.3
$\alpha 2\beta 2$					
A-769662	apo	2.7 ± 0.3	3.3 ± 0.3	3.9 ± 0.4	3.3
	holo	4.2 ± 0.5	3.2 ± 0.4	4.5 ± 0.4	4.0
	holo+ATP	2.1 ± 0.2	2.2 ± 0.2	1.9 ± 0.1	2.1
SC4	apo	3.0 ± 0.3	2.8 ± 0.2	3.4 ± 0.4	3.1
	holo	3.3 ± 0.3	2.9 ± 0.3	3.1 ± 0.3	3.1
	holo+ATP	2.3 ± 0.2	2.5 ± 0.2	2.7 ± 0.2	2.5
PF-739	apo	3.2 ± 0.3	2.9 ± 0.2	3.4 ± 0.5	3.2
	holo	4.1 ± 0.6	2.7 ± 0.4	3.1 ± 0.3	3.3
	holo+ATP	2.7 ± 0.4	3.0 ± 0.5	3.0 ± 0.4	2.9

4.1.2. Dynamics of AMPK complexes

ED analysis was used to identify the differences in the conformational motions of the protein and disclose the changes in the protein dynamics. The main essential motions in the $\alpha 2\beta 1$ and $\alpha 2\beta 2$ apo forms are related to the core of α -helices in the α subunit and the β -sheets in β subunit. The motion of these two regions tends to bring the two domains closer and move them apart in a concerted motion. There are some differences in the extent of these motions in some elements, such as deformation of the CBM and the P-loop. In total, binding the activator and ATP reduces the conformational variance of the first essential motions of $\alpha 2\beta 1$ bound to the activators (**Figure 11**). In the holo states, binding of the activator leads to a synchronous motion between P-loop and the CBM domain. Subsequent binding of ATP reduces the overall flexibility of the protein regarding the P-loop, the α -helices and the β subunit proximal to the ADaM site, revealing a synergy between activator and ATP in increasing the stiffness of the protein.

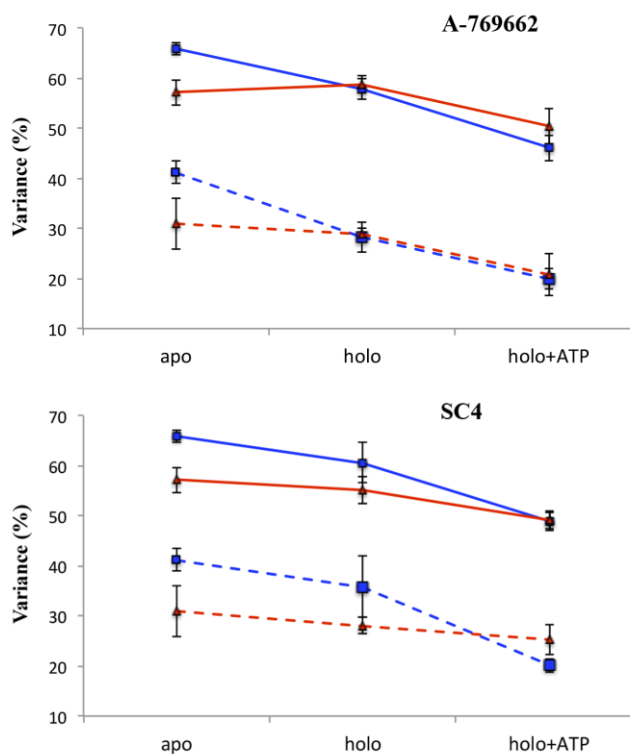


Figure 11. Contribution to the structural variance (%) of the first essential motion (dashed line) and the sum of the first three essential motions (solid line) of the protein backbone for apo, holo and holo+ATP species of AMPK. Values determined for $\alpha 2\beta 1$ and $\alpha 2\beta 2$ isoforms are shown in blue and red, respectively. Values averaged for the three replicas run for each system (bars denote the standard deviation).

In contrast to the $\alpha 2\beta 1$, this dynamical behavior is less regular in $\alpha 2\beta 2$. The structural variance of the holo form is similar or even higher than the apo state. Binding of activator can increase the flexibility of the P-loop, but the response in other structural elements is less regular.

DCCM analysis was performed to examine the correlation in the motions of residues along a trajectory, thus complementing the ED analysis. The results show a similar trend for A-769662, SC4 and PF-739 activators bound to AMPK. Binding of activator and ATP results in a decrease in the dynamical correlation of the residues, this effect being more noticeable in the case of $\alpha 2\beta 1$ (**Figure 12**). However, it reinforces the correlation between the motion of the CBM domain and the P-loop. Furthermore, the similarity indexes between DCCM maps are 0.91/0.84 for the different replicas of $\alpha 2\beta 1$ with A-769662/SC4, but it decreases to 0.75 for $\alpha 2\beta 2$ complexes.

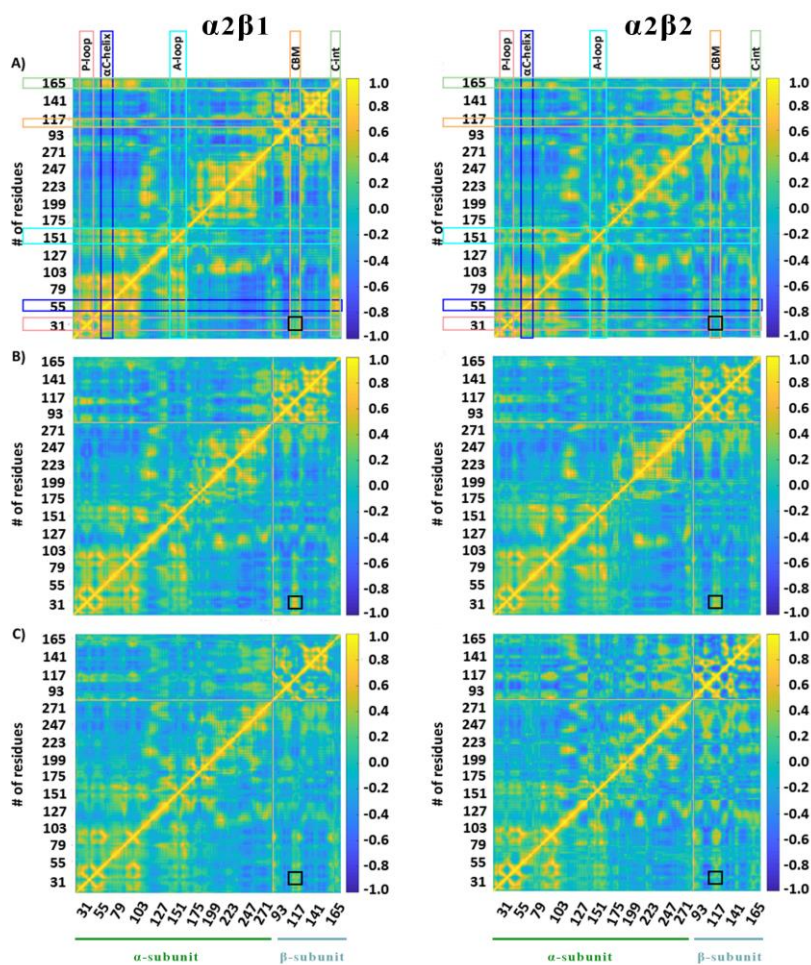


Figure 12. Dynamic cross-correlation (DCC) matrices for A) apo, B) holo and C) holo+ATP complexes of (left) $\alpha2\beta1$ and (right) $\alpha2\beta2$ with A-769662. Regions colored in yellow/blue show high correlated/anticorrelated fluctuations. The structural subunits and the most important regions of AMPK systems are indicated.

To see how binding of activator can affect ATP-binding site, we determined the structural resemblance of the ATP-binding site along the simulations using holo+ATP systems as reference. For holo+ATP species, the residues in the ATP-binding site sample a reduced conformational space for $\alpha2\beta1$ and $\alpha2\beta2$, while the apo species exhibits a wider distribution in the case of A-769662 and SC4. Unexpectedly, regarding the PF739, the apo shows a narrower distribution with one main peak centered at 2.0\AA for both $\alpha2\beta1$ and $\alpha2\beta2$ species. This fact is completely different to the results obtained for A-769662 and SC4, where the apo structure show larger distributions, or even a bimodal mode and the conformations sampled by the apo are very similar to those conformations already sampled by the holo+ATP species. The most distinctive features appear from the distribution profiles of the holo complexes (Figure 13). In $\alpha2\beta1$, binding of activator shifts the distribution toward the holo+ATP. In comparison, the RMSD profile of $\alpha2\beta2$ reveals that the activator leads to structures in which the ATP-binding site has less resemblance to holo+ATP systems.

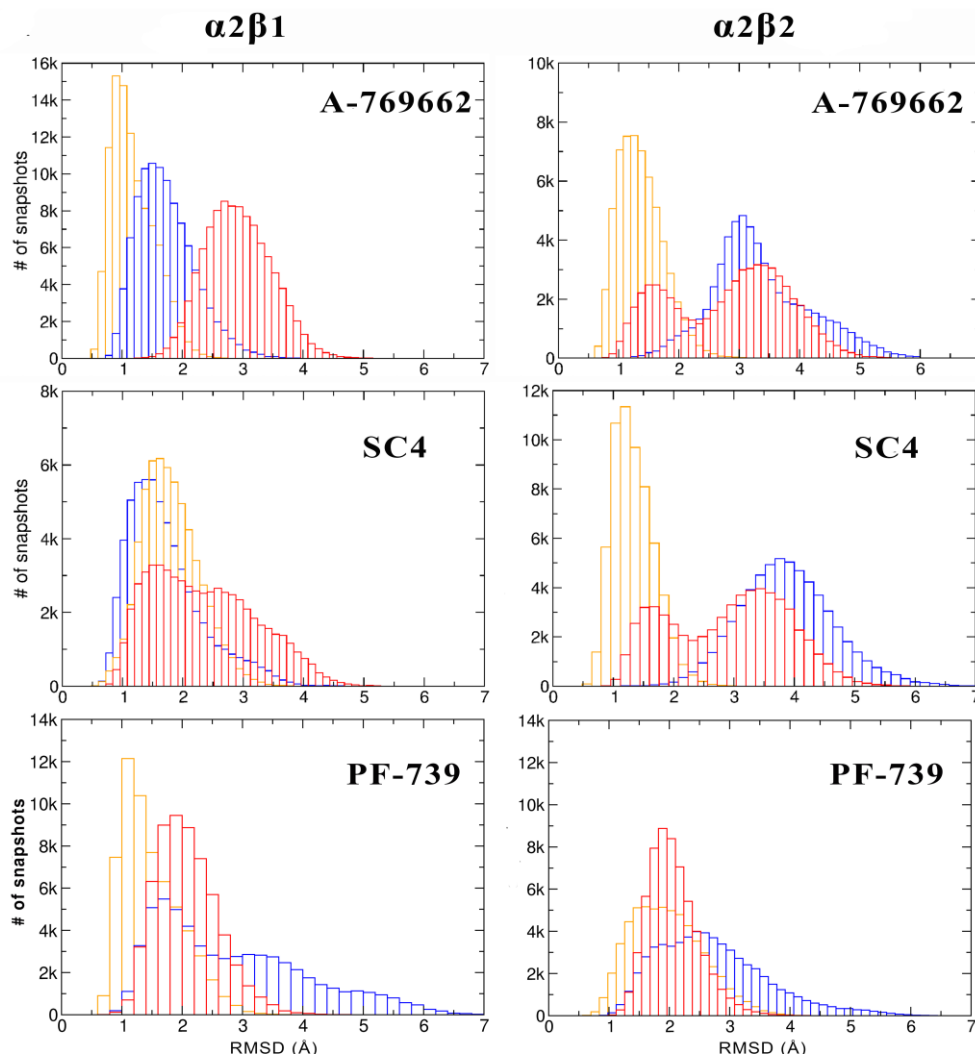


Figure 13. Distribution of the positional deviation (RMSD; Å) of the structures sampled along the trajectories run for apo (red), holo (blue), and holo+ATP (orange) for the residues that shape the ATP-binding site.

4.1.3. Comparing the network of interaction in $\beta 1$ and $\beta 2$ containing complexes

To investigate the selective activation of $\beta 1$ versus $\beta 2$ complexes, we examined the network of interactions between the activator, P-loop and CBM unit (ADaM site). The results obtained for $\alpha 2\beta 1$ with A-769662 shows that there is a salt bridge interaction between $\alpha\text{Asp}88$ and $\beta\text{Arg}83$, and at the outer edge of the ADaM site $\beta\text{Ser}108\text{p}$ adopts two arrangements and interacts with $\alpha\text{Lys}31$ and $\alpha\text{Lys}29$. Regarding the $\alpha 2\beta 2$, the interaction of $\beta\text{Arg}82$ and $\alpha\text{Asp}88$ is lost because of the mutation of $\beta 1\text{Asn}111 \rightarrow \beta 2\text{Asp}111$. In this case $\beta 2\text{Asp}111$ forms electrostatic interactions with $\beta 2\text{Arg}82$ instead of $\alpha\text{Asp}88$. Moreover, this residue forms a salt bridge interaction with $\beta\text{Ser}108\text{p}$. In contrast to the dual arrangement in $\alpha 2\beta 1$, $\beta\text{Ser}108\text{p}$ maintains a fixed orientation in $\alpha 2\beta 2$. These changes can be seen in the interaction pathways identified by Weighted Implementation of Suboptimal Paths (WISP; **Figure 14**). The CBM domain and p-

loop of $\alpha 2\beta 1$ are connected via three pathways. One of these paths includes the β Ser108p, while the other two paths involve the ligand via the triad β Arg83- α Asp88-A-769662. However, in the holo $\alpha 2\beta 2$ only the β Ser108p pathway is found.

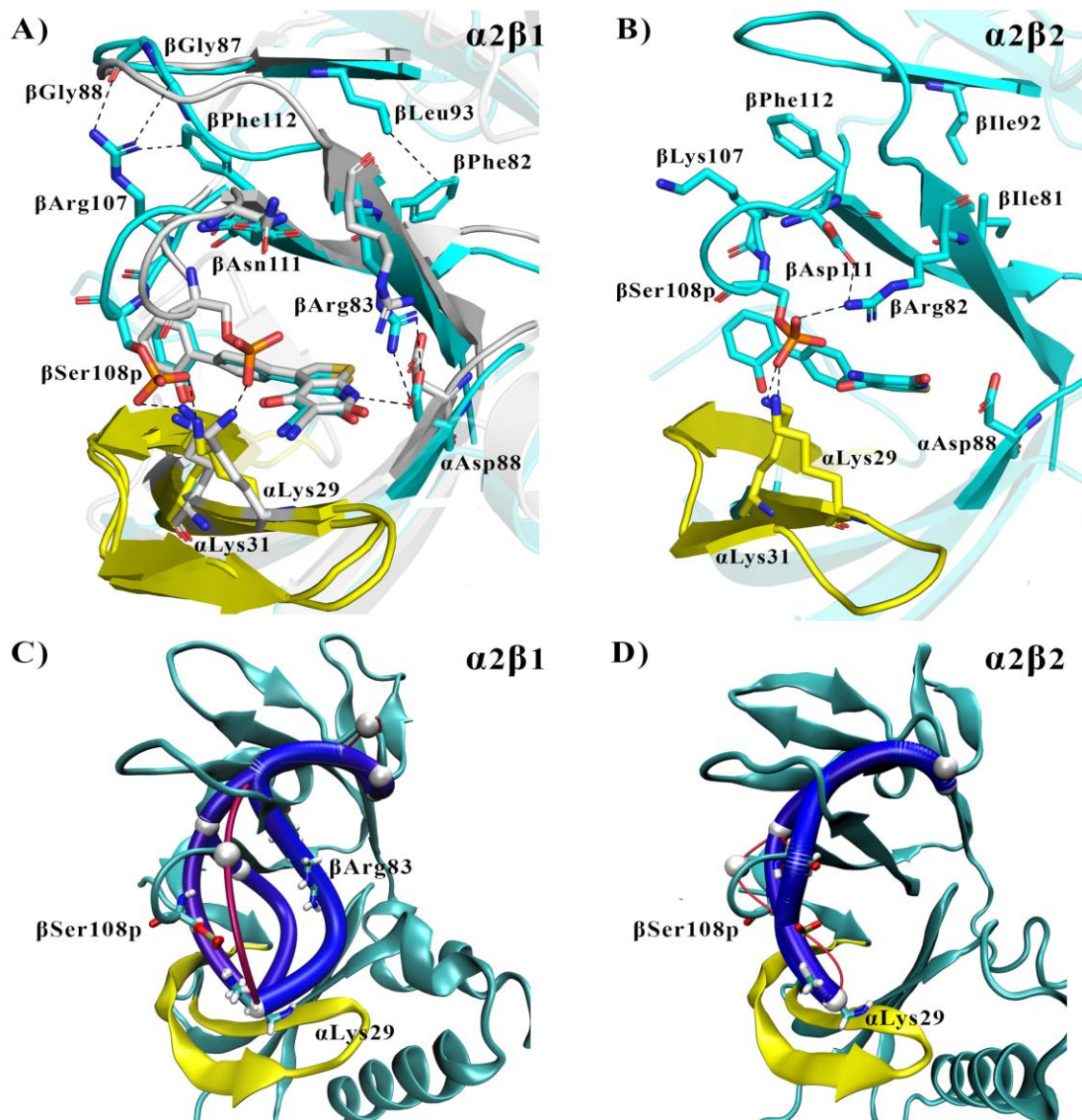


Figure 14. Representation of selected interactions between the CBM, P-loop and A-769662 in holo A) $\alpha 2\beta 1$ and B) $\alpha 2\beta 2$. The two arrangements of β Ser108p in holo $\alpha 2\beta 1$ are shown in cyan and grey (P-loop in yellow). Major interaction networks obtained from WISP analysis for C) $\alpha 2\beta 1$ and D) $\alpha 2\beta 2$ systems.

With regard to the holo forms with SC4, there are two arrangements of $\beta 1$ Arg83 ($\beta 2$ Arg82). In the first case, $\beta 1$ Arg83 ($\beta 2$ Arg82) forms a salt bridge with α Asp88. Alternatively, in the second arrangement, $\beta 1$ Arg83 ($\beta 2$ Arg82) interacts with β Ser108p and forms a cation- π interaction with SC4 and an interaction with β Asp111 (**Figure 15**). This dual arrangement is due to the electrostatic influence of the carboxylate group of SC4. Three different pathways are found with the WISP analysis in both isoforms, involving β Ser108p and the triad $\beta 1$ Arg83

Results summary

(β 2Arg82)-SC4- α Lys29. These findings assist us to understand the mild selectivity of SC4 between α 2 β 1 and α 2 β 2.

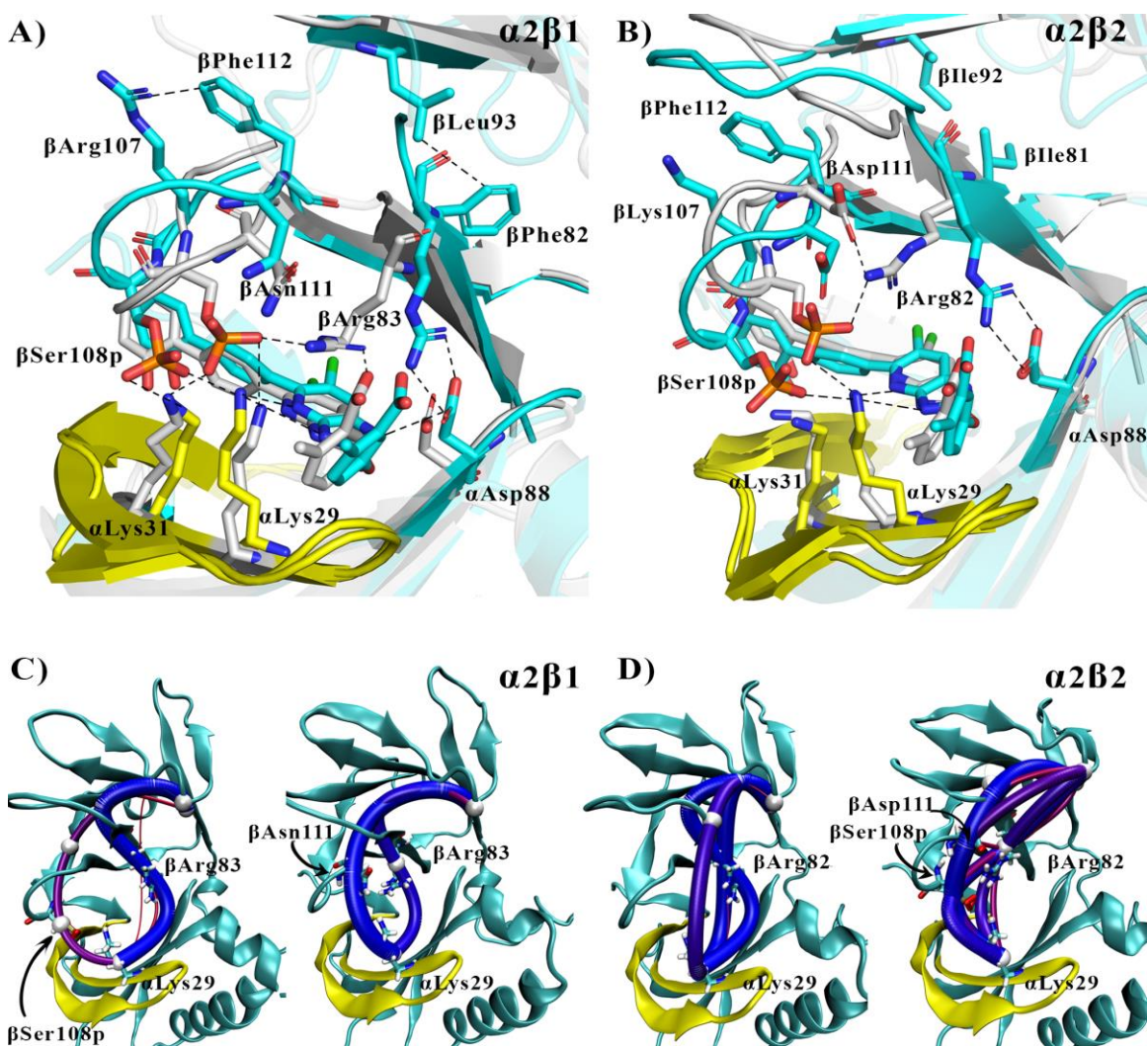


Figure 15. Representation of selected interactions between the CBM, P-loop and SC4 in holo A) α 2 β 1 and B) α 2 β 2. The distinct conformations arrangements are shown in cyan and grey colors (P-loop in yellow). Major interaction networks obtained from WISP analysis for C) α 2 β 1 and D) α 2 β 2 systems.

In the case of PF-739, we can divide the interaction network in two regions. The first one corresponds to the salt bridge interaction formed between the β 1Arg83 with the α Asp88, which at the same time is also forming a hydrogen bond interaction with the PF-739. This pattern is maintained for both α 2 β 1 and α 2 β 2 isoforms (Figure 16). In the second region, the β Ser108p makes salt bridge interactions with α Lys29 and α Lys31. Moreover, an additional contact conserved for all the species along the simulation regards to the hydrogen bond between α Lys31 and the hydroxymethyl-cyclopropyl group.

On the other hand, the motion of the sugar-like mannitol appendage after 500ns of MD simulations involves an important change in the P-loop, favoring a new interaction between α Lys29 of the P-loop and the ligand. This orientational change has a higher critical effect on the $\alpha 2\beta 2$ species due to the substitution of $\beta 1$ Asn111 to $\beta 2$ Asp111. This substitution results in two different conformations of the $\beta 2$ Arg82.

Regarding the WISP analysis for the holo $\alpha 2\beta 1$, the CBM domain and the P-loop are connected via two main pathways. One involves the pathway through the β Ser108p and the ligand, while the second path involves the triad $\beta 1$ Arg83- α Asp88-PF739. However, only one pathway is observed for holo $\alpha 2\beta 2$, which involves the connection between α - and β - subunits through the β Ser108p (**Figure 16**).

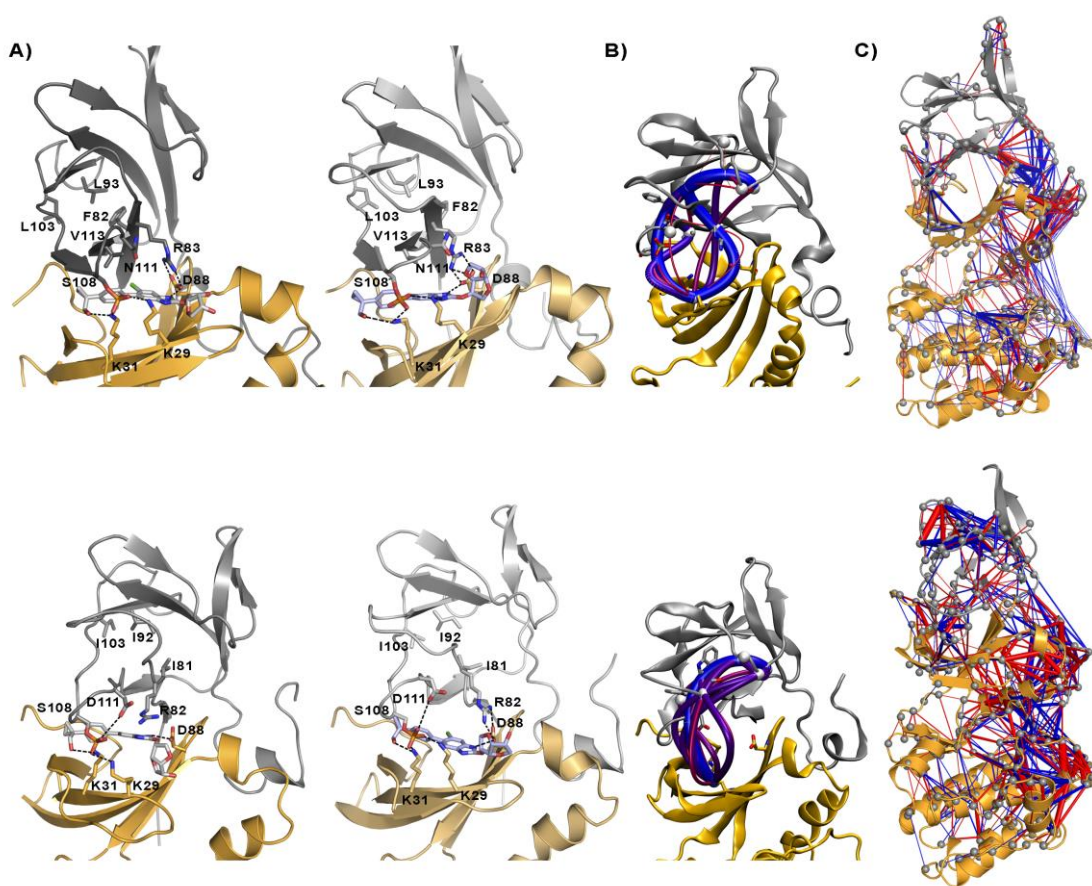


Figure 16. A) Representation of selected interactions between the CBM, P-loop and PF-739 in holo $\alpha 2\beta 1$ (top) and $\alpha 2\beta 2$ (bottom) species. The α -subunit is shown in orange cartoon, while the β -subunit is shown in grey cartoons. PF-739 is shown in grey and light blue sticks in the ADaM site. Selected polar interactions are highlighted as dashed lines. B) Major interaction networks obtained from WISP analysis. C) Subtraction average of the dynamical contact network¹¹¹ of apo – holo states for $\alpha 2\beta 1$ (top) and $\alpha 2\beta 2$ (bottom) species. The edges coloured in red show atomic couples of contacts formed in the holo state with respect to the apo state, while the blue edges represent contacts that are lost in the holo state.

Results summary

4.1.4. Identifying a new AMPK modulator

In this study, we introduced a novel indolic compound as modulator of endothelial AMPK that acts as a mixed-type inhibitor.⁴⁴ It means that IND6 may bind the ATP-binding site leading to competitive inhibition of the enzyme, but also suggest that IND6 may regulate the AMPK activity through an alternate mechanism, presumably involving binding to an additional pocket. The binding modes of IND6 and SBI-0206965, which is another mixed-type inhibitor, have been investigated by MD simulations, enzymatic assays and surface plasmon resonance (SPR).

In the case of SBI-0206965, the compound remains stably bound to the ATP-binding site in all the simulation time and this binding was assisted by two hydrogen bonds between SBI and α Val98 of the hinge region of ATP-binding site (**Figure 17**). On the other hand, the MD simulations of IND6 show larger fluctuations, which is reflected in the separation from α Val98 and α Glu96 to IND6.

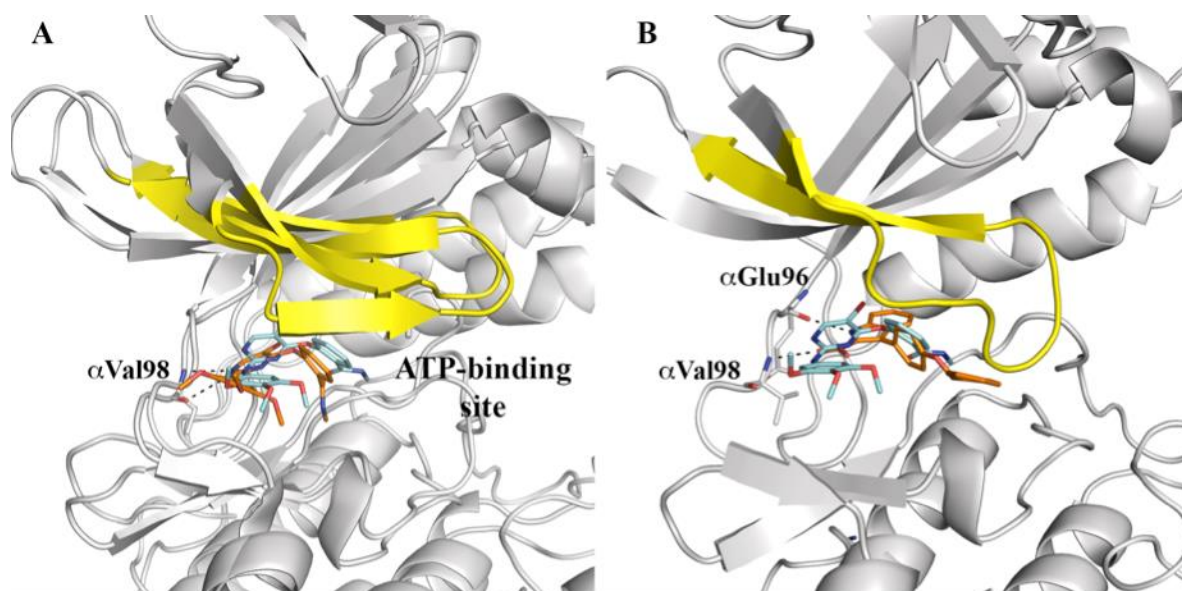


Figure 17. Representation of the binding mode of SBI-020695 and IND6 to the ATP-binding site of AMPK. A) Superposition of the X-ray crystallographic structure of AMPK bound to SBI-020695 and a representative snapshot taken from the MD simulations. B) Representative snapshot of the MD simulation run for the AMPK-IND6 complex. The P-loop is highlighted in yellow. SBI-020695 and IND6 are shown in sticks. C atoms colored in cyan and orange shown the X-ray and final position of the ligand in MD simulation, respectively. Hydrogen atoms have been removed for the sake of clarity.

To investigate the modulatory effect of IND6, we performed other MD simulations where IND6 was located in the ADaM site (**Figure 18**). These studies show that IND6 could adopt two different binding modes. In the first one, IND6 is bound to the hydrophobic cavity in the ADaM site and two salt bridges can be observed between IND6 with α Lys31 and α Lys33. In this case

the top of the P-loop points to the N-terminus of the α C-helix, so the ATP-binding site will be accessible for the binding of ATP mimicking the role of A-769662 in this binding site. In the second case, IND6 adopts a displaced position in the ADaM site, shifted toward the α C-helix and locates on the top of the P-loop, forming electrostatic interactions with α Lys31 and α Lys33 and also, a cation- π interaction with β Arg83. This binding mode results in a structural distortion of the P-loop and makes the ATP-binding site unable to accommodate ATP.

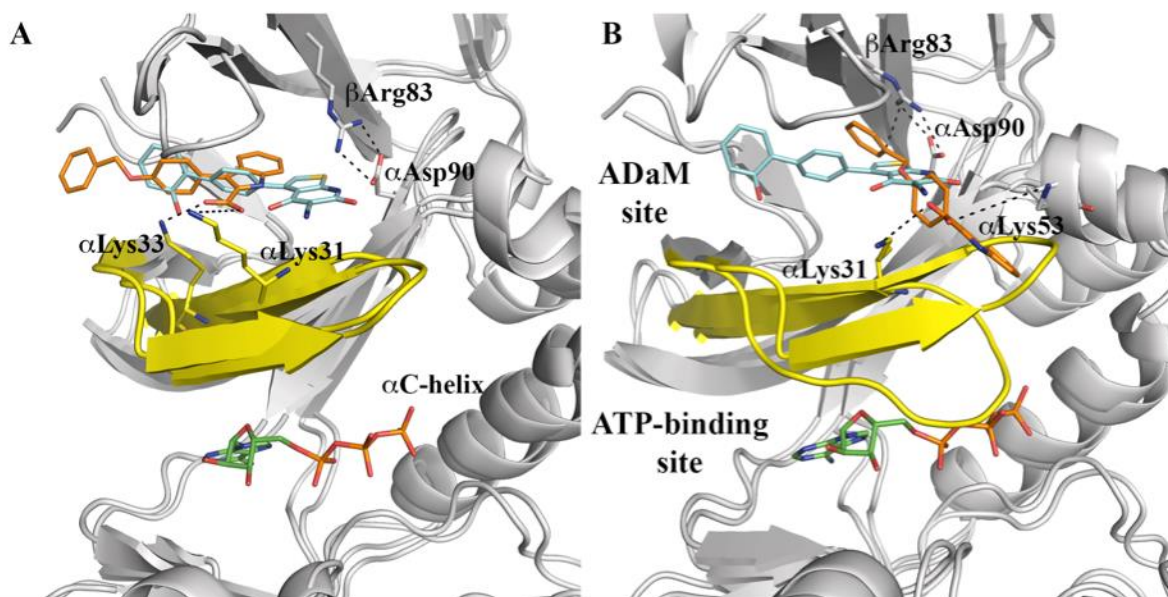


Figure 18. Superposition of the X-ray structure of the ternary complex formed by AMPK bound to A-769662 and ATP (C atoms in cyan and green, respectively), and a representative snapshot of IND6 (C atoms in orange) bound to the ADaM site. (A) Activating binding mode. (B) Non-competitive binding mode. The P-loop is shown in yellow. Selected interactions with protein residues are indicated as dashed lines. Hydrogen atoms have been removed for the sake of clarity.

4.2. Exploring the probable reductase partner for trHbN

As noted in the Introduction, we aimed to identify the potential reductase partner for truncated hemoglobin N (trHbN) in Mtb. Many pathogenic microorganisms have the nitric oxide detoxification (NOD) mechanism mediated by hemoglobins.⁶⁸ In this process, a globin domain in conjugation with a reductase partner performs the conversion of toxic NO produced by macrophages to harmless nitrate anion¹¹² (eq 8).



TrHbN has a potent NO deoxygenase activity despite lacking a reductase domain. The mechanism by which this protein recovers itself during NO detoxification and also, its reductase partner is still unknown.¹¹³ On the other hand, trHbN is not a druggable target due to the lack of typical druggability hallmarks in its structure, which restricts its suitability for drug discovery. However, if we can interfere with the redox process required to regenerate the ferrous species from ferric form at the end of the chemical reaction, the survival of the bacillus would be seriously compromised. Therefore, identifying and characterizing of the reductase partner is crucial for inhibiting its ability to act as a redox partner and preventing its ability to form a functional complex with trHbN.

Heme iron coordination in trHbN is well suited for conducting O₂/NO chemistry for NO dioxygenation,¹¹⁴ which is modulated by two structural features in this protein; i) Pre-A region (12 residues) needed for optimal NOD activity,⁵⁷ and ii) the protein tunnel system, which is formed by short and long branches that facilitates ligand entry to the heme site.¹¹⁵ The O₂ reaches the heme cavity via the tunnel short branch, which triggers dynamical alternation to the protein backbone and regulates the opening of the long branch, so facilitating the access of NO to the O₂-bound heme.⁵⁰

The NOD function relies on the stabilization of O₂ into the deoxygenated heme, rapid migration of NO to the heme cavity, and a fast electron transfer from an electron donor to recover the ferrous state to ferric and restart a new cycle.^{65,66} This project was carried out to answer the question that how trHbN can perform efficiently the NOD function which depends on the structural and functional compatibility with a reductase partner.

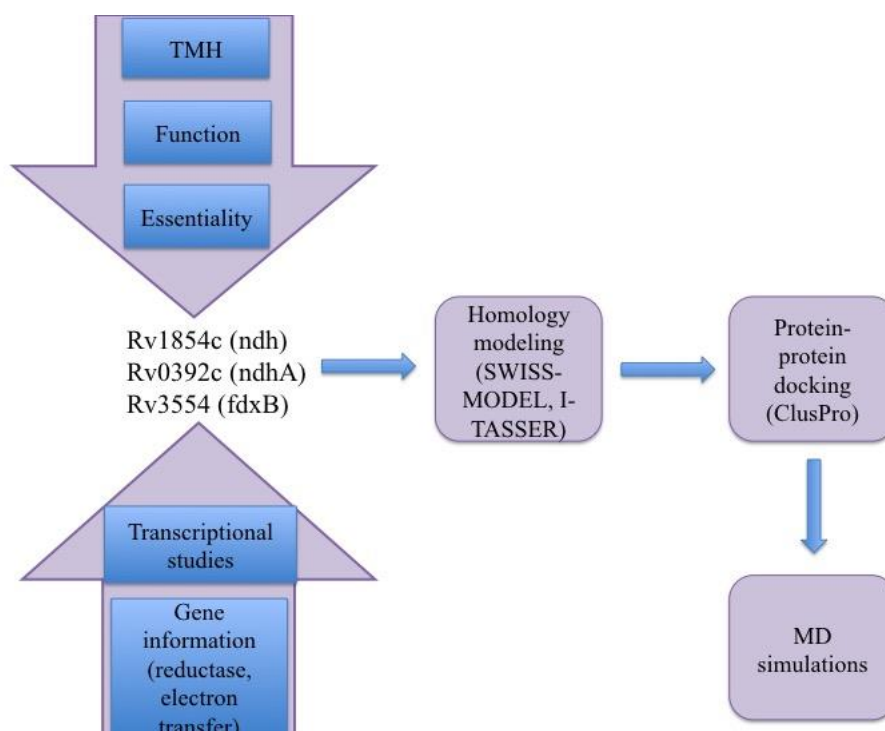
4.2.1. Searching for the reductase candidates

In the first step, since trHbN is found in membrane and the periplasmic space, we searched in the Tuberculist¹¹⁶ database to find the most probable reductase partners with respect to different important characteristics, such as having a transmembrane helix (TMH) attached to the membrane and their function as a reductase or electron transfer proteins. The results of this step

led us to the identification of three putative reductases: two type II NAD dehydrogenases *ndh*, *ndhA*, and one ferredoxin reductase *fdxB* (Scheme 3).

4.2.2. Building the 3D structures

For further investigations, we needed to have the three-dimensional (3D) structures of the candidates. These structures were built by homology modeling using two webserver; SWISS-MODEL⁹⁸ and I-TASSER⁹⁷ excluding the residues related to the TMH, because in this step we needed the part of the protein which interacts with trHbN and adding this part to the protein sequence using to build the 3D structure might introduce some artifact in the final structure. Both of these webserver performed well in the last Critical Assessment of Protein Structure Prediction (CASP) challenge. In particular, I-TASSER has been consistently ranked amongst the top methods in the community-wide CASP experiments for accurate 3D structure prediction.¹¹⁷ Comparison of the results obtained from the two homology modeling algorithms revealed that there are no significant differences between the modeled structures, except in some parts that formed by some linker loops. Because of the better performance of I-TASSER predictor, we continued with the structure obtained from this webserver.



Scheme 3. Schematic information of the bioinformatics strategy undertaken to identify putative reductases relevant for mediating in NO scavenging in *Mtb*.

Results summary

4.2.3. Protein-protein interaction

ClusPro⁹⁹ was used to define the interaction between trHbN (PDB: 1IDR) and the reductase candidates. Amongst the various protein-protein poses, the best ones were chosen based on the population of poses in the docked solutions, the shortest distance from the geometrical centers of heme and FAD, and the lowest energy specified by the algorithm of ClusPro (Table 4). The program uses PIPER program, which is based on the Fast Fourier Transformation correlation approach. In the case of ndh and ndhA, the distances between the heme and FAD cofactor were around 11 Å, whereas the distance in fdxB is enlarged to 25 Å (**Figure 19**). The large distance between the FAD and heme cofactors in trHbN and fdxB occurs due to the accumulation of various loops in the interaction areas of these two proteins. To avoid this problem, we divided the sequence of fdxB into two parts and we built one 3D structure with I-TASSER for each part and then, we docked these two structures by the use of Cluspro webserver to make the complete protein. After accomplishing this step, we again performed the protein-protein interaction between trHbN and fdxB and the distance between cofactors was reduced to 15Å.

Table 4. The characterization of ten protein-protein poses of trHbN-ndh obtained from ClusPro.

Cluster	Members	Representative	Weighted Score
0	95	Center	-649.0
		Lowest Energy	-743.6
1	67	Center	-824.5
		Lowest Energy	-824.5
2	52	Center	-743.8
		Lowest Energy	-765.0
3	45	Center	-714.2
		Lowest Energy	-716.3
4	43	Center	-675.7
		Lowest Energy	-703.6
5	38	Center	-674.3
		Lowest Energy	-742.2
6	32	Center	-646.5
		Lowest Energy	-797.2
7	32	Center	-709.4
		Lowest Energy	-765.4
8	29	Center	-696.7
		Lowest Energy	-744.2
9	27	Center	-747.2
		Lowest Energy	-747.2

The 3D structures and the interaction between trHbN and potential reductase partners are important points to select the most promising candidates; however, other factors need to be considered. The fundamental role of *ndh* in Mtb respiration is supported by a wide range of biochemical¹¹⁸ and transcriptional studies.¹¹⁹ The Mtb genome contains two copies of *ndh* genes (*ndh* and *ndhA*). *Ndh* and *ndhA* share 67% sequence identity. A strain of Mtb in which *ndh* has been disrupted is nonviable,¹²⁰ however a *ndhA* deletion mutant of Mtb can be easily isolated.¹²¹ So, *ndh* is essential for Mtb while *ndhA* could be deleted without causing an adverse effect *in vitro*.¹¹³ Moreover, based on the studies of Yano *et al*,¹²² trHbN appears to be a druggable target as phenothiazines (a class of antipsychotic drugs) are reported to be inhibitors of this protein),¹²³ so this fact could solve the druggability problem of trHbN. Additionally, this protein is absent from the mammalian genome, which reinforces its relevance as a potential target. All these findings led us to choose *ndh* as the potential reductase partner for the next steps.

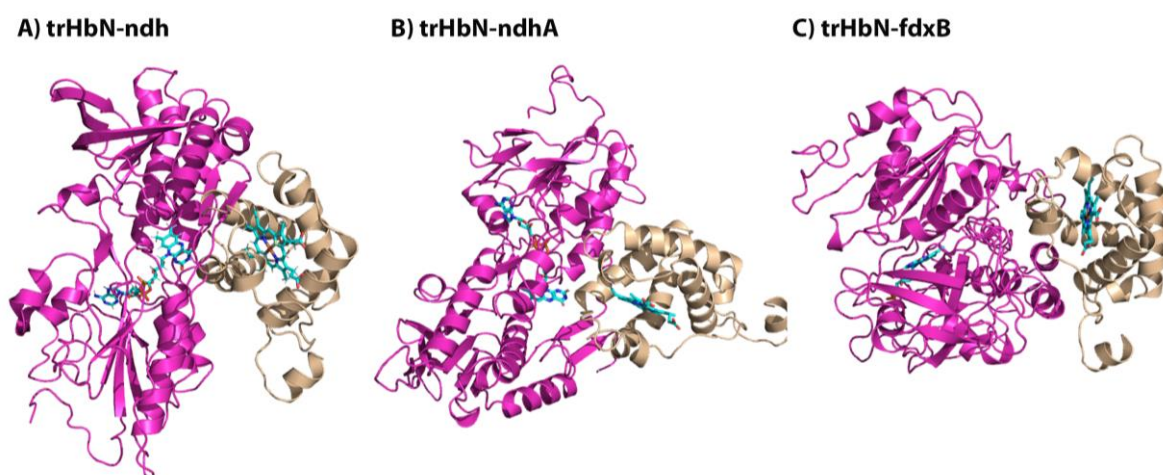


Figure 19. The representation of the dimeric complex trHbN (wheat) – reductase (Magenta) for A) *ndh*, B) *ndhA* and C) *fdxB* obtained after protein-protein docking. The FAD cofactor and heme are indicated with sticks.

In the next step, different MD simulations were performed using AMBER16 package¹²⁴ for six replicas, including three replicas without restraints and three with restraints in the position of FAD in order to avoid the artefactual movements of this cofactor. In the restrained simulations, the RMSD showed that both trHbN and *ndh* proteins fluctuate in the first 200ns of the simulations, and then they remain stable until the end of the simulations with an average RMSD of 2.8 ± 0.4 Å. In the case of the replicas without restraints, two replicas were stabilized after the first 50ns of simulations and the third one after 300ns of MD.

The final snapshots of one of the replicas without restraints were selected to add the NADH cofactor, and therefore one MD simulation was performed for 500ns with the presence of NADH and FAD. The RMSD value of the complex containing all cofactors was larger than the

Results summary

previous RMSD. This increase is basically due to the RMSD of the ndh reductase, which needs to adapt its conformation to the NADH cofactor. The RMSD of trHbN was stable along 500ns of the simulation. At the end of this step, we could achieve a stable and refined complex with an acceptable distance between FAD and heme in order to analyze the electron transfer rate between these two cofactors.

The estimation of electron transfer (ET) was performed by the Beratan algorithm.^{125,126} The ET rate was determined for 4000 frames along the trajectory (**Figure 20**). The results indicated an ET rate of 3.4×10^{-5} eV. This rate is comparable with the electron coupling reported for *E.coli* flavoHb (7.1×10^{-5} eV), a system that is known to cycle electrons very efficiently. The analysis of ET pathways helped us to identify the important residues for electron transfer; so, Met87 and Phe90 with water molecules have a major contribution to the ET path. Water molecules play an important role in mediating the transition between FAD and Phe90, but if Met87 involves in the transfer is very likely to replace the water as a bridge between the FAD and Phe90.

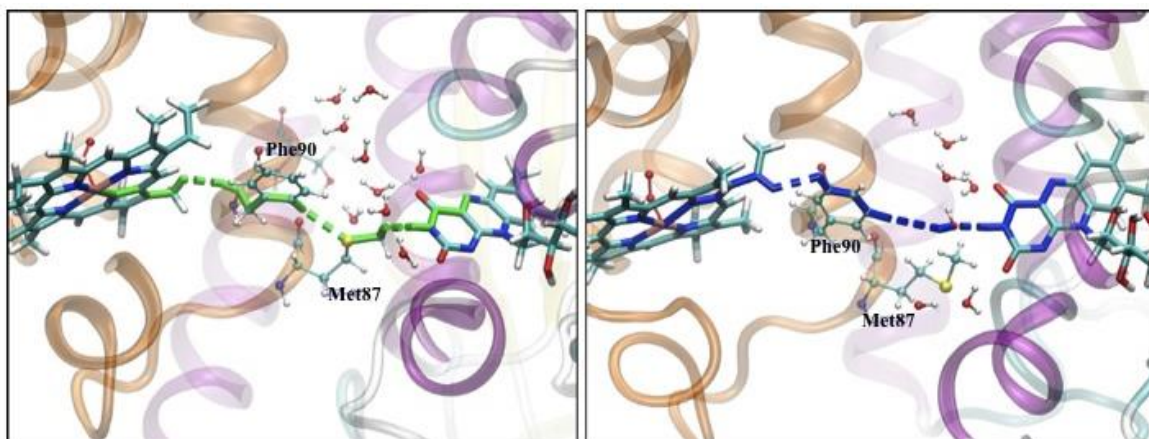


Figure 20. Representation of the major electron transfer pathways. Heme, FAD, Met87, and Phe90 are shown as cyan sticks.

With respect to all this information, we showed a compatible redox partner to maintain the trHbN's function, presenting ndh as a potential reductase partner for trHbN, and hence a promising target against TB (**Figure 21**). All in all, we have the 3D model of this candidate and the best pose of interaction between this partner and trHbN. Moreover, a potential binding pocket in the interface between both proteins, key residues at this interface and directly related to the ET path have been identified, which could be utilized for the drug design strategies.

On the other hand, we are trying to express and purify these two proteins in order to estimate the electron transfer rate and confirm our hypothesis in vitro environment.

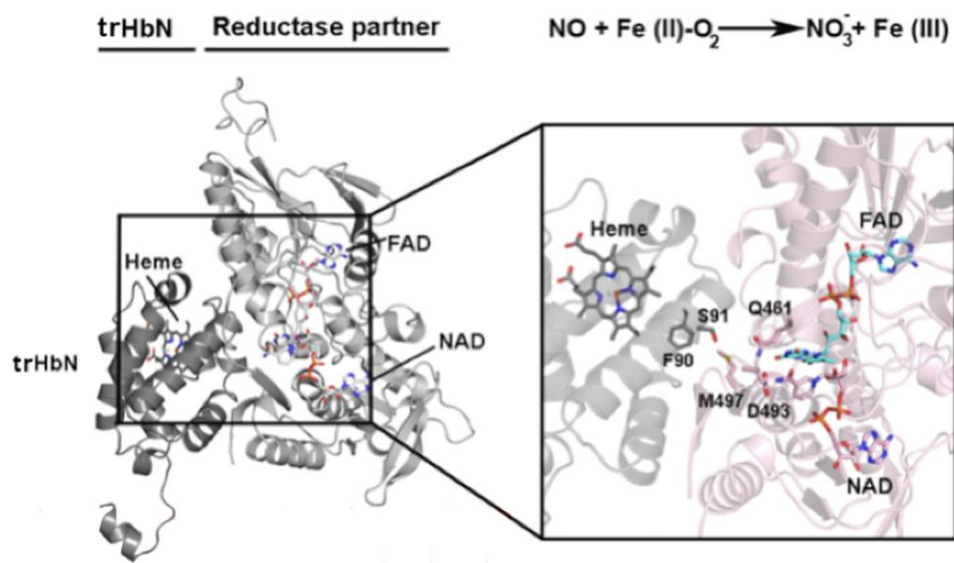


Figure 21. Representation of the in silico 3D model of the complex between ndh and trHbN. The position of the heme cofactor in trHbN, and both FAD and NADH in ndh is shown, together with a schematic representation of the arrangement relative to the Mtb membrane. Selected residues that may participate in the protein-protein interface and mediate electron transfer are shown on the right.

CHAPTER

5

DISCUSSION

5. Discussion

5.1. AMPK

5.1.1. Study of the direct activation mechanism of AMPK

In the first part of our studies, we have disclosed the direct activation mechanism of AMPK by A-769662, which binds to the ADaM site located at the interface of α - and β - subunits. This binding cavity is not recognized in other kinases, so AMPK has gained interest as a potential therapeutic target, especially in the context of pathological dysfunction.

A-769662 activates AMPK independently of α -Thr172 phosphorylation; instead it needs the β Ser108 to be phosphorylated. Binding of A-769662 enhances the $\alpha 2\beta 1$ isoform of AMPK up to more than 90-fold when p β Ser108 is present. On the other hand, A-769662 demonstrates selectivity for $\alpha 2\beta 1$ and it is not active in isoform $\alpha 2\beta 2$, while another activator, SC4, can activate both $\beta 1$ and $\beta 2$ -containing AMPK complexes, although a slightly higher activation is observed in $\alpha 2\beta 1$. These finding raises a question about the underlying molecular mechanisms implicated in the enzyme activation. Therefore, in the second part of our research, we focused on the role played by β subunit in the selectivity of different isoforms of AMPK via binding of A-769662, SC4 and PF-739.

We can conclude that the presence of the activator and ATP affects the protein flexibility, specially regarding the motion of the N-terminus of the α subunit and the regulatory domain of the β subunit, which are more sensitive to the presence of ligand. Also, a decrease in the RMSF values is observed for the region that encloses the P-loop, since the activator is located at the interface formed by the CBM domain and the P-loop. This effect is less apparent for the $\alpha 2\beta 2$ complexes, suggesting that the activator induces a weaker structural stabilization upon binding to $\alpha 2\beta 2$ isoform.

Comparison the results obtained from ED analysis shows a significant reduction in the overall flexibility of the protein in holo+ATP systems, especially regarding the P-loop and the α C and C-interacting helices but also in the elements of the β subunit proximal to the ADaM site, which are much less flexible in the holo states. These findings reflect the synergy between activator and ATP in increasing the stiffness of the $\alpha 2\beta 1$ AMPK complex and can be viewed as a mechanism to translate the binding energy in to changes in the mechanical sensitivity of the protein dynamics that should facilitate the enzymatic catalysis of AMPK. The dynamical observed for $\alpha 2\beta 1$ is less regular in $\alpha 2\beta 2$, as the structural variance of the first essential motion in the apo system remains unaltered or is even increased in the presence of the activator. In the holo+ATP system of $\alpha 2\beta 2$ there is lack of a consistent pattern of structural flexibility.

The output of the DCCM and the RMSD of ATP-binding sites are in agreement with the previous results. Inspection of the DCCM matrices allows the examination of the impact of ligand binding on the dynamical behavior of the enzyme. The gross features of the apo systems can be recognized in the matrices of holo and holo+ATP, but there is a significant reduction in the extent of the correlated motions. Moreover, in the plots of the holo, we can see a reinforcement in the correlation between the motion of the CBM domain and the P-loop, which suggests that binding of the activator has a direct influence on the preorganization of the ATP-binding site and the conformational rearrangement of the P-loop can enable the adoption of a proper conformation well suited for binding of ATP. Although an overall gross similarity exists in the dynamical behavior of both $\alpha 2\beta 1$ and $\alpha 2\beta 2$, the results show some subtle differences between the DCCM of these two isoforms. This information suggests that the $\alpha 2\beta 2$ complex has a larger resilience to the structural modulation exerted by the activator, while $\alpha 2\beta 1$ is more sensitive to the conformational changes induced upon binding of the activator.

From the RMSD of the ATP-binding site along the trajectory, we can see that in the case of holo+ATP, the residues shaping the ATP-binding site (which includes the P-loop), sample a reduced conformational space with a peak in the narrow population (1.0-1.6Å) for both $\alpha 2\beta 1$ and $\alpha 2\beta 2$. The most distinctive features emerge from the distribution profiles of the holo species. For $\alpha 2\beta 1$, binding of the activator shifts the holo distribution toward the holo+ATP, as noted in the narrowing of the distribution profile with regard to the apo species. This finding represents the concept that the activator acts as a molecular glue that connects the fluctuations of the CBM domain in the β subunit with the conformational flexibility of the P-loop and enhances the tendency to preorganize the P-loop for binding to ATP. In contrast, in $\alpha 2\beta 2$ complexes, the activator leads to structures where the ATP-binding site has less resemblance to the holo+ATP system. These differences mainly arise from the P-loop, saying that the activator is less effective in inducing the remodeling of the P-loop.

In order to find the reason of the selectivity of $\alpha 2\beta 1$ and $\alpha 2\beta 2$ complexes with A-769662, SC4, and PF-739, key residues required for the formation of the allosteric network that connects the ADaM and ATP-binding sites through interaction with ligands were studied. In particular, the interactions between α Asp88 and $\beta 1$ Arg83 and between β pSer108 and α Lys29 are crucial for the enzyme activation in the case of $\alpha 2\beta 1$ complexes with A-769662. The different interaction networks observed upon binding of A-769662 to $\alpha 2\beta 1$ supports the effective role of the activator in assisting synergistically the fluctuation of the CBM unit and the P-loop which facilitates the preorganization of the P-loop for binding to ATP. This effect is weakened in $\alpha 2\beta 2$ species in which the $\beta 1$ Asn111 is substituted with $\beta 2$ Asp111. $\beta 2$ Asp111 competes with α Asp88 in forming electrostatic interactions with $\beta 2$ Arg82. In holo $\alpha 2\beta 1$, the CBM domain and the P-loop are connected via three pathways, whereas only the β Ser108 pathway is found in holo

Discussion

$\alpha 2\beta 2$. These changes highlight the inefficacy of A-769662 in acting as a molecular glue between the CBM domain and the P-loop. On the other hand, the changes of residues $\beta 1\text{Phe}82$, $\beta 1\text{Leu}93$ and $\beta 1\text{Leu}103$ to $\beta 2\text{Ile}81$, $\beta 2\text{Ile}92$ and $\beta 2\text{Ile}103$, which are embedded in a hydrophobic cluster in the CBM domain, result in enlarged fluctuations of the hydrophobic core in $\alpha 2\beta 2$ and the conformational change observed for $\beta 2\text{Tyr}125$ and $\beta 2\text{Phe}96$.

In the case of SC4, there is a dual interaction network that connects the CBM domain and the P-loop via $\beta\text{Ser}108\text{p}$, specially the $\beta\text{Arg}83\text{-SC}4\text{-}\alpha\text{Lys}29$ triad. These findings can pay a way to understand the mild selectivity between $\beta 1$ and $\beta 2$ containing AMPK. This trait can be attributed to the electrostatic effect of the carboxylate group of SC4 assisting the adoption of the two arrangements of $\beta 1\text{Arg}83$ ($\beta 2\text{Arg}82$), relieving the effect of the $\beta 1\text{Asn}111$ substitution. This likely contributes to explain the 6-fold reduction in the activation potency determined for $\alpha 2\beta 2\gamma 1$ upon substitution of the SC4 pyridine-like nitrogen in the imidazo[4,5-b] pyridine ring for carbon.

Regarding PF-739, we can observe the interaction networks in two regions, one corresponds to the $\beta 1\text{Arg}83$ and $\alpha\text{Asp}88$ and the second region is related to the interactions formed by $\beta\text{Ser}108\text{p}$. The motion of the PF-739 along the simulations results in an important change in the P-loop, favoring new interactions between $\alpha\text{Lys}29$ and the P-loop and also $\beta 1\text{Arg}83$ and the ligand. The orientational change of the ligand has a higher critical effect on the $\alpha 2\beta 2$ due to the substitution of $\beta 1\text{Asn}111 \rightarrow \beta 2\text{Asp}111$, which provokes two different conformations of the $\beta 2\text{Arg}82$. In the case where the sugar-like mannitol group is facing towards the P-loop, the $\beta 2\text{Arg}82$ is forming an interaction with the $\beta 2\text{Asp}111$, losing all the contact with the $\alpha\text{Asp}88$. Also, the $\beta 2\text{Asp}111$ is in contact with the $\beta\text{Ser}108\text{p}$, which maintains its interactions with both $\alpha\text{Lys}29$ and $\alpha\text{Lys}31$ from the P-loop. These interactions reinforce the connection between α - and β -subunits. However, when the sugar-like mannitol is facing up, the $\beta 2\text{Arg}82$ forms the salt bridge with the $\alpha\text{Asp}88$ and loses all the contact with the $\beta 2\text{Asp}111$. This conformation is less stable and less sampled during the simulation. These observations are also found in the major interaction networks. In holo $\alpha 2\beta 1$, there are two pathways that connect the CBM domain and the P-loop, one involves the $\beta\text{Ser}108\text{p}$ and the ligand and the other one involves the triad $\beta 1\text{Arg}83\text{-}\alpha\text{Asp}88\text{-PF-739}$. In holo $\alpha 2\beta 2$, there is only one pathway through $\beta\text{Ser}108\text{p}$. These results agree with the interaction formed between $\beta 2\text{Arg}82$ and $\beta 2\text{Asp}111$, which weakened considerably the interaction between $\beta 2\text{Arg}82$ and $\alpha\text{Asp}88$. These results can explain the higher affinity of PF-739 towards the $\beta 1$ -isoform.

5.1.2. AMPK modulator

In the third part of our studies, which was in collaboration with the research group of Profs. A. Castro and M. S. Fernández, we investigated a novel indolic derivate, IND6. IND6 seems to behave as a paradoxical activator of the endothelial AMPK $\alpha 1\beta 1\gamma 1$ in EA.Hy926 cells, although exhibits a mixed-type inhibition in the enzymatic assays. The binding mode of the compounds has been assessed by combining MD simulations, enzymatic and SPR assays.

After screening of the compounds for binding to the AMPK $\alpha 1\beta 1\gamma 1$ using SPR technique, their effect on the enzymatic activity of AMPK was assessed through a luminescent assay with the recombinant AMPK. All the compounds reduce the activity of AMPK $\alpha 1\beta 1\gamma 1$ at 30 μ M, therefore they can be considered as inhibitors. IND6 was selected as representative compound to perform a detailed evaluation of the biological effect on the enzyme activity. Then, IND6 was subjected to enzymatic kinetic analysis to examine its competition with ATP. On the other hand, SPR sensorgrams showed that IND6 presented an additive effect on the binding to AMPK, proposing that it can bind at different binding sites.

Due to its mixed-type inhibition, we investigated the binding of IND6 (and SBI-0206965 as reference) to the ATP-binding site using MD simulations. In the simulations related to SBI-0206965, this compound remains stably bound in the ATP-binding site assisted by two hydrogen bonds, in contrast the MD simulations for IND6 reveals larger fluctuations of the ligand in the binding site. This behavior can be related to the flexibility of the benzyloxy moiety and the non-planarity of the central benzene ring of the indole ring. Also, it enhances the fluctuations of the P-loop results in the enhancement of the hydrogen bond distances formed between IND6 with α Val98. These traits agree with the 27-fold lower potency of IND6 relative to SBI-0206965. This information agrees with the ability of IND6 to bind to ATP-binding site leading to competitive inhibition of the enzyme.

To examine the binding of IND6 to the ADaM site, additional MD simulations were carried out, leading to the identification of two distinct binding modes. In the first case, IND6 is deeply bound into the hydrophobic cavity of the ADaM site. It leaves the ATP-binding site accessible for the binding of ATP and the conformation sampled for the P-loop is similar to the AMPK complex bound to A-769662. So, IND6 might mimic the role of A-769662 in this binding mode. In the second binding mode, IND6 protrudes from the ADaM site toward the α C-helix and locates on the top of the P-loop. This binding mode induces a structural distortion of the P-loop, making the ATP-binding site unable to accommodate ATP and it can explain the non-competitive mechanism of the mixed-type inhibition on IND6.

5.2. Exploring the reductase partner for trHbN

With an estimated of 10.4 million people infected with tuberculosis, it is the ninth leading cause of death worldwide, causing 1.72 million deaths in 2018 (WHO). Multidrug resistance, drug side effects, and co-infections make it necessary to find novel and safety antimycobacterial agents. Our project aimed to search for novel therapeutic approaches targeting the molecular basis of the nitrosative stress resistance of Mtb. In particular, the aim is to target the putative reductase partner required to sustain the NO scavenging activity and perturb both the electron respiratory chain in Mtb and the NO scavenging activity that contributes to the survival of the bacillus. TrHbN converts NO into harmless nitrate anion in a process with the oxidation of heme Fe(II) to Fe(III). Since the conversion of NO to nitrate requires the oxidation of ferrous to ferric state, we attempted to discern the potential influence of putative reductases to regenerate efficiently the ferrous form of trHbN.

Considering this hypothesis, in the first step, we followed an integrative approach that combined i) the structural and functional information available for trHbN, particularly regarding the localization in the cell membrane and the cell wall, ii) a bioinformatics analyses of the Mtb genome searching for reductase and electron transfer proteins, and iii) a search of the data available regarding transcriptional changes in early phases of infection.^{127–129} The analysis led to three putative reductases: two type II NAD dehydrogenase, *ndh* and *ndhA*, and a ferredoxin reductase, *fdxB*.

After recognition of the proteins, we built the 3D structures with two different programs with the highest accuracy, I-TASSER and SWISS-MODEL. The output of these 3D predictors helped us to solve the 3D structures of three candidates, providing accurate information about the specific fold of the protein. The results of both predictors indicate that there are no significant differences between the two structures obtained. The RMSD of 2Å for the backbone atoms confirm the highest similarity between the 3D structures built by I-TASSER and SWISS-MODEL. Since I-TASSER is the best server for full-length 3D protein structural predictor based on CASP reports,¹¹⁷ we chose the I-TASSER model and we performed the protein-protein docking to prove the trHbN-reductase interaction and to identify the factors implicated in it.

Amongst the 10 poses obtained from ClusPro webserver for each reductase-trHbN complex, we paid attention to the distance between FAD and Heme to accomplish the electron transfer efficiently and also the lowest energy as specified by using balanced ClusPro algorithm. In the case of *ndh* and *ndhA*, the distance from the geometrical centers of heme and FAD was around 11Å, but regarding the *fdxB*, it was enlarged to 25Å. This large distance was affected by the predicted 3D structure, especially regarding the arrangement of loops in the interface, but a remodeling study led to a new 3D structural model where the distance between FAD and heme. was to 15Å.

For the next step, we decided to choose one candidate to perform the MD simulations. Among the three, the best candidate for drug discovery is *ndh* due to these findings: i) *ndh* and *ndhA* share a high degree of amino acid similarity, but gene knockout showed that *ndh* gene is essential while *ndhA* could be deleted in *Mtb* without any adverse effect, ii) *ndh* is druggable, as noted in a recent identification of two compounds targeting *ndh*,¹³⁰ and iii) *ndh* is absent from the mammalian genome and can be considered as a potential target.

Based on these factors, we selected *ndh* and we performed different MD simulations with and without restraints in the position of FAD. After preliminary rearrangements, the complex was found to remain stable until the end of the simulations, resulting in an average RMSD of $2.8 \pm 0.4 \text{ \AA}$. The regions with the highest flexibility correspond to residues 100-121 of trHbN and residues 300-350 and 370-400 of *ndh*, which are involved in the protein-protein interface and loop structures, which are located in the outer parts of the protein.

The ET between two proteins was estimated to have an electron transfer rate of $3.7 \times 10^{-5} \text{ eV}$, which is acceptable in light of the results reported for other systems that are known to cycle electrons very efficiently.¹²⁶ There are two main ET pathways that link the isoalloxazine ring of FAD in *ndh* to the heme cofactor of trHbN involving residues Met87 and Phe90 of trHbN. Water molecules play an important role in mediating the transition between FAD and Phe90, however when Met87 participates in the path, there are usually no water molecules involved, which confirm the importance of these residues in conducting the ET. So, blocking the path of these residues may affect the potential of the reductase partner.

Finally, a potential binding pocket has been identified in the interface between both proteins that can be used for the drug discovery strategies.



CHAPTER

CONCLUSION

6. Conclusion

This section summarizes the main conclusions found in the course of this work:

6.1. AMPK

- The presence of the activator in $\alpha 2\beta 1$ isoform of AMPK affects the protein flexibility particularly regarding the motion of the N-terminus of the catalytic α -subunit and the regulatory domain of the β -subunit.
- In $\alpha 2\beta 1$ isoform, the activator acts as a glue, filling the space between the β -subunit and the N-terminal domain of the α -subunit, favoring the binding of ATP and results in the increase of the AMPK activity.
- The shape and the size of the ATP-binding pocket are allosterically regulated by the binding of activator and tend to adopt a topology well suited for ATP binding.
- The results of the structural analysis revealed key residues required for the formation of the allosteric network; the interaction between Asp88 and β Arg83 and between β Ser108 and α Lys29 are critical for the enzyme activation.
- The role of the activator in facilitating the adoption of pre-organized conformations of the ATP-binding site is weakened in holo $\alpha 2\beta 2$, which results in decreased efficacy of A-769662.
- The lower efficacy of the activator in $\alpha 2\beta 2$ isoform is related to the occurrence of the $\beta 1$ Asn111 \rightarrow $\beta 2$ Asp111 substitution, which affects the pattern of interaction formed by the activator upon binding to the ADaM site.
- Changes of residues $\beta 1$ Phe82, $\beta 1$ Leu93 and $\beta 1$ Leu103 to $\beta 2$ Ile81, $\beta 2$ Ile92 and $\beta 2$ Ile103 which are embedded in a hydrophobic cluster in the CBM domain, leads to enlarged fluctuation of the hydrophobic core in the holo $\alpha 2\beta 2$ and also the conformational change observed for $\beta 2$ Tyr125 and $\beta 2$ Phe96 resulting in local remodeling of the CBM unit.
- In the case of SC4, the mild selectivity between $\beta 1$ and $\beta 2$ -containing AMPK can be explained by the resemblance of the dual interaction networks that connect the CBM unit and the P-loop via β Ser108p and particularly the β Arg83-SC4- α Lys29 triad.
- The electrostatic influence of the carboxylate group present in SC4 facilitates the adoption of the two arrangements of β Arg83 (β Arg82), tempering the effect of the $\beta 1$ Asn111 \rightarrow $\beta 2$ Asp111 substitution.

Conclusion

- The single substitution of $\beta 1\text{Asn111} \rightarrow \beta 2\text{Asp111}$ can change the interaction networks formed surrounded the activator and induce better mechanical responses of the $\alpha 2\beta 1$ isoform towards the interaction of PF-739 in comparison to the $\alpha 2\beta 2$.
- Even in the case of a pan-activator like the PF-739, still subtle residue substitutions in the ADaM site are responsible for difference in affinity towards the isoforms.
- We were able to indicate key molecular features that mediate the selective activation of $\alpha 2\beta 1$ and $\alpha 2\beta 2$ isoforms, explaining the different selectivity of A-769662, SC4 and PF-739.

6.2. TrHbN

- We demonstrated a novel mechanism by which trHbN interacts with a compatible redox partner to maintain its NOD activity. In particular, ndh emerges as a potential partner for trHbN function.
- The complex formed between ndh and trHbN proteins remain stable during the MD simulations, enabling the identification of the residues that define the protein-protein interface.
- The assessment of the ET pathways helped us to identify important residues for the electron transfer. Met87 and Phe90 with water molecules have a major contribution to the ET path.
- Potential binding pocket in the interface between both proteins, key residues at this interface and directly related to the ET path have been identified, which could be utilized for the drug design strategies

BIBLIOGRAPHY

Bibliography

- (1) World Health Organization. The top 10 causes of death <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death> (accessed Mar 17, 2020).
- (2) Figueras, A.; Laporte, J. R. Failures of the Therapeutic Chain as a Cause of Drug Ineffectiveness. *Br. Med. J.* **2003**, *326* (7395), 895–896. <https://doi.org/10.1136/bmj.326.7395.895>.
- (3) Dcosta, V. M.; King, C. E.; Kalan, L.; Morar, M.; Sung, W. W. L.; Schwarz, C.; Froese, D.; Zazula, G.; Calmels, F.; Debruyne, R.; et al. Antibiotic Resistance Is Ancient. *Nature* **2011**, *477* (7365), 457–461. <https://doi.org/10.1038/nature10388>.
- (4) Bush, K.; Courvalin, P.; Dantas, G.; Davies, J.; Eisenstein, B.; Huovinen, P.; Jacoby, G. A.; Kishony, R.; Kreiswirth, B. N.; Kutter, E.; et al. Tackling Antibiotic Resistance. *Nat. Rev. Microbiol.* **2011**, *9* (12), 894–896. <https://doi.org/10.1038/nrmicro2693>.
- (5) Alberts B, Johnson A, Lewis J, et al. *Molecular Biology of the Cell*; 2002.
- (6) Berg, J. M.; Stryer, L.; Tymoczko, J. L. Biochemistry / Jeremy M. Berg, John L. Tymoczko, Lubert Stryer. **2012**.
- (7) Kuzmanov, U.; Emili, A. Protein-Protein Interaction Networks: Probing Disease Mechanisms Using Model Systems. *Genome Med.* **2013**, *5* (4), 1–12. <https://doi.org/10.1186/gm441>.
- (8) Safari-Alighiarloo, N.; Taghizadeh, M.; Rezaei-Tavirani, M.; Goliaei, B.; Peyvandi, A. A. Protein-Protein Interaction Networks (PPI) and Complex Diseases. *Gastroenterol. Hepatol. from Bed to Bench* **2014**, *7* (1), 17–31. <https://doi.org/10.22037/ghfbb.v7i1.511>.
- (9) Stank, A.; Kokh, D. B.; Fuller, J. C.; Wade, R. C. Protein Binding Pocket Dynamics. *Acc. Chem. Res.* **2016**, *49* (5), 809–815. <https://doi.org/10.1021/acs.accounts.5b00516>.
- (10) Proteomics: studying proteins to fight diseases <https://publichealthmatters.blog.gov.uk/2014/03/17/proteomics-studying-proteins-to-fight-disease/> (accessed Mar 17, 2020).
- (11) Tekir, S. D.; Çakir, T.; Ülgen, K. Ö. Infection Strategies of Bacterial and Viral Pathogens through Pathogen-Human Protein-Protein Interactions. *Front. Microbiol.* **2012**, *3* (FEB), 1–11. <https://doi.org/10.3389/fmicb.2012.00046>.
- (12) Crespo, A.; Martı, M. A.; Kalko, S. G.; Morreale, A.; Orozco, M.; Gelpi, J. L.; Luque, F. J. Theoretical Study of the Truncated Hemoglobin HbN : Exploring the Molecular Basis of the NO Detoxification Mechanism. **2005**, No. 7, 4433–4444. <https://doi.org/10.1021/ja0450004>.
- (13) Gonzalez, M. W.; Kann, M. G. Chapter 4: Protein Interactions and Disease. *PLoS Comput. Biol.* **2012**, *8* (12). <https://doi.org/10.1371/journal.pcbi.1002819>.
- (14) Notes, L. Bioinformatics III Structural Bioinformatics and Genome Analysis. *Bioinformatics* **2007**, *2007* (August 2006).
- (15) Núñez, S.; Venhorst, J.; Kruse, C. G. Target-Drug Interactions: First Principles and Their Application to Drug Discovery. *Drug Discov. Today* **2012**, *17* (1–2), 10–22. <https://doi.org/10.1016/j.drudis.2011.06.013>.
- (16) Copeland, R. A. The Drug–Target Residence Time Model: A 10-year Retrospective. **2016**, *15* (February), 87–95.
- (17) Copeland, R. A.; Pompliano, D. L.; Meek, T. D. PE R S PEC T I V E S Drug – Target Residence Time and Its Implications for Lead Optimization.
- (18) Roskoski, R. A Historical Overview of Protein Kinases and Their Targeted Small Molecule Inhibitors. *Pharmacol. Res.* **2015**, *100*, 1–23. <https://doi.org/10.1016/j.phrs.2015.07.010>.
- (19) Neumann, D.; Viollet, B. *Methods and Protocols*; 2018; Vol. 2. <https://doi.org/10.1093/cdn/nzy042>.
- (20) Xiao, B.; Sanders, M. J.; Underwood, E.; Heath, R.; Carmena, D.; Jing, C.; Walker, P. A.; Eccleston, J. F.; Lesley, F.; Saiu, P.; et al. Europe PMC Funders Group Structure of Mammalian

- AMPK and Its Regulation by ADP. *Nature* **2011**, 472 (7342), 230–233.
<https://doi.org/10.1038/nature09932>.Structure.
- (21) Aledavood, E.; Moraes, G.; Lameira, J.; Castro, A.; Luque, F. J.; Estarellas, C. Understanding the Mechanism of Direct Activation of AMP-Kinase: Toward a Fine Allosteric Tuning of the Kinase Activity. *J. Chem. Inf. Model.* **2019**, 59 (6), 2859–2870.
<https://doi.org/10.1021/acs.jcim.8b00890>.
 - (22) Day, E. A.; Ford, R. J.; Steinberg, G. R. AMPK as a Therapeutic Target for Treating Metabolic Diseases. *Trends Endocrinol. Metab.* **2017**, 28 (8), 545–560.
<https://doi.org/10.1016/j.tem.2017.05.004>.
 - (23) Luo, Z.; Saha, A. K.; Xiang, X.; Ruderman, N. B. AMPK, the Metabolic Syndrome and Cancer. *Trends Pharmacol. Sci.* **2005**, 26 (2), 69–76. <https://doi.org/10.1016/j.tips.2004.12.011>.
 - (24) Penfold, L.; Woods, A.; Muckett, P.; Nikitin, A. Y.; Kent, T. R.; Zhang, S.; Graham, R.; Pollard, A.; Carling, D. CaMKK2 Promotes Prostate Cancer Independently of AMPK via Increased Lipogenesis. *Cancer Res.* **2018**, 78 (24), 6747–6761. <https://doi.org/10.1158/0008-5472.CAN-18-0585>.
 - (25) Hardie, D. G. AMP-Activated Protein Kinase as a Drug Target. *Annu. Rev. Pharmacol. Toxicol.* **2007**, 47 (1), 185–210. <https://doi.org/10.1146/annurev.pharmtox.47.120505.105304>.
 - (26) Calabrese, M. F.; Rajamohan, F.; Harris, M. S.; Caspers, N. L.; Magyar, R.; Withka, J. M.; Wang, H.; Borzilleri, K. A.; Sahasrabudhe, P. V.; Hoth, L. R.; et al. Structural Basis for AMPK Activation: Natural and Synthetic Ligands Regulate Kinase Activity from Opposite Poles by Different Molecular Mechanisms. *Structure* **2014**, 22 (8), 1161–1172.
<https://doi.org/10.1016/j.str.2014.06.009>.
 - (27) Yan, Y.; Zhou, X. E.; Xu, H. E.; Melcher, K. Structure and Physiological Regulation of AMPK. *Int. J. Mol. Sci.* **2018**, 19 (11). <https://doi.org/10.3390/ijms19113534>.
 - (28) Garcia, D.; Shaw, R. J. AMPK: Mechanisms of Cellular Energy Sensing and Restoration of Metabolic Balance. *Mol. Cell* **2017**, 66 (6), 789–800.
<https://doi.org/10.1016/j.molcel.2017.05.032>.
 - (29) Carling, D.; Thornton, C.; Woods, A.; Sanders, M. J. AMP-Activated Protein Kinase: New Regulation, New Roles? *Biochem. J.* **2012**, 445 (1), 11–27. <https://doi.org/10.1042/BJ20120546>.
 - (30) Chen, L.; Wang, J.; Zhang, Y. Y.; Yan, S. F.; Neumann, D.; Schlattner, U.; Wang, Z. X.; Wu, J. W. AMP-Activated Protein Kinase Undergoes Nucleotide-Dependent Conformational Changes. *Nat. Struct. Mol. Biol.* **2012**, 19 (7), 716–718. <https://doi.org/10.1038/nsmb.2319>.
 - (31) Wu, L.; Zhang, L.; Li, B.; Jiang, H.; Duan, Y.; Xie, Z.; Shuai, L.; Li, J.; Li, J. AMP-Activated Protein Kinase (AMPK) Regulates Energy Metabolism through Modulating Thermogenesis in Adipose Tissue. *Front. Physiol.* **2018**, 9 (FEB), 1–23. <https://doi.org/10.3389/fphys.2018.00122>.
 - (32) King, M. W. AMPK: Master Metabolic Regulator
<https://themedicalbiochemistrypage.org/ampk.php> (accessed Apr 26, 2020).
 - (33) Gowans, G. J.; Hawley, S. A.; Ross, F. A.; Hardie, D. G. AMP Is a True Physiological Regulator of Amp-Activated Protein Kinase by Both Allosteric Activation and Enhancing Net Phosphorylation. *Cell Metab.* **2013**, 18 (4), 556–566. <https://doi.org/10.1016/j.cmet.2013.08.019>.
 - (34) Carling, D.; Mayer, F. V.; Sanders, M. J.; Gamblin, S. J. AMP-Activated Protein Kinase: Nature’s Energy Sensor. *Nat. Chem. Biol.* **2011**, 7 (8), 512–518.
<https://doi.org/10.1038/nchembio.610>.
 - (35) Vazquez-Martin, A.; Vellon, L.; Quirós, P. M.; Cufí, S.; De Galarreta, E. R.; Ferraros, C. O.; Martin, A. G.; Martin-Castillo, B.; Lopez-Otin, C.; Menendez, J. A. Activation of AMP-Activated Protein Kinase (AMPK) Provides a Metabolic Barrier to Reprogramming Somatic Cells into Stem Cells. *Cell Cycle* **2012**, 11 (5), 974–989. <https://doi.org/10.4161/cc.11.5.19450>.
 - (36) Cool, B.; Zinker, B.; Chiou, W.; Kifle, L.; Cao, N.; Perham, M.; Dickinson, R.; Adler, A.; Gagne, G.; Iyengar, R.; et al. Identification and Characterization of a Small Molecule AMPK Activator That Treats Key Components of Type 2 Diabetes and the Metabolic Syndrome. *Cell Metab.* **2006**, 3 (6), 403–416. <https://doi.org/10.1016/j.cmet.2006.05.005>.
 - (37) Hardie, D. G. AMPK - Sensing Energy While Talking to Other Signaling Pathways. *Cell Metab.*

2014, 20 (6), 939–952. <https://doi.org/10.1016/j.cmet.2014.09.013>.

- (38) Willows, R.; Sanders, M. J.; Xiao, B.; Patel, B. R.; Martin, S. R.; Read, J.; Wilson, J. R.; Hubbard, J.; Gamblin, S. J.; Carling, D. Phosphorylation of AMPK by Upstream Kinases Is Required for Activity in Mammalian Cells. *Biochem. J.* **2017**, *474* (17), 3059–3073. <https://doi.org/10.1042/BCJ20170458>.
- (39) Scott, J. W.; Ling, N.; Issa, S. M. A.; Dite, T. A.; O'Brien, M. T.; Chen, Z. P.; Galic, S.; Langendorf, C. G.; Steinberg, G. R.; Kemp, B. E.; et al. Small Molecule Drug A-769662 and AMP Synergistically Activate Naive AMPK Independent of Upstream Kinase Signaling. *Chem. Biol.* **2014**, *21* (5), 619–627. <https://doi.org/10.1016/j.chembiol.2014.03.006>.
- (40) Xiao, B.; Sanders, M. J.; Carmena, D.; Bright, N. J.; Haire, L. F.; Underwood, E.; Patel, B. R.; Heath, R. B.; Walker, P. A.; Hallen, S.; et al. Structural Basis of AMPK Regulation by Small Molecule Activators. *Nat. Commun.* **2013**, *4*, 1–10. <https://doi.org/10.1038/ncomms4017>.
- (41) Ngoei, K. R. W.; Langendorf, C. G.; Ling, N. X. Y.; Hoque, A.; Varghese, S.; Camerino, M. A.; Walker, S. R.; Bozikis, Y. E.; Dite, T. A.; Ovens, A. J.; et al. Structural Determinants for Small-Molecule Activation of Skeletal Muscle AMPK A2 β 2 γ 1 by the Glucose Importogog SC4. *Cell Chem. Biol.* **2018**, *25* (6), 728–737.e9. <https://doi.org/10.1016/j.chembiol.2018.03.008>.
- (42) Jørgensen, N. O.; Kjøbsted, R.; Larsen, M. R.; Birk, J. B.; Andersen, N. R.; Albuquerque, B.; Schjerling, P.; Miller, R.; Carling, D.; Pehmøller, C. K.; et al. Direct Small Molecule ADaM-Site AMPK Activators Reveal an AMPK γ 3-Independent Mechanism for Blood Glucose Lowering. *Mol. Metab.* **2021**, No. June, 101259. <https://doi.org/10.1016/j.molmet.2021.101259>.
- (43) Cokorinos, E. C.; Delmore, J.; Reyes, A. R.; Albuquerque, B.; Kjøbsted, R.; Jørgensen, N. O.; Tran, J. L.; Jatkar, A.; Cialdea, K.; Esquejo, R. M.; et al. Activation of Skeletal Muscle AMPK Promotes Glucose Disposal and Glucose Lowering in Non-Human Primates and Mice. *Cell Metab.* **2017**, *25* (5), 1147–1159.e10. <https://doi.org/10.1016/j.cmet.2017.04.010>.
- (44) Santa Gómez, M.; Aledavood, E.; Beroiz, M.; Lagartera, L.; Vega-Martín, E.; Gil-Ortega, M.; Cumella, J.; Pérez, C.; Luque, F. J.; Estarellas, C.; et al. Novel Indolic AMPK Modulators Increase Nitric Oxide Release in Human Endothelial Cells. **2021**.
- (45) World Health Organization. Global tuberculosis report 2019 https://www.who.int/tb/publications/global_report/en/ (accessed Mar 15, 2020).
- (46) Cadena, A. M.; Fortune, S. M.; Flynn, J. L. Heterogeneity in Tuberculosis. *Nat. Rev. Immunol.* **2017**, *17* (11), 691–702. <https://doi.org/10.1038/nri.2017.69>.
- (47) Schön, T.; Miotto, P.; Köser, C. U.; Viveiros, M.; Böttger, E.; Cambau, E. Mycobacterium Tuberculosis Drug-Resistance Testing: Challenges, Recent Developments and Perspectives. *Clin. Microbiol. Infect.* **2017**, *23* (3), 154–160. <https://doi.org/10.1016/j.cmi.2016.10.022>.
- (48) Singh, V.; Mizrahi, V. Identification and Validation of Novel Drug Targets in Mycobacterium Tuberculosis. *Drug Discov. Today* **2017**, *22* (3), 503–509. <https://doi.org/10.1016/j.drudis.2016.09.010>.
- (49) Caminero, J. A.; Sotgiu, G.; Zumla, A.; Migliori, G. B. Best Drug Treatment for Multidrug-Resistant and Extensively Drug-Resistant Tuberculosis. *Lancet Infect. Dis.* **2010**, *10* (9), 621–629. [https://doi.org/10.1016/S1473-3099\(10\)70139-0](https://doi.org/10.1016/S1473-3099(10)70139-0).
- (50) Oliveira, A.; Singh, S.; Bidon-Chanal, A.; Forti, F.; Martí, M. A.; Boechi, L.; Estrin, D. A.; Dikshit, K. L.; Luque, F. J. Role of PheE15 Gate in Ligand Entry and Nitric Oxide Detoxification Function of Mycobacterium Tuberculosis Truncated Hemoglobin N. *PLoS One* **2012**, *7* (11). <https://doi.org/10.1371/journal.pone.0049291>.
- (51) Bidon-Chanal, A.; Martí, M. A.; Crespo, A.; Milani, M.; Orozco, M.; Bolognesi, M.; Luque, F. J.; Estrin, D. A. Ligand-Induced Dynamical Regulation of NO Conversion in Mycobacterium Tuberculosis Truncated Hemoglobin-N. *Proteins Struct. Funct. Genet.* **2006**, *64* (2), 457–464. <https://doi.org/10.1002/prot.21004>.
- (52) Pawaria, S.; Lama, A.; Raje, M.; Dikshit, K. L. Responses of Mycobacterium Tuberculosis Hemoglobin Promoters to In Vitro and In Vivo Growth Conditions □. **2008**, *74* (11), 3512–3522. <https://doi.org/10.1128/AEM.02663-07>.
- (53) Heroux, M. S.; Mohan, A. D.; Olsen, K. W. Ligand Migration in the Truncated Hemoglobin of

- Mycobacterium Tuberculosis. **2011**, *63* (March), 214–220. <https://doi.org/10.1002/iub.438>.
- (54) Pesce, A.; Couture, M.; Dewilde, S.; Guertin, M.; Yamauchi, K.; Ascenzi, P.; Moens, L.; Bolognesi, M. A Novel Two-over-Two α -Helical Sandwich Fold Is Characteristic of the Truncated Hemoglobin Family. **2000**, *19* (11).
- (55) Perutz, M. F. AFFINITY OF HEMOGLOBIN : Influence of Structure of the Globin on the Heme Iron. **1979**.
- (56) Milani, M.; Pesce, A.; Ouellet, Y.; Ascenzi, P.; Guertin, M.; Bolognesi, M. Mycobacterium Tuberculosis Hemoglobin N Displays a Protein Tunnel Suited for O₂ Diffusion to the Heme. **2001**, *20* (15).
- (57) Lama, A.; Pawaria, S.; Bidon-chanal, A.; Anand, A.; Gelpí, L.; Arya, S.; Martí, M.; Estrin, D. A.; Luque, F. J.; Dikshit, K. L. Role of Pre-A Motif in Nitric Oxide Scavenging by Truncated Hemoglobin, HbN, of Mycobacterium Tuberculosis * □. **2009**, *284* (21), 14457–14468. <https://doi.org/10.1074/jbc.M807436200>.
- (58) Milani, M.; Pesce, A.; Nardini, M.; Ouellet, H.; Ouellet, Y.; Dewilde, S.; Bocedi, A.; Ascenzi, P.; Guertin, M.; Moens, L.; et al. Structural Bases for Heme Binding and Diatomic Ligand Recognition in Truncated Hemoglobins. **2005**, *99*, 97–109. <https://doi.org/10.1016/j.jinorgbio.2004.10.035>.
- (59) Ascenzi, P.; Bolognesi, M.; Milani, M.; Guertin, M.; Visca, P. Mycobacterial Truncated Hemoglobins : From Genes to Functions. **2007**, *398*, 42–51. <https://doi.org/10.1016/j.gene.2007.02.043>.
- (60) Mukai, M.; Ouellet, Y.; Ouellet, H.; Guertin, M.; Yeh, S. R. NO Binding Induced Conformational Changes in a Truncated Hemoglobin from Mycobacterium Tuberculosis. *Biochemistry* **2004**, *43* (10), 2764–2770. <https://doi.org/10.1021/bi035798o>.
- (61) Martí, M. A.; Bidon-Chanal, A.; Crespo, A.; Yeh, S. R.; Guallar, V.; Luque, F. J.; Estrin, D. A. Mechanism of Product Release in NO Detoxification from Mycobacterium Tuberculosis Truncated Hemoglobin N. *J. Am. Chem. Soc.* **2008**, *130* (5), 1688–1693. <https://doi.org/10.1021/ja076853+>.
- (62) Singh, S.; Thakur, N.; Oliveira, A.; Petruk, A. A.; Hade, M. D.; Sethi, D.; Bidon-chanal, A.; Martí, M. A.; Datta, H.; Parkesh, R.; et al. Mechanistic Insight into the Enzymatic Reduction of Truncated Hemoglobin N of Mycobacterium Tuberculosis. **2014**, *289* (31), 21573–21583. <https://doi.org/10.1074/jbc.M114.578187>.
- (63) Smagghe, B. J.; Trent, J. T.; Hargrove, M. S. NO Dioxygenase Activity in Hemoglobins Is Ubiquitous in Vitro, but Limited by Reduction in Vivo. *PLoS One* **2008**, *3* (4). <https://doi.org/10.1371/journal.pone.0002039>.
- (64) Yano, T.; Lin-Sheng, L.; Weinstein, E.; Teh, J. S.; Rubin, H. Steady-State Kinetics and Inhibitory Action of Antitubercular Phenothiazines on Mycobacterium Tuberculosis Type-II NADH-Menaquinone Oxidoreductase (NDH-2). *J. Biol. Chem.* **2006**, *281* (17), 11456–11463. <https://doi.org/10.1074/jbc.M508844200>.
- (65) Gardner, P. R.; Gardner, A. M.; Martin, L. A.; Dou, Y.; Li, T.; Olson, J. S.; Zhu, H.; Riggs, A. F. Nitric-Oxide Dioxygenase Activity and Function of Flavohemoglobins. Sensitivity to Nitric Oxide and Carbon Monoxide Inhibition. *J. Biol. Chem.* **2000**, *275* (41), 31581–31587. <https://doi.org/10.1074/jbc.M004141200>.
- (66) Eich, R. F.; Li, T.; Lemon, D. D.; Doherty, D. H.; Curry, S. R.; Aitken, J. F.; Mathews, A. J.; Johnson, K. A.; Smith, R. D.; Phillips, G. N.; et al. Mechanism of NO-Induced Oxidation of Myoglobin and Hemoglobin. *Biochemistry* **1996**, *35* (22), 6976–6983. <https://doi.org/10.1021/bi960442g>.
- (67) Fago, A.; Mathews, A. J.; Moens, L.; Dewilde, S.; Brittain, T. The Reaction of Neuroglobin with Potential Redox Protein Partners Cytochrome B5 and Cytochrome C. *FEBS Lett.* **2006**, *580* (20), 4884–4888. <https://doi.org/10.1016/j.febslet.2006.08.003>.
- (68) Gardner, P. R. Nitric Oxide Dioxygenase Function and Mechanism of Flavohemoglobin, Hemoglobin, Myoglobin and Their Associated Reductases. *J. Inorg. Biochem.* **2005**, *99* (1), 247–266. <https://doi.org/10.1016/j.jinorgbio.2004.10.003>.

- (69) Ouellet, H.; Ouellet, Y.; Richard, C.; Labarre, M.; Wittenberg, B.; Wittenberg, J.; Guertin, M. Truncated Hemoglobin HbN Protects Mycobacterium Bovis from Nitric Oxide. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (9), 5902–5907. <https://doi.org/10.1073/pnas.092017799>.
- (70) Depeursinge, A.; Racoceanu, D.; Iavindrasana, J.; Cohen, G.; Platon, A.; Poletti, P.-A.; Muller, H. Fusing Visual and Clinical Information for Lung Tissue Classification in HRCT Data. *Artif. Intell. Med.* **2010**, *10*, ARTMED1118. <https://doi.org/10.1016/j>.
- (71) Sperger, T.; Sanhueza, I. A.; Schoenebeck, F. Computation and Experiment: A Powerful Combination to Understand and Predict Reactivities. *Acc. Chem. Res.* **2016**, *49* (6), 1311–1319. <https://doi.org/10.1021/acs.accounts.6b00068>.
- (72) Smith, S. j.; Sutcliffe, B. T. The Development of Computational Chemistry in the United Kingdom. *Rev. Comput. Chem.* **1997**, *10*, 271–316.
- (73) Computer-Aided Drug Discovery <https://www.solutions.bocsci.com/computational-chemistry.htm> (accessed Apr 2, 2021).
- (74) Computational Chemistry <https://www.acs.org/content/acs/en/careers/college-to-career/chemistry-careers/computational-chemistry.html> (accessed Mar 17, 2020).
- (75) Cohen, N. C. Guide Book on Molecular Modelling on Drug Design. In *Guide Book on Molecular Modelling on Drug Design*; Academic press limited publication, London, 1996.
- (76) Young, D. C. No Title. In *Computational Chemistry: A Practical Guide for Applying Techniques to Real-World Problems*; John Wiley & Sons Inc., 2001.
- (77) Friesner, R. A. Ab Initio Quantum Chemistry: Methodology and Applications. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6648–6653. <https://doi.org/10.1073/pnas.0408036102>.
- (78) Termath, V. Quantum Mechanics in Chemistry. *Zeitschrift für Phys. Chemie* **1998**, *205* (Part_1), 135–135. https://doi.org/10.1524/zpch.1998.205.part_1.135.
- (79) Reviews in Computational Chemistry - Google Books [https://books.google.es/books?hl=en&lr=&id=YhceVOq1K5UC&oi=fnd&pg=PA45&dq=semi+empirical+methods&ots=_K7YxJpzWS&sig=3Dq9SoJikRMg1uLoaa_gwdvu88s&redir_esc=y#v=onepage&q=semi empirical methods&f=false](https://books.google.es/books?hl=en&lr=&id=YhceVOq1K5UC&oi=fnd&pg=PA45&dq=semi+empirical+methods&ots=_K7YxJpzWS&sig=3Dq9SoJikRMg1uLoaa_gwdvu88s&redir_esc=y#v=onepage&q=semi%20empirical%20methods&f=false) (accessed Mar 19, 2020).
- (80) Hehre, W. J. *A Guide to Molecular Mechanics and Quantum Chemical Calculations*; 2003.
- (81) Landau, R. H.; Páez, M.; Gould, H.; Tobochnik, J. *Computational Physics*; 1999; Vol. 67. <https://doi.org/10.1119/1.19197>.
- (82) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd Editio.; 2004.
- (83) Leach, A. R. *Molecular Modelling: Principles and Applications*, Second edi.; 2001.
- (84) Darian, E.; Gannett, P. M. Application of Molecular Dynamics Simulations to Spin-Labeled Oligonucleotides. *J. Biomol. Struct. Dyn.* **2005**, *22* (5), 579–593. <https://doi.org/10.1080/07391102.2005.10507028>.
- (85) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins Struct. Funct. Genet.* **2006**, *65* (3), 712–725. <https://doi.org/10.1002/prot.21123>.
- (86) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91* (1–3), 43–56. [https://doi.org/10.1016/0010-4655\(95\)00042-E](https://doi.org/10.1016/0010-4655(95)00042-E).
- (87) Theoretical Biophysics Group. NAMD User Guide. *Univ. Illinois, Urbana, USA* **2016**.
- (88) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; et al. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLoS Comput. Biol.* **2017**, *13* (7), 1–17. <https://doi.org/10.1371/journal.pcbi.1005659>.
- (89) Bowers, K. J.; Chow, D. E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In *SC'06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*; IEEE, 2006; p 43.

- (90) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174. <https://doi.org/10.1002/jcc.20035>.
- (91) Martín-García, F.; Papaleo, E.; Gomez-Puertas, P.; Boomsma, W.; Lindorff-Larsen, K. Comparing Molecular Dynamics Force Fields in the Essential Subspace. *PLoS One* **2015**, *10* (3), 1–16. <https://doi.org/10.1371/journal.pone.0121114>.
- (92) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280. <https://doi.org/10.1021/j100142a004>.
- (93) Lapinski, L. VIMAIIOlq ~ SPE ~ PY Comparison of Ab Initio HF / 6-31G **, HF / 6-31 + + G ** and MP2 / 6-31G ** Calculated Infrared Spectra with Matrix Isolation Spectra. **1995**, *8*, 331–342.
- (94) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21* (12), 1049. [https://doi.org/10.1002/1096-987x\(200009\)21:12<1049::aid-jcc3>3.3.co;2-6](https://doi.org/10.1002/1096-987x(200009)21:12<1049::aid-jcc3>3.3.co;2-6).
- (95) Kukol, A. Molecular Modeling of Proteins. **2008**, No. February 2008, 390. <https://doi.org/10.1007/978-1-59745-177-2>.
- (96) Protein Data Bank.
- (97) Yang, J.; Zhang, Y.; Rohini, K.; Srikumar, P. S.; Anbarasu, K.; Structural, L.; Suite, B.; Ec, A.; Security, D. C.; Roy, A.; et al. Function Prediction. *Curr. Protoc. Bioinforma.* **2011**, *9* (December), 5–9. <https://doi.org/10.1038/nprot.2010.5.I-TASSER>.
- (98) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. SWISS-MODEL: An Automated Protein Homology-Modeling Server. *Nucleic Acids Res.* **2003**, *31* (13), 3381–3385. <https://doi.org/10.1093/nar/gkg520>.
- (99) Kozakov, D.; Hall, D. R.; Xia, B.; Porter, K. A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S. The ClusPro Web Server for Protein-Protein Docking. *Nat. Protoc.* **2017**, *12* (2), 255–278. <https://doi.org/10.1038/nprot.2016.169>.
- (100) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential Dynamics of Proteins. *Proteins Struct. Funct. Bioinforma.* **1993**, *17*, 412–425.
- (101) Shkurti, A.; Goni, R.; Andrio, P.; Breitmoser, E.; Bethune, I.; Orozco, M.; Laughton, C. A. PyPcazip: A PCA-Based Toolkit for Compression and Analysis of Molecular Simulation Data. *SoftwareX* **2015**, *5*, 44–50. <https://doi.org/10.1016/j.softx.2016.04.002>.
- (102) Karplus, M.; Ichiye, T. Comment on a “Fluctuation and Cross-Correlation Analysis of Protein Motions Observed in Nanosecond Molecular Dynamics Simulations.” *J. Mol. Biol.* **1995**, *252* (4), 492–503. <https://doi.org/10.1006/jmbi.1995.0514>.
- (103) O’Rourke, K. F.; Gorman, S. D.; Boehr, D. D. Biophysical and Computational Methods to Analyze Amino Acid Interaction Networks in Proteins. *Comput. Struct. Biotechnol. J.* **2016**, *14*, 245–251. <https://doi.org/10.1016/j.csbj.2016.06.002>.
- (104) Van Wart, A. T.; Durrant, J.; Votapka, L.; Amaro, R. E. Weighted Implementation of Suboptimal Paths (WISP): An Optimized Algorithm and Tool for Dynamical Network Analysis. *J. Chem. Theory Comput.* **2014**, *10* (2), 511–517. <https://doi.org/10.1021/ct4008603>.
- (105) MUNDAY, M. R.; CAMPBELL, D. G.; CARLING, D.; HARDIE, D. G. Identification by Amino Acid Sequencing of Three Major Regulatory Phosphorylation Sites on Rat Acetyl-CoA Carboxylase. *Eur. J. Biochem.* **1988**, *175* (2), 331–338. <https://doi.org/10.1111/j.1432-1033.1988.tb14201.x>.
- (106) Witters, L. A.; Gao, G.; Kemp, B. E.; Quistorff, B. Hepatic 5'-AMP-Activated Protein Kinase: Zonal Distribution and Relationship to Acetyl-CoA Carboxylase Activity in Varying Nutritional States. *Arch. Biochem. Biophys.* **1994**, *308* (2), 413–419.
- (107) Winder, W. W.; Hardie, D. G. Inactivation of Acetyl-CoA Carboxylase and Activation of AMP-Activated Protein Kinase in Muscle during Exercise. *Am. J. Physiol. - Endocrinol. Metab.* **1996**, *270*, 299–304.

- (108) Sanders, M. J.; Grondin, P. O.; Hegarty, B. D.; Snowden, M. A.; Carling, D. Investigating the Mechanism for AMP Activation of the AMP-Activated Protein Kinase Cascade. *Biochem. J.* **2007**, *403* (1), 139–148. <https://doi.org/10.1042/BJ20061520>.
- (109) Giordanetto, F.; Karis, D. Direct AMP-Activated Protein Kinase Activators: A Review of Evidence from the Patent Literature. *Expert Opin. Ther. Pat.* **2012**, *22* (12), 1467–1477. <https://doi.org/10.1517/13543776.2012.743994>.
- (110) Scott, J. W.; Galic, S.; Graham, K. L.; Foitzik, R.; Ling, N. X. Y.; Dite, T. A.; Issa, S. M. A.; Langendorf, C. G.; Weng, Q. P.; Thomas, H. E.; et al. Inhibition of AMP-Activated Protein Kinase at the Allosteric Drug-Binding Site Promotes Islet Insulin Release. *Chem. Biol.* **2015**, *22* (6), 705–711. <https://doi.org/10.1016/j.chembiol.2015.05.011>.
- (111) Gheeraert, A.; Pacini, L.; Batista, V. S.; Vuillon, L.; Lesieur, C.; Rivalta, I. Exploring Allosteric Pathways of a V-Type Enzyme with Dynamical Perturbation Networks. *J. Phys. Chem. B* **2019**, *123* (16), 3452–3461. <https://doi.org/10.1021/acs.jpcc.9b01294>.
- (112) Forrester, M. T.; Foster, M. W. Protection from Nitrosative Stress: A Central Role for Microbial Flavohemoglobin. *Free Radic. Biol. Med.* **2012**, *52* (9), 1620–1633. <https://doi.org/10.1016/j.freeradbiomed.2012.01.028>.
- (113) Pathania, R.; Navani, N. K.; Gardner, A. M.; Gardner, P. R.; Dikshit, K. L. Nitric Oxide Scavenging and Detoxification by the Mycobacterium Tuberculosis Haemoglobin, HbN in Escherichia Coli. *Mol. Microbiol.* **2002**, *45* (5), 1303–1314. <https://doi.org/10.1046/j.1365-2958.2002.03095.x>.
- (114) Yeh, S. R.; Couture, M.; Ouellet, Y.; Guertin, M.; Rousseau, D. L. A Cooperative Oxygen Binding Hemoglobin from Mycobacterium Tuberculosis. Stabilization of Heme Ligands by a Distal Tyrosine Residue. *J. Biol. Chem.* **2000**, *275* (3), 1679–1684. <https://doi.org/10.1074/jbc.275.3.1679>.
- (115) Milani, M.; Pesce, A.; Ouellet, Y.; Ascenzi, P.; Guertin, M.; Bolognesi, M. Mycobacterium Tuberculosis Hemoglobin N Displays a Protein Tunnel Suited for O₂ Diffusion to the Heme. *EMBO J.* **2001**, *20* (15), 3902–3909. <https://doi.org/10.1093/emboj/20.15.3902>.
- (116) No Title <https://mycobrowser.epfl.ch> (accessed Nov 20, 2020).
- (117) Moul, J.; Fidelis, K.; Kryshchak, A.; Schwede, T.; Tramontano, A. Critical Assessment of Methods of Protein Structure Prediction: Progress and New Directions in Round XI. *Proteins Struct. Funct. Bioinforma.* **2016**, *84* (April), 4–14. <https://doi.org/10.1002/prot.25064>.
- (118) Weinstein, E. A.; Yano, T.; Li, L. S.; Avarbock, D.; Avarbock, A.; Helm, D.; McColm, A. A.; Duncan, K.; Lonsdale, J. T.; Rubin, H. Inhibitors of Type II NADH:Menaquinone Oxidoreductase Represent a Class of Antitubercular Drugs. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (12), 4548–4553. <https://doi.org/10.1073/pnas.0500469102>.
- (119) Boshoff, H. I. M.; Myers, T. G.; Copp, B. R.; McNeil, M. R.; Wilson, M. A.; Barry, C. E. The Transcriptional Responses of Mycobacterium Tuberculosis to Inhibitors of Metabolism. Novel Insights into Drug Mechanisms of Action. *J. Biol. Chem.* **2004**, *279* (38), 40174–40184. <https://doi.org/10.1074/jbc.M406796200>.
- (120) Sasseti, C. M.; Rubin, E. J. Genetic Requirements for Mycobacterial Survival during Infection. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (22), 12989–12994. <https://doi.org/10.1073/pnas.2134250100>.
- (121) McAdam, R. A.; Quan, S.; Smith, D. A.; Bardarov, S.; Betts, J. C.; Cook, F. C.; Hooker, E. U.; Lewis, A. O.; Woollard, P.; Everett, M. J.; et al. Characterization of a Mycobacterium Tuberculosis H37Rv Transposon Library Reveals Insertions in 351 ORFs and Mutants with Altered Virulence. *Microbiology* **2002**, *148* (10), 2975–2986. <https://doi.org/10.1099/00221287-148-10-2975>.
- (122) Yano, T.; Lin-Sheng, L.; Weinstein, E.; Teh, J. S.; Rubin, H. Steady-State Kinetics and Inhibitory Action of Antitubercular Phenothiazines on Mycobacterium Tuberculosis Type-II NADH-Menaquinone Oxidoreductase (NDH-2). *J. Biol. Chem.* **2006**, *281* (17), 11456–11463. <https://doi.org/10.1074/jbc.M508844200>.
- (123) Shirude, P. S.; Paul, B.; Roy Choudhury, N.; Kedari, C.; Bandodkar, B.; Ugarkar, B. G. Quinolonyl Pyrimidines: Potent Inhibitors of NDH-2 as a Novel Class of Anti-TB Agents. *ACS*

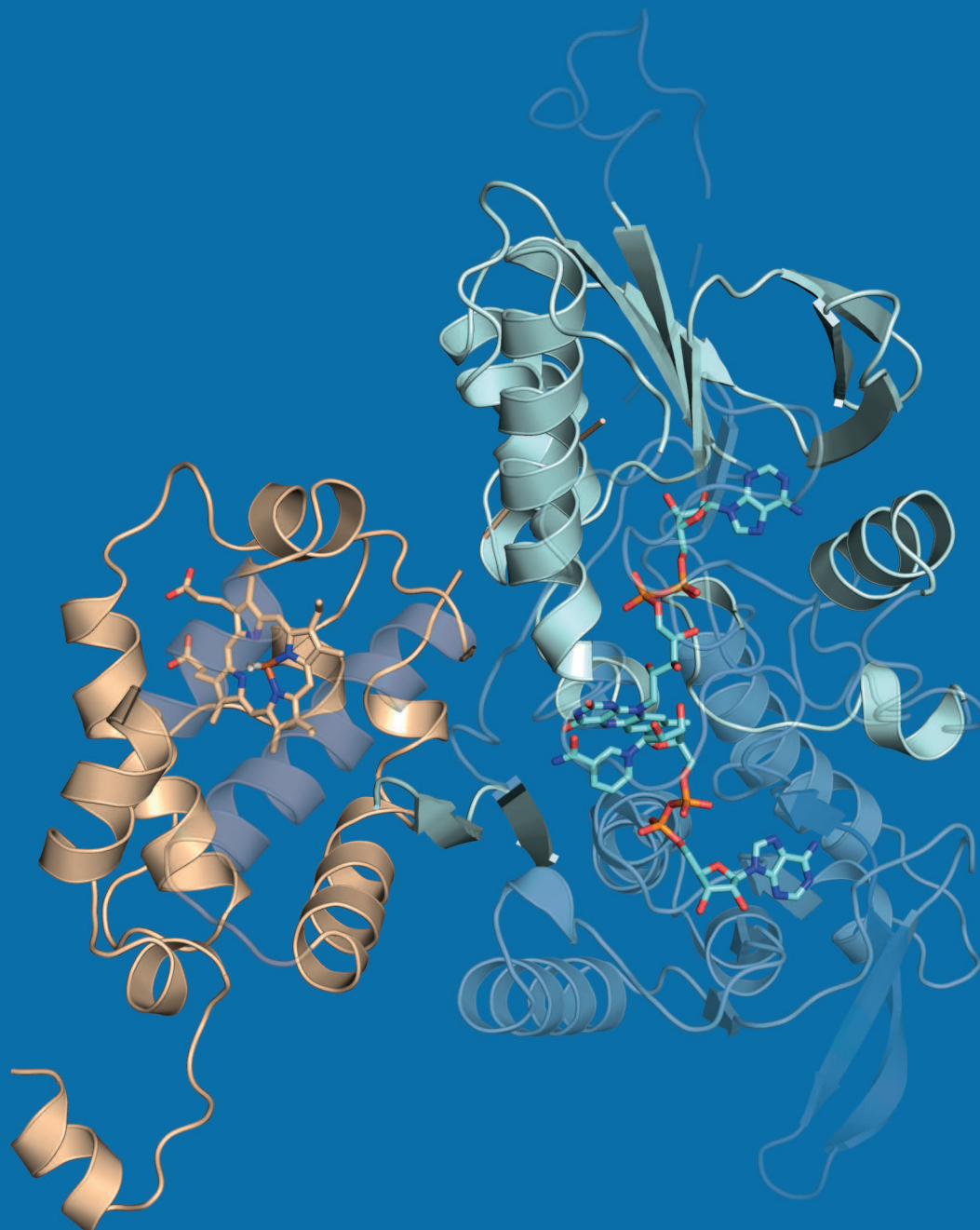
Med. Chem. Lett. **2012**, 3 (9), 736–740. <https://doi.org/10.1021/ml300134b>.

- (124) Case, D. A.; Betz, R. .; Botello-Smith, W.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; TGiese, T. J.; Gohlke, H.; Goetz, A. W.; et al. Amber 2016. *Univ. California, San Fr.* **2016**.
- (125) Beratan, D.; Betts, J.; Onuchic, J. Protein Electron Transfer Rates Set by the Bridging Secondary and Tertiary Structure. *Science (80-.)*. **1991**, 252, 1285–1288.
- (126) Beratan, D.; Onuchic, J.; Winkler, J.; Gray, H. Electron-Tunneling Pathways in Proteins. *Science (80-.)*. **1992**, 258 (5089), 1740–1741.
- (127) Torrents, E.; Sahlin, M.; Biglino, D.; Gräslund, A.; Sjöberg, B. M. Efficient Growth Inhibition of Bacillus Anthracis by Knocking out the Ribonucleotide Reductase Tyrosyl Radical. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102 (50), 17946–17951. <https://doi.org/10.1073/pnas.0506410102>.
- (128) Hade, M. D.; Sethi, D.; Datta, H.; Singh, S.; Thakur, N.; Chhaya, A.; Dikshit, K. L. Truncated Hemoglobin O Carries an Autokinase Activity and Facilitates Adaptation of Mycobacterium Tuberculosis under Hypoxia. *Antioxidants Redox Signal.* **2020**, 32 (6), 351–362. <https://doi.org/10.1089/ars.2018.7708>.
- (129) Arya, S.; Sethi, D.; Singh, S.; Hade, M. D.; Singh, V.; Raju, P.; Chodiseti, S. B.; Verma, D.; Varshney, G. C.; Agrewala, J. N.; et al. Truncated Hemoglobin, HbN, Is Post-Translationally Modified in Mycobacterium Tuberculosis and Modulates Host-Pathogen Interactions during Intracellular Infection. *J. Biol. Chem.* **2013**, 288 (41), 29987–29999. <https://doi.org/10.1074/jbc.M113.507301>.
- (130) Harbut, M. B.; Yang, B.; Liu, R.; Yano, T.; Vilchèze, C.; Cheng, B.; Lockner, J.; Guo, H.; Yu, C.; Franzblau, S. G.; et al. Small Molecules Targeting Mycobacterium Tuberculosis Type II NADH Dehydrogenase Exhibit Antimycobacterial Activity. *Angew. Chemie - Int. Ed.* **2018**, 57 (13), 3478–3482. <https://doi.org/10.1002/anie.201800260>.



UNIVERSITAT DE
BARCELONA

Facultat de Farmàcia i Ciències de l'Alimentació
Universitat de Barcelona



Elnaz Aledavood
2021