UNIVERSITY OF NAPLES "FEDERICO II"

DEPARTMENT OF PHARMACY

School of Medicine and Surgery

PhD in "Pharmaceutical Science"

XXXIV CYCLE

"Computationally-driven design of bile acids multi-targeting modulators for metabolic and inflammatory diseases"

PhD Tutor **Candidate**

Dr. Bruno Catalanotti **Dr. Bianca Fiorillo**

PhD Coordinator Prof. Maria Valeria D'Auria

“Remember to look up at the stars and not down at your feet.
Try to make sense of what you see
and wonder about what makes the Universe exist.
Be curious.
And however difficult life may seem,
there is always something you can do and succeed at.
It matters that you don’t just give up.”
Stephen Hawking

Alla mia famiglia, il mio tutto.

Index

Summary 1

Publications 4

List of abbreviations 7

Chapter I: Introduction 9

1.1 Bile acids (BAs) 10

1.2 Bile acids receptors: Nuclear Receptors (NRs) 13

1.3 Bile acids receptors: G-protein coupled receptors (GPCRs) 17

1.4 Other targets for bile acids receptors' (BARs) ligands 19

1.4.1 *Cysteinyl Leukotriene receptor 1 ($CysLT_1R$)* 19

1.4.2 SARS-CoV-2 Spike protein 22

1.4.3 Angiotensin converting enzyme 2 (ACE2) 25

1.5 Theoretical Methods 28

1.5.1 Molecular docking 28

1.5.2 Virtual screening (VS) 29

1.5.3 Molecular dynamics (MD) 29

1.5.4 Metadynamics (MetaD) 32

1.5.5 Bias exchange (BE) 34

Chapter II: Results and Discussion 35

2.1 Design of dual modulators against the nuclear retinoic acid-related orphan receptor γ (ROR-γ) and the membrane receptor G-protein bile acid receptor 1 (GPBAR1) 36

2.1.1 Pharmacological activity 37

2.1.2 Computational studies 38

2.2 Dual modulators against Cysteinyl Leukotriene Receptor 1($CysLT_1R$) and G-Protein coupled bile acid receptor 1 (GPBAR1) 47

2.2.1 Pharmacological activity 47

2.2.2 Computational studies 53

2.3 Natural and semi-synthetic steroidal agent acting on Spike protein Receptor Binding Domain (RBD) of SARS-CoV-2 65

2.3.1 Computational studies 65

2.3.2 In vitro binding assays 76

2.4 Bile acids derivatives as potent angiotensin converting enzyme 2 (ACE2) activators 83

2.4.1 Computational studies 83

2.4.2 In vitro enzymatic assays 100

CHAPTER III: Side projects 105

3.1 The TBC1D31/praja2 complex controls primary ciliogenesis through PKA-directed OFD1 ubiquitylation 107

3.1.1 Introduction 107

3.1.2 Results and Discussion 110

3.2 Discovery of pelargonidin as a potential inhibitor of the SARS-CoV-2 interaction and angiotensin converting enzyme 2 (ACE2) 122

3.2.1 Introduction 122

3.2.2 Results and Discussion 123

Conclusions 138

Computational Methods 142

4.1 Design of dual modulators against the nuclear retinoic acid-related orphan receptor γ (ROR-γ) and the membrane receptor G-protein bile acid receptor 1 (GPBAR1) 143

4.2 Dual modulators against Cysteinyl Leukotriene Receptor 1($CysLT_1R$) and G-Protein coupled bile acid receptor 1 (GPBAR1) 147

4.3 Natural and semi-synthetic steroidal agent acting on Spike protein Receptor Binding Domain (RBD) of SARS-CoV-2 151

4.4 Bile acids derivatives as potent angiotensin converting enzyme 2 (ACE2) activators 154

4.5 The TBC1D31/praja2 complex controls primary ciliogenesis through PKA-directed OFD1 ubiquitylation 158

4.6 *Discovery of pelargonidin as a potential inhibitor of the SARS-CoV-2 interaction and angiotensin converting enzyme 2 (ACE2)* 161

References 163

Summary

During the PhD period, my research activity was focused on the applications of both classical and advanced computational techniques to the medicinal chemistry field.

I had the opportunity to participate in various drug discovery campaigns, dealing with different kinds of biological targets.

This thesis highlights the major research activities in which I was involved, with a particular focus on the pharmaceutical design and optimization of new compounds targeting bile acids (BA) receptors, including both Nuclear (NRs) and G-Protein Coupled (GPCRs) receptors. The research activity of the doctorate began by a deep comprehension of state of the art regarding the dual activity of bile acid derivatives targeting the steroidal receptors and, in particular, with activity towards the nuclear retinoic acid-related orphan receptor γ (ROR-γ) and the membrane receptor G-protein bile acid receptor 1 (GPBAR1) for the treatment of metabolic diseases. Subsequently, the mechanism of action and the pharmacological activity of bile acid derivatives were explored on other different targets.

Therefore, my research activity was also extended to the identification of ligands with no bile acids scaffold endowed with pharmacological activity on bile acid receptors, like GPBAR1, leading to the discovery of quinoline scaffold derivatives with dual activity on both GPBAR1 and the cysteinyl leukotriene receptor 1 ($CysLT_1R$). In particular, starting from the identification of REV5901, the first GPBAR1 agonist with a quinoline moiety studied, and exploring its scaffold, its derivatives with double activity on GPBAR1 and $CysLT_1R$ were discovered with therapeutic potential in the treatment of colitis and other inflammatory processes.

The work conducted on membrane receptors was initially carried out during my visiting period at the Faculty of Biomedical Sciences of the Università della Svizzera italiana (USI) in Lugano, under the supervision of Professor Vittorio Limongelli, which lasted only two months due to the COVID-19 emergency, but continued in remote collaboration.

Additionally, the epidemiological emergency due to the COVID-19 disease was addressed to defeat the current pandemic situation starting a drug-repurposing virtual screening campaign was realized in order to identify natural and semisynthetic FDA approved drugs binding the SARS-CoV-2 region binding domain (RBD) of the Spike protein and host cell receptor angiotensin-converting enzyme 2 (ACE2). Interestingly, among the FDA approved drugs, some natural and semisynthetic steroidal scaffold compounds showed promising results to inhibit the Spike RBD/ACE2 interaction.

Furthermore, I was involved in other scientific projects not related to the identification of dual-activity compounds targeting bile acid receptors. First, thanks to the collaboration with Professor Antonio Feliciello, of the Department of Molecular Medicine and Medical Biotechnologies at the University of Naples-Federico II in Naples, the defects underlying the biogenesis of primary cilia, which cause ciliopathies, have been investigated combining *in vitro, in vivo* and *in silico* approaches. Second, in the frame of the collaboration with Professor Stefano Fiorucci of the Department of Medicine and Surgery of the University of Perugia, the ability of AHR pelargonidin agonist to down-regulate the expression of ACE2 and, therefore, to exert anti-inflammatory effects useful in the prevention and treatment of SARS-CoV-2 infection was investigated.

The last section of this thesis, finally, describes the theoretical computational methods employed to carry out the PhD research activity. During these three years of PhD, I have acquired good knowledge about computational methods like docking

calculation, Molecular Dynamics and Metadynamics simulations, Molecular and Essential Dynamics trajectory analysis with the aim to facilitate the identification and development of new potential drug candidates.

Publications

1. Fiorucci, S.; Rapacciuolo, P.; Fiorillo, B.; Roselli, R.; Marchianò, S.; Di Giorgio, C.; Bordoni, M.; Bellini, R.; Cassiano, C.; Conflitti, P.; Catalanotti, B.; Limongelli, V.; Sepe, V.; Biagioli, M.; Zampella, A. Discovery of a potent and orally active dual GPBAR1/CysLT1R modulator for the treatment of metabolic fatty liver disease. *Frontiers in Pharmacology* **2022**, *accepted manuscript,* Manuscript ID: 858137.

2. Fiorillo, B.; Marchianò, S.; Moraca, F.; Sepe, V.; Carino, A.; Rapacciuolo, P.; Biagioli, M.; Limongelli, V.; Zampella, A.; Catalanotti, B.; Fiorucci, S. Discovery of Bile Acid Derivatives as Potent ACE2 Activators by Virtual Screening and Essential Dynamics. *J Chem Inf Model* **2022**, *62*, 196-209.

3. Fiorillo, B.; Sepe, V.; Conflitti, P.; Roselli, R.; Biagioli, M.; Marchianò, S.; De Luca, P.; Baronissi, G.; Rapacciuolo, P.; Cassiano, C.; Catalanotti, B.; Zampella, A. Limongelli V, Fiorucci S. Structural Basis for Developing Multitarget Compounds Acting on Cysteinyl Leukotriene Receptor 1 and G-Protein-Coupled Bile Acid Receptor 1. *J Med Chem.* **2021**, *64*, 16512-16529.

4. Biagioli, M.; Marchianò, S.; Roselli, R.; Di Giorgio, C.; Bellini, R.; Bordoni, M.; Gidari, A.; Sabbatini, S.; Francisci, D.; Fiorillo, B.; Catalanotti, B.; Distrutti, E.; Carino, A.; Zampella, A.; Costantino, G.; Fiorucci, S. Discovery of a AHR pelargonidin agonist that counter-regulates Ace2 expression and attenuates ACE2-SARS-CoV-2 interaction. *Biochem Pharmacol.* **2021**, *188*, 114564.

5. Senatore, E.; Chiuso, F.; Rinaldi, L.; Intartaglia, D.; Delle Donne, R.; Pedone, E.; Catalanotti, B.; Pirone, L.; Fiorillo, B.; Moraca, F.; Giamundo, G.; Scala, G.; Raffeiner, A.; Torres-Quesada, O.; Stefan, E.; Kwiatkowski, M.; van Pijkeren, A.; Morleo, M.; Franco, B.; Garbi, C.; Conte, I.; Feliciello, A. The TBC1D31/praja2 complex controls primary ciliogenesis through PKA-directed OFD1 ubiquitylation. *EMBO J.* **2021**, *40*, e106503.

6. Carino, A.; Moraca, F.; Fiorillo, B.; Marchianò, S.; Sepe, V.; Biagioli, M.; Finamore, C.; Bozza, S.; Francisci, D.; Distrutti, E.; Catalanotti, B.; Zampella, A.; Fiorucci, S. Hijacking SARS-CoV-2/ACE2 Receptor Interaction by Natural and Semi-synthetic Steroidal Agents Acting on Functional Pockets on the Receptor Binding Domain. *Front Chem*. **2020**, *8*, 572885.

Poster presentation

• "Computational studies for the design of modulators of ROR-γ receptor"; (12th European Workshop in Drug Design, Certosa di Pontignano, Siena, Italy, May 2019).

Oral communications

• "Molecular basis of praja2 and TBC1D31 interaction"; (Paul Ehrlich Euro-PhD Network Virtual Meeting - PEVM2021, July 2021).

• "Identification of quinoline-based compounds with dual activity against $CysLT_1R$ and GPBAR1"; (Computationally Driven Drug Discovery Meeting- CDDD 7th Virtual Meeting, June 2021).

Certifications

• "Introduction to Computational Antibody Engineering"; (Schrödinger online course, December 2021).

• "High performance Molecular Dinamics"; (CINECA, Via dei Tizii, 6b, 00185 Roma, Italia, Octuber 2019).

• "BImBS 2019 - BioInformatics meets BioSimulations in protein and DNA studies: from theory to practice"; (USI, Via G. Buffi, 6900, Lugano, Switzerland, Octuber 2019).

Grants

• SuperComputing Resource Allocation – ISCRA-CLASS C project award on MARCONI100 and GALILEO (BDW) at CINECA (BO), March 2020.

• SuperComputing Resource Allocation – ISCRA-CLASS C project award on MARCONI A2 (KNL) and GALILEO (BDW) at CINECA (BO), March 2019.

List of abbreviations

BAs, bile acids; CA, cholic acid; CDCA, chenodeoxycholic acid; DCA, deoxycholic acid; LCA, lithocholic acid; FXR, farnesoid X receptor, GPBAR1, G-protein bile acid receptor; UDCA, ursodeoxycholic acid; CYP7A1, cholesterol-7α-hydroxylase; BSEP, bile salt export pump; ASBT , apical bile acid transporter; ROR, retinoic acid-related orphan receptors; NRs, nuclear receptors; FXREs, FXR response elements; PPARα, peroxisome proliferator-activated receptor alpha; DBD, DNA binding domain; LBD, ligand binding domain; GPCRs, G-protein coupled receptors; IL, intracellular loop; EL, extracellular loop; cAMP, cyclic adenosine monophosphate; GLP-1, glucagon-like peptide 1; NASH, nonalcoholic steatohepatitis; $CysLT_1R$, cysteinyl Leukotriene receptor 1; LTR, leukotriene receptors; $CysLT_2R$, cysteinyl leukotriene receptor 2; LTB4-R1, leukotriene B4 receptor 1; LTB4-R2, leukotriene B4 receptor 2; OXER1, oxoeicosanoid receptor 1; TMPRSS2, transmembrane serine protease 2; NTD, N-terminal domain; CTD, c-terminal domain; RBD, receptor binding domain; RBM, receptor-binding motif; ACE2, angiotensin-converting enzyme 2; HR1, heptapeptide 1 domain; HR2, heptapeptide 2 domain; Ang, angiotensin; RAS, renin-angiotensin system; AT1R, angiotensin type 1 receptor; MasR, Mas G-protein coupled receptor; VCAM, vascular cells adhesion molecules; DIZE, diminanazene aceturate; TNFα, tumor necrosis factor a; IL, interleukin; OFDI, type I digital orofacial syndrome; PKA, protein kinase A; AC, adenylate cyclase; AKAP, A-kinase anchor protein; AhR, aryl hydrocarbon receptor; VS, virtual screening; MD, molecular dynamic; MetaD, metadynamic; FES, free energy surface; CV, collective variable; BE, bias exchange; RMSD, root mean square deviation; ARE, response element; CRE, cAMP responsive element; cryo-EM, Cryogenic electron microscopy; TM, transmembrane; SAR, structure–activity relationship; TLCA, taurolithocolic acid; T-UDCA, tauro-ursodeoxycholic acid; G-UDCA, glyco-ursodeoxycholic acid; OCA, obeticholic acid;

G-CDCA, glyco-chenodeoxycholic acid; G-UDCA, glyco-ursodeoxycholic acid; Com, community; PC, principal component; SASA, Solvent Accessible Surface Analysis; RMSF, root mean square fluctuation; CoIp, co-immunoprecipitation; IEC, intestinal epithelial cells; FA, fatty acid; POPC, phosphatidylcholine; GUI, graphical user interface; PRCG, the Polak-Ribiere conjugate gradient; GAFF, General Amber Force Field; CHL, cholesterol; EM, electron microscopy; PDB, protein data bank; LGA, Lamarckian genetic algorithm; RESP, restrained electrostatic potential; PME, particle mesh Ewald; PCA, principal component analysis; NMWiz, normal mode wizard; REP, upstream of a reporter; TLCA, taurolithocolic acid.

CHAPTER I

Introduction

1.1 Bile acids (BAs)

Bile acids (BAs) are steroidal acids that, together with cholesterol, phospholipids and bilirubin, comprise the principal constituents of bile. They can be classified in: i) primary bile acids, synthesized from cholesterol by the liver; ii) secondary bile acids, resulting from bacterial actions in the colon. Primary bile acids include cholic acid (CA) and chenodeoxycholic acid (CDCA), which have been recognized as functioning as the main farnesoid X receptor (FXR) ligands in humans.[1] Secondary bile acids, like deoxycholic acid and lithocholic acid (DCA and LCA, respectively) generated by intestinal microbiota, are preferential ligands for G-protein bile acid receptor 1 (GPBAR1).[1,2] Ursodeoxycholic acid (UDCA), which is a primary bile acid in mice, but found also in trace in humans where it is not classifiable as primary or secondary bile acid, is, along with CDCA, the only bile acid approved for clinical use, and is a weak agonist for GPBAR1 and considered a neutral or weak antagonist toward FXR.[3]

These endogenous molecules, produced in the liver and intestine by the catabolism of cholesterol,[1,4] are synthesized in the liver through two main pathways, the classical and the alternative: the first one begins with the enzymatic reaction of cholesterol-7α-hydroxylase (CYP7A1) and produces cholic acid (CA) and chenodeoxycholic acid (CDCA); the second one also initiated by CYP27A1, mainly produces CDCA. Subsequently, during metabolism, after being conjugated with taurine or glycine in the C24 position, the BAs are secreted into the biliary system via the canalicular bile salt export pump (BSEP) and the protein 2 associated with multidrug resistance, arrive in the gallbladder, where they are stored and released in the small intestine only after the ingestion of meals. Therefore, once released in the intestine, they are subject to deamidation and 7α-dehydroxylation by the intestinal microbiota, leading to the formation of secondary BAs, deoxycholic acid (DCA) and lithocholic acid (LCA). The secondary BAs are then absorbed in the distal ileum by the enterocyte sodium-

dependent apical bile acid transporter (ASBT) transferred from the apical membrane to the basolateral membrane by the ileal bile acid binding protein and pumped into the portal circulation by transporters such as OSTα/β.[5] The secondary BAs that are not absorbed in the distal ileum reach the colon and are subjected to de-conjugation by intestinal microbiota and passively reabsorbed by the colonocytes or excreted with feces.[6]

BAs also act as signal molecules that negatively regulate their own biosynthesis.[7]

Furthermore, they have important roles in the regulation of lipid, glucose and energy metabolism and, for such reason, they are principally responsible for the digestion of food fats and oil. In particular, they function as micelle-forming surfactants, which encapsulate nutrients, facilitating their absorption.[8]

Recently, they are increasingly being appreciated as complex metabolic integrators and signaling factors, allowing a number of bile-acid-activated signaling pathways to become attractive therapeutic targets for metabolic disorders: for example, obesity-related metabolic diseases, including type 2 diabetes mellitus and non-alcoholic fatty liver disease, are associated with dysregulation of bile acid homeostasis.[9]

BAs act as steroid hormones and perform various metabolic actions through interaction with several nuclear hormones receptors, including the nuclear retinoic acid receptor-related orphan receptor (ROR) and FXR.[10,11,12,13]

Another important bile acid receptor is the cell membrane receptor G protein-coupled bile acid receptor 1 (GPBAR1), considering that it may be involved in metabolic, endocrine and neurological functions.[14,15]

Since the activation of these different signaling pathways by bile acids regulate not only their own synthesis and enterohepatic recirculation, but also the homeostasis of triglycerides, cholesterol and glucose, these bile acid-controlled signaling pathways

have become the source of promising new drug targets for the treatment of common metabolic and liver diseases.

1.2 Bile acids receptors: Nuclear Receptors (NRs)

Nuclear receptors (NRs) are a family of ligand-regulated intracellular transcription factors that regulate metabolism and inflammation. NRs can be classified into two large classes based on their mechanism of action and subcellular distribution. Small lipophilic substances such as natural hormones spread through the cell membrane and can bind to nuclear receptors located in the cytosol (type I NR) or to nuclear receptors located in the nucleus (type II NR) of the cell. In both cases, the binding causes a conformational change in the receptor which, depending on the class of the receptor, triggers a cascade of downstream events that direct the NRs towards the sites of regulation of the DNA transcription. In this site the receptor can be involved in an up-regulation or down-regulation of gene expression and, for such reason, they are classified as transcription factors.[16,17] Therefore, they can function as homo/heterodimers[18] and work with other proteins to regulate the expression of specific genes, thereby controlling the development, homeostasis, and metabolism of the organism.

NRs dysfunction is related to various diseases such as cancer, metabolic and autoimmune disorders and, for such reason, they are considered as promising pharmacological targets.[19]

Among the family of NRs, FXR is one of the best known and studied bile acid receptors, present in the liver and intestine,[13,20,21,22] regulating the transcription of key genes involved in the synthesis and transport of bile acids.

Like other NRs, once activated, FXR translocates to the nucleus where it forms a heterodimer with RXR)[23,24] and binds to the hormonal response elements on DNA, called FXR response elements (FXREs), which up- or down-regulate the expression of some genes.[20]

It has been demonstrated that the activation of FXR indirectly causes the transactivation of the expression of CYP7A1, the enzyme responsible for the catabolism of cholesterol and, therefore, for the synthesis of bile acids,[25,26] indeed, accelerates their excretion.

Therefore, thanks to these multiple capacities, FXR can be considered a real bile acid sensor, necessary to maintain the functionality of the enterohepatic circulation of bile acids and to protect liver cells from possible overloads of cellular bile acids.

FXR activation has also been shown to inhibit lipogenesis and promote the oxidation of free fatty acids by activating peroxisome proliferator-activated receptor alpha (PPARα).[27]

Furthermore, activation of FXR in diabetic mice reduces plasma glucose and improves insulin sensitivity, while inactivation of FXR has the opposite effect.

Retinoic acid-related orphan receptors (RORs) are another member of NRs, a family of ligand-regulated intracellular transcription factors that are activated by steroid hormones to regulate the expression of target genes linked to metabolism and inflammation.

RORs represent key regulators of many physiological processes: for instance, the ROR-α (NR1F1, RORA or RZRα) and ROR-γ (NR1F3, RORC or TOR) isoforms play important roles in glucose and lipid metabolism, while ROR-β (NR1F2, RORB or RZRβ) could be involved in the control of the circadian rhythm. ROR-γ isoform is highly expressed in the thymus (the thymus-specific isoform is referred to as ROR-γt), muscle, testis, pancreas, prostate, heart, intestine and liver.[28,29,30]

Despite their different functional roles, the three RORs isoforms share the general structural characteristics of NRs: 12 α-helices (H1-12) and 2 β-strands, the presence of a highly conserved DNA binding domain (DBD) and a carboxy-terminal ligand binding domain (LBD), this latter includes the H12 that can adopt multiple

conformations modulating the interaction with other co-activators and/or co-factors (Figure 1).

Different ligands (agonist, inverse agonist or antagonist) can be used to modulate the capability of NRs influencing the H12 positioning (Figure 1), which exists in a closed active form, stabilized by an H-bond between His479 and Tyr502, and in an open inactive form, in which such H-bond is disrupted.

ROR-γ agonists, such as 25-hydroxycholesterol, induce a conformational change into the protein structure. In particular, the agonist binding stabilizes the H-bond between His479 and Tyr502 on H11 and H12, respectively. This receptor stabilization allows the recruitment of transcriptional coactivators, such as SRC2, which leads to the modulation and promotion of target gene transcription. Instead, inverse agonists of ROR-γ, such as digoxin, through a path of interaction which includes Leu324, Trp317, His479 and Tyr502, disrupt the H-bond which stabilize H11 and H12 in a closed and active conformation. The H-bond disruption causes a conformational change of H12, which prevents the recruitment of the transcriptional coactivator and repress target gene expression.[12]

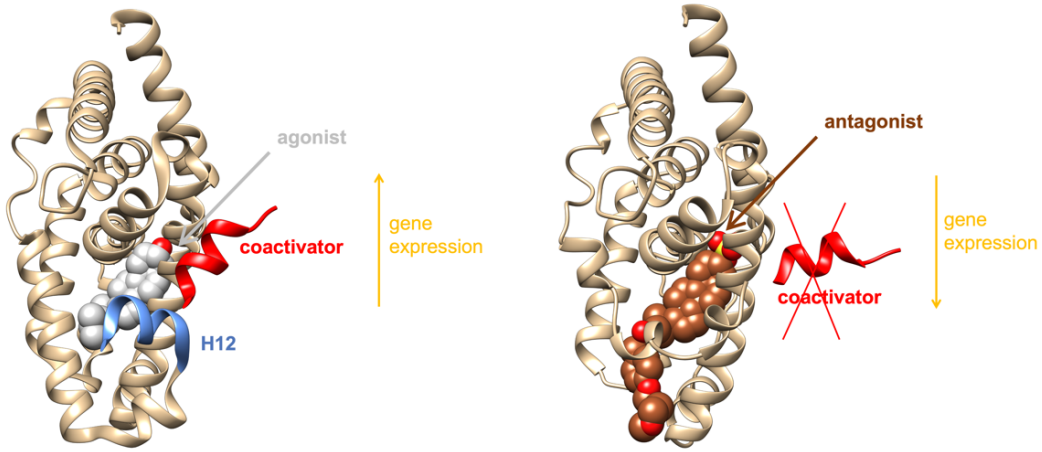


Figure 1. ***Structural basis for the mechanism of action of RORs agonists, inverse agonists and antagonists.***

Several cholesterol intermediates and metabolites function as natural ligands of ROR-α and ROR-γ and act as agonists or inverse agonists (Figure 2): side chain oxygenated sterols (e.g., 20 α-hydroxycholesterol) are high affinity ROR-γ agonists;[12] while sterols oxygenated at the 7-position, (e.g., 7-hydroxycholesterol) function as inverse agonists for ROR-γ.[31] Studies revealed that ROR-α and ROR-γ exhibit quite a wide binding specificity for many sterols.[21,22]

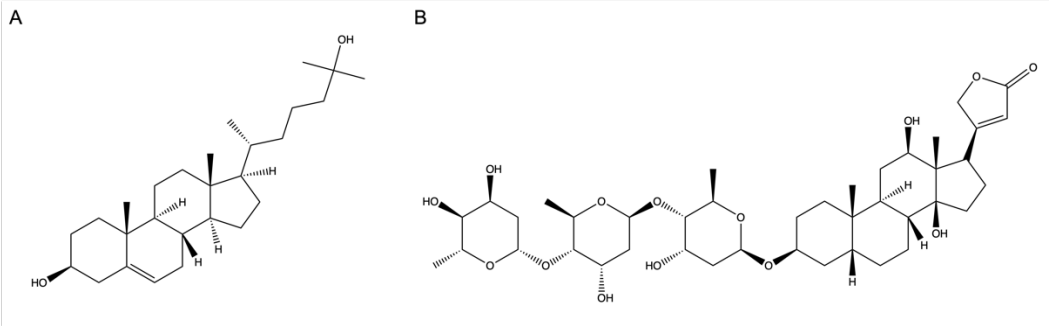


Figure 2. ***Previously reported RORs ligands: A) (25R)-hydroxycholesterol, agonists; B) digoxin, inverse agonists.***

Many synthetic inverse agonists have also been identified for the ROR-γt isoform for the treatment of inflammatory diseases, while ROR-γt agonists have been proposed for use as immuno-oncology agents to activate the immune system to treat cancer.[32] Both of them can have a steroidal scaffold or not steroidal scaffold.

Recent advances have established that the inhibition of RORs is a promising therapeutic approach for the treatment of autoimmune diseases, metabolic disorders and some cancers.[33]

1.3 Bile acids receptors: G-protein coupled receptors (GPCRs)

In the last decade, multi-target drugs have achieved considerable interest for their advantages in the treatment of complex health conditions linked to drug resistance problems or in favoring co-synergism in therapies.

In pathologies related to lipid and glucose disorders in which ROR receptors are involved,[33] an emerging role in medicinal chemistry research is also played by the GPBAR1 (also known as TGR5), which is a member of one of the largest membrane protein family, G-protein coupled receptor (GPCRs), targeted by approximately 40% of the marketed drugs.[2] GPCRs can be activated by several types of ligands, including light-sensitive compounds, odors, pheromones, hormones, and neurotransmitters, and their size varies from small molecules to peptides and large proteins. GPCRs general structure consists of an extracellular N-terminus, followed by seven transmembrane (7-TM) α-helices (TM1 to TM7) connected by three intracellular (IL1 to IL3) and three extracellular loops (EL1 to EL3), and finally an intracellular C-terminus (Figure 3).

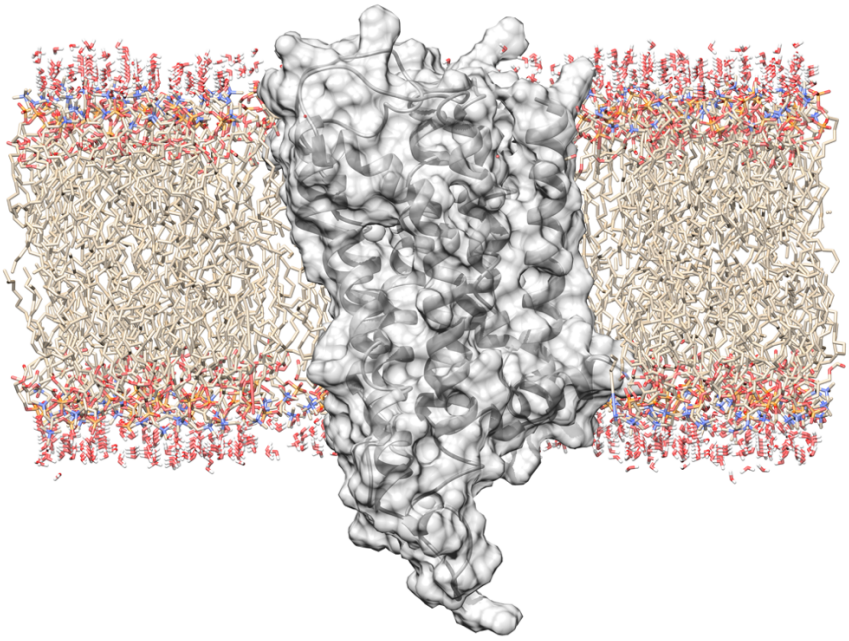


Figure 3. ***Representation of the 3D model of GPBAR1 embedded in the phospholipid bilayer.*** GPBAR1 is represented in grey cartoon, the membrane is represented as tan sticks and water molecules are displayed with their explicit hydrogens.

GPBAR1 is highly expressed in the liver, intestine, gallbladder, brown adipose tissue, muscles, and immune cells.[34,35]

Targeting GPBAR1 has demonstrated being a valid strategy to contrast hepatic inflammation, steatohepatitis, biliary diseases and metabolic syndromes.[36]

Bile acids are natural agonists of GPBAR1 and their binding induces an increase in intracellular cAMP levels, which in turn activates specific intracellular signaling cascades and internalization of the receptor.[34]

In entero-endocrine L cells, ligand-mediated activation of GPBAR1 stimulates the secretion of the glucagon-like peptide 1 (GLP-1), whose function is to enhance insulin release from the pancreas, thereby regulating levels of blood glucose, gastrointestinal motility and appetite and with the final aim of attenuating insulin resistance. In muscle and brown adipose tissue, ligand binding to GPBAR1 induce an increase in energy expenditure and oxygen consumption, which in turn increases the basal metabolism.[37]

Furthermore, GPBAR1 is known for having potent anti-inflammatory effects.[38] In fact, it has been observed that its activation suppresses endothelial dysfunction in rodent models of nonalcoholic steatohepatitis (NASH).[39,40]

All these data reveal GPBAR1 as a promising target in the pharmacotherapy of enterohepatic and metabolic disorders.[14,41,42]

1.4 Other targets for bile acids receptors' (BARs) ligands

In the frame of the discovery of multi-target compounds, the PhD project has evolved in the analysis of other targets for ligands of bile acid receptors (BARs). In particular, the main characters were cysteinyl leukotriene receptor 1 ($CysLT_1R$), SARS-CoV-2 Spike protein and ACE2.

First, considering that drug reprofiling, or repositioning, is a growing area in the drug development process as an attractive strategy to take advantage of previous preclinical and clinical investigations and to invest in candidates originally designed for targeting different contexts, my research group had previously discovered the capability of alpha-pentyl-3-[2-quinolinylmethoxy] benzyl alcohol - REV5901 - (Figure 4) to activate GPBAR1 with an EC_{50} of 2.5 μM and inhibit $CysLT_1R$ with an IC_{50} of 1.1 μM.[43] Finding new drugs capable of simultaneously modulating $CysLT_1R$ and GPBAR1 has proven useful in the therapeutic potential of treating colitis and other inflammatory processes.

Furthermore, due to the coronavirus disease 2019 (COVID-19) pandemic and in light of the urgent need to identify novel approaches to be used in the emergency phase, my research group and I have embarked on an exploratory virtual screening campaign aimed at repurposing natural substances and clinically available drugs as potential anti-SARS-CoV-2 agents by targeting Spike SARS-CoV-2 viral protein, as well as the human target of the spike, the angiotensin converting enzyme ACE2.

Interestingly, we discovered that bile acids can interact with both targets.

1.4.1 Cysteinyl Leukotriene receptor 1 ($CysLT_1R$)

Leukotriene receptors (LTR) are another pharmacologically relevant subfamily of class A GPCRs composed of five members: $CysLT_1R$ and 2 $CysLT_2R$, leukotriene B4 receptor 1 (LTB4-R1) and 2 (LTB4-R2), and Oxoeicosanoid receptor 1 (OXER1). The

endogenous ligands of these receptors are leukotrienes, which are eicosanoids derived from the oxidation of arachidonic acid and act as mediators in inflammatory processes. In particular, the cysteinyl leukotrienes (CysLT) C4 (LTC4), D4 (LTD4) and E4 (LTE4) are endogenous ligands of $CysLT_1R$ and $CysLT_2R$ with different potency and affinity (Figure 4).[44]

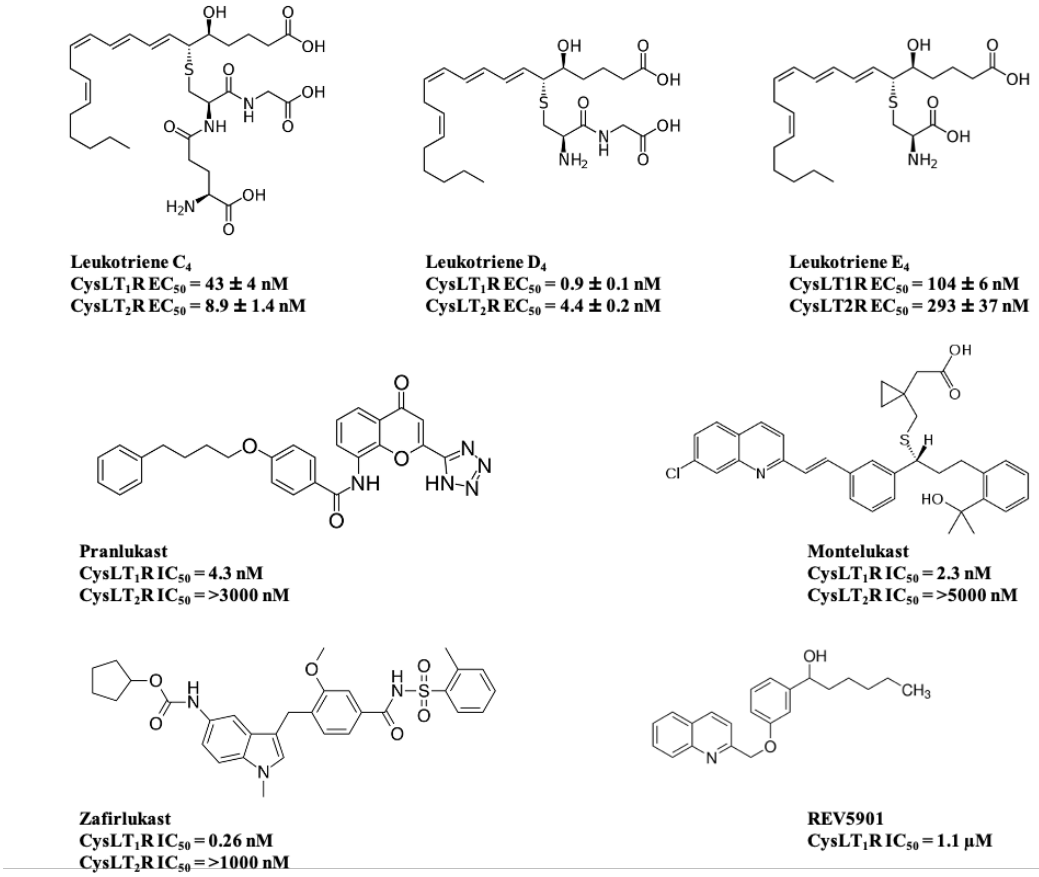


Figure 4. ***Endogenous ligands and previously reported CysLT1R antagonists.*** Credit for the figure go to Ref.[45]

Leukotrienes activate CysLTRs by generating cellular responses through intracellular interaction with $G_{q/11}$ or $G_{i/o}$ proteins, with the final goal in regulating cytokine secretion, vascular permeability, fibrosis, bronchoconstriction and recruitment of effector cells and mucus.[46,47,48,49,50,51] Among the various isoforms,

$CysLT_1R$ is known to mediate allergic and hypersensitivity reactions and its activation leads to pathological conditions such as asthma and allergic rhinitis.

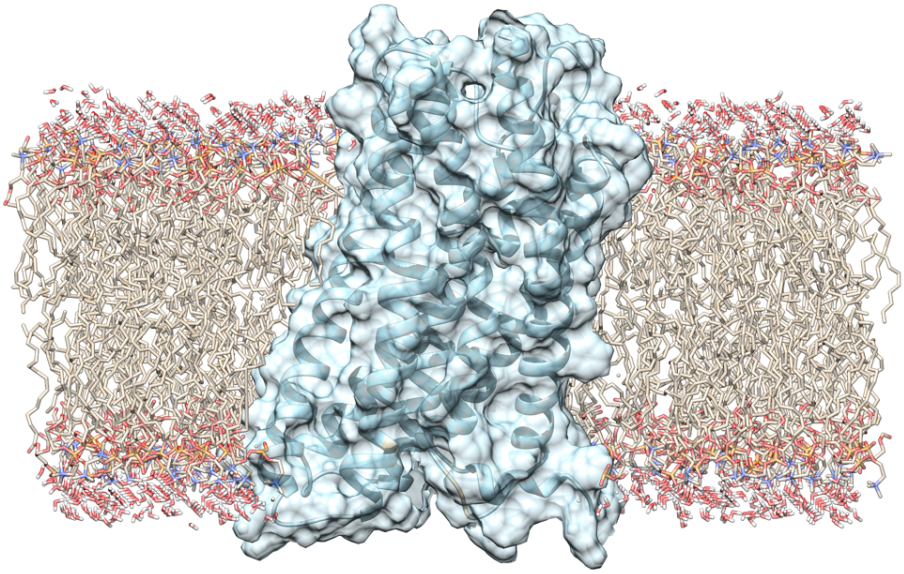


Figure 5. ***Representation of the 3D model of $CysLT_1R$ embedded in the phospholipid bilayer.*** $CysLT_1R$ is represented in cyan cartoon, the membrane is represented as tan sticks and water molecules are displayed with their explicit hydrogens.

The $CysLT_1R$ receptor (Figure 5) is highly expressed in bronchial fibroblasts, airway mucosal cells, pulmonary smooth muscle and several types of pulmonary leukocytes, especially in asthmatic patients.[52,53] It is also present in the small intestine, pancreas, liver, colon, and vascular endothelial cells.[54]

The discovery of the first $CysLT_1R$ antagonists montelukast, zafirlukast and pranlukast, the first antagonists of $CysLT_1R$ (Figure 4), had a great impact on the treatment of asthma and respiratory morbidity and, thus, the interest in the development of other $CysLT_1R$ antagonists has always become higher, reaching, for some of the new ligands, the steps of preclinical and clinical trials.[47]

$CysLT_1R$ have been shown to be of great pharmacological interest as they could act on several independent cellular pathways leading to an anti-inflammatory effect with

reduced cytokine secretion and expression of chemokines, monocyte differentiation, vasodilation, macrophage migration and modulation of cell proliferation.[46,47]

1.4.2 SARS-CoV-2 Spike protein

COVID-19 disease is characterized by a respiratory tract infection caused by severe acute respiratory syndrome SARS-CoV-2, a coronavirus belonging to the *Coronaviridae* family first identified in the city of Wuhan in China in December 2019.[55]

Figure 6. ***Representation in cartoon of Spike protein (PDB ID 6zb5).***[56]

SARS-CoV-2 is a single-stranded RNA virus, presenting numerous glycosylated Spike proteins (S protein) (Figure 6), trimeric structures which surround the virus like

a "crown" and mediate the infection binding the target cell. In particular, Spike protein binds the human angiotensin converting enzyme receptor (ACE2),[57,58,59,60,61,62] a transmembrane type 1 mono carboxypeptidase expressed on the surface of epithelial cells and this event, together with the proteolytic breakdown of ACE2 by the serine protease2 (TMPRSS2), allows the virus to enter the cell.

There are two possible processes for incorporating the viral RNA genome: an endocytic pathway or a non-endocytic pathway.[63] The interaction, therefore, takes place through recognition of the host cell and fusion of the viral membranes and host cells.[64]

Therefore, a pre-fusion and a post-fusion conformation can be identified.

The Spike protein has three-dimensional "mushroom-like" conformations, with the S1 subunit forming the head, and the S2 subunit forming the stem. The pre-fusion conformation involves three S1 binding regions positioned at the top of the S2 trimeric fusion membrane. Conversely, the post-fusion conformation is a coiled-coil containing only the S2 subunit.

The S1 subunit is composed of:[65,66]

• N-terminal domain (NTD) (residues Gln14–Ser305);

• C-terminal and Receptor binding domain, CTD/RBD (residues Arg319–

Phe541), responsible for the interaction with the host cell through ACE2 recognition mediated by the receptor-binding motif (RBM), a part of the RBD making direct contacts with ACE2;

• FP domain (residues Thr788–Pro806);

• Repetitive heptapeptide 1 domain (HR1) (residues V912–V984);

• Repetitive heptapeptide 2 domain (HR2) (residues Leu1163–Arg1213);

• TM domain (residues Arg1213–Arg1237);

• Cytoplasmic domain fusion (CT) (residues Arg1237–His1273).

Structural studies have revealed that the S1-CTD region adopts a mixed population of conformational states, characterized by the angles between the RBDs, that seem to switch continuously between an "up" conformation that allows the binding to ACE2 and a "down" conformation that prevents it (Figure 7).[67] When the binding occurs, the S1-CTD portion stabilizes in its "up" conformation, representing the highest energy state, in which the protein is ready to bind with the host cell receptor.[68]

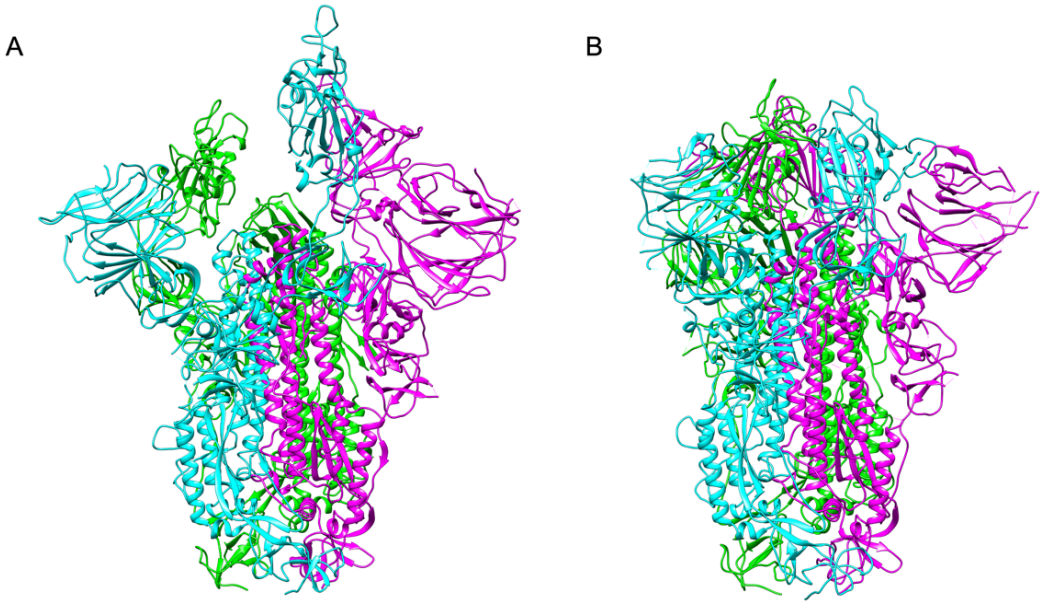


Figure 7. ***Representation in cartoon of A) "Up" (PDB ID 6crx)[69] and B) "Down" (PDB ID 6zb5)[56] conformational states of the Spike S1-CTD portion.***

Furthermore, the binding of the S protein RBD to ACE2 activates the proteolytic cleavage of the S2 subunit by TMPRSS2, enabling the entry of the virus into cells, promoting the viral replication and cell-to-cell transmission and the consequent spread of the coronavirus throughout the host. [66,67,70,71,72] Furthermore, a second S2 cleavage site was identified at the S2 ′ site which is believed to be essential for the activation of the protein responsible for membrane fusion.

1.4.3 Angiotensin converting enzyme 2 (ACE2)

ACE2 is a membrane protein, expressed in several tissues, such as arteries, heart, lung, gastrointestinal tract and kidney,[73,74,75] and it is important in blood pressure regulation, cardiac/renal function and fluid and electrolyte balance.

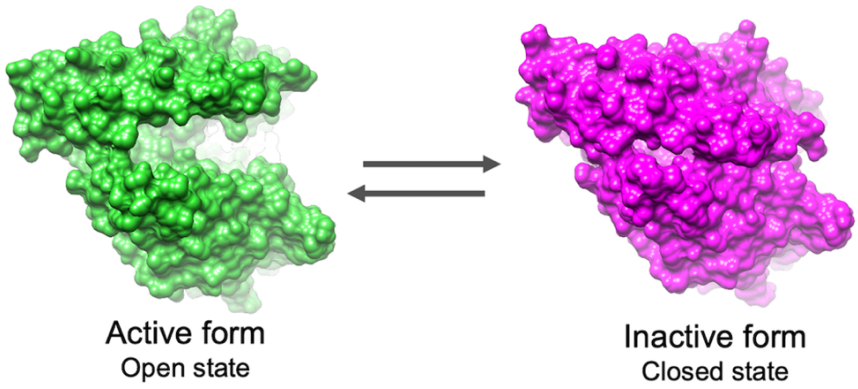


Figure 8. ***Surface representation of ACE2 in different conformations.***

In recent years, great pharmacological interest has been directed to the study of the biological role of ACE2, essentially related to the extracellular enzymatic conversion of the vasoconstrictor peptide angiotensin (Ang) II into the vasodilator peptide Ang (1-7).

ACE2 exists in two conformations: an active-open state and an inactive-closed state (Figure 8). In fact, in the active form, the enzyme increases the Angiotensin (1-7) and decreases the Angiotensin II concentrations. In particular, in the renin-angiotensin system (RAS), the angiotensin converting enzyme (ACE) converts Ang I to Ang II, which activates the Ang II type I receptor (AT1R), with the final result of vasoconstriction, cardiac and vascular hypertrophy, increased fluid retention, oxidative stress and tissue fibrosis.[76,77,78,79] In contrast, ACE2 hydrolyzes Ang II into Ang (1-7), with the final effect of reducing Ang II which, through AT1R signaling, prevents vasoconstrictor and proinflammatory effects. Furthermore, the increase in Ang (1-7) causes the activation of the receptor coupled to the G Mas protein (MasR),

thus determining vasodilation and the reduction of oxidative stress and inflammation due to the release of nitric oxide, bradykinin. and prostaglandin E2.[80,81,82] This is due to the ability of Ang (1-7) to mediate leukocyte migration, cytokine expression and release and fibrogenic pathways.[80] Hence, it has been demonstrated that inactivation of ACE2 results in increased vascular inflammation and atherosclerosis in ApoE knockout mice due to increased expression of vascular cell adhesion molecules (VCAM), chemokines, cytokines and MMPs.[83] On the contrary, the activation of ACE2 and, thus, the consequent increase of Ang (1–7) and activation of MasR, on the one hand, determines a decrease in the expression of inflammatory mediators, such as Il-1β, Il-6, TNFα and MCP-1, on the other hand, causes an increased expression of the anti-inflammatory cytokine, Il-10.[83,84,85]

Moreover, pharmacological studies have also reported the relationship between the ACE2/Ang (1-7)/MasR axis and cancer growth suggesting the therapeutic potential of ACE2 regulation in cancer treatment.[86]

The interaction with SARS-CoV-2 induces downregulation of ACE2 expression that results in the alteration of the physiological balance between Ang II and Ang (1-7), increasing the Ang II-mediated RAS signaling and, on the other side, depleting the protective effects mediated by ACE2/Ang (1-7)/MasR axis. These phenomena have been studied in COVID-19 patients, who have experienced cytokine storm and coagulopathy.[87,88] Therefore, modulating the activity of ACE2 could represent a potential approach to limit the damage due to the excessive inflammatory response in COVID-19 disease.[89]

Despite its important biological role, a very limited number of ACE2 activators are known,[90,91] like the diminazene aceturate (DIZE), an anti-trypanosomiasis veterinary drug that is also capable to activate ACE2.[92] In experimental models, dize has been shown to have protective effects in the case of hypertension, kidney disease, liver

injury, myocardial infarction, ischemia[93,94,95,96] and finally, it attenuates inflammation in a NF-κB dependent manner.[97] For these protective effects, the use of DIZE as a therapeutical agent for COVID-19 patients had been proposed,[98] but it demonstrated cytotoxic side effects at therapeutic doses.[99]

1.5 Theoretical Methods

During my PhD activity, I have exploited the advantages of the applications of both classical and advanced computational techniques to the field of medicinal chemistry, which include molecular docking calculations, virtual screening studies, molecular dynamics (MD) and metadynamics (MetaD) simulations.

These methodologies are discussed in the following section.

1.5.1 Molecular docking

Molecular docking is a widely applied methodology in the field of computational chemistry and to predict the preference of a molecule to orient itself with respect to another in order to form a stable complex according to the interactions that are formed with the target.[100] The preferred orientation, explained as binding mode, can also be used to predict the strength of the association or the binding affinity between two molecules using the scoring functions.

The affinity between two molecules is expressed through the dissociation constant. The formation of a ligand-protein complex:

$$\mathrm{L} + \mathrm{P} \rightleftharpoons \mathrm{LP}$$

and the corresponding dissociation constant is defined:

$$K_D = \frac{[\mathrm{L}]\,[\mathrm{P}]}{[\mathrm{LP}]}$$

where [P], [L] and [LP] represent the molar concentration (M) of protein, ligand and complex, respectively. The smaller the dissociation constant, the stronger the binding of the ligand and, therefore, the greater the affinity between ligand and protein.

Scoring functions, on the other hand, are mathematical functions used to roughly predict the binding affinity between two molecules after they have been docked.

Scoring functions are widely used in molecular modeling applications, like the Virtual Screening (VS) databases[101] of candidate ligands, typically used to identify new small molecules that bind to a protein target of interest.

1.5.2 Virtual Screening (VS)

Virtual screening (VS) is another computational technique used in drug discovery to identify small molecules that are more likely to bind to a drug target, typically a protein receptor or enzyme.[102,103] There are two types of screening techniques, ligand-based (LBVS) and structure-based (SBVS).[104]

The ligand-based method is to be preferred when structural information of the receptor is absent. This technique is based, in fact, on the principle that having a set of ligands with known activity against a specific target, a common pharmacophore model capturing their structural features can be built, thus allowing to identify, from a database novel compounds, those binding the same target as the known compounds do.[105,106] The small molecule database can also be exploited in the search for molecules with a shape similar to that of known active ingredients, as these molecules will more likely adapt better to the target binding site.

On the other hand, the structure-based (SBVS) method works directly with the 3D structure of a macromolecular target or a macromolecule/ligand complex. Differently from LBSV, SBVS takes into account the complementary chemical features of the target active site and its spatial relationship.

1.5.3 Molecular dynamics (MD)

Another computational methodology is MD that allows to reproduce and predict, using appropriate approximations, the microscopic behavior of the system under

examination, i.e., the stability of proteins, conformational changes, protein folding and molecular recognition of proteins, DNA, membranes, complexes. Classical MD simulations allow to extract the trajectories of atoms by integrating Newton's equation of motion for a system of N atoms: the solution, that is the coordinates of the system as a function of time, represents the trajectories of the atoms that make up the molecule,

$$\mathbf{F}_i = m_i \mathbf{a}_i$$

where

$$i = 1, \ldots, N$$

From the knowledge of the forces acting on each atom, it is possible to determine the acceleration of each atom in the system. The integration of the equations of motion allows obtaining a trajectory that describes the variation over time of the positions, velocities and accelerations of each particle,

with

$$\mathbf{F}_i = -\frac{\partial V(\mathbf{r})}{\partial \mathbf{r}_i}$$

and V (r), the potential, is a function of the positions of the atoms.

In this equation it is assumed that the motion of atomic nuclei can be described by classical dynamics, if we consider that the distance in the energy levels of the degrees of freedom involved is << kT, where k is the Boltzmann constant and T the temperature.

The potential energy of the system is a function of the atomic positions of all the atoms that composes the system. It is defined as the sum of internal (or "binding") and external (or "non-binding") terms:

$$V(r) = V_{bound} + V_{unbound}$$

where, V_{bound} is:

$$V_{bound} = V_{bond\text{-}stretch} + V_{angle\text{-}bend} + V_{torsions}$$

In order to calculate a trajectory, it is, therefore, necessary to know the initial positions of the atoms, an initial distribution of velocities and acceleration, determined as the gradient of a potential energy function: the initial positions can be obtained from experimental structures resolved with the X-ray diffraction methodologies or by magnetic-nuclear resonance (NMR); The distribution of the initial velocities are derived from a random Maxwellian distribution corresponding to the desired temperature (T):

$$p(\vec{v_i}) = \sqrt{\frac{m_i}{2\pi K_B T}} \exp\left[-m_i \vec{v_i}^2 / 2K_B T\right]$$

where $p(v_i)$ is the probability that the i-th atom has velocity v_i at temperature T; finally, the initial random distribution of the speeds is chosen in such a way that there is no total momentum of the momentum, that is:

$$\vec{P} = \sum_{i=1}^{N} m_i \vec{v_i} = 0$$

Among the popular sets of parameters (force fields) for MD simulations of proteins, we can cite for example AMBER,[107] GROMOS,[108] CHARM[109] and OPLS.[110] They all use the potential function expression for all the atoms of the simulated system except for the GROMOS (and CHARMM19 force field) force field in which a united atom description is used for non-polar hydrogens.

In MD simulations the consideration of the solvent, typically water for systems of biological interest, can be explicit or implicit. In the explicit description, the solvent molecules and their atomic force field are added in the simulation box to the experimental density. In the implicit description, however, the solvent is treated as a dielectric medium in which the system is immersed. The latter is, on the one hand, a

more approximate description but on the other, it is also computationally much more efficient since in many practical cases the solvent constitutes the majority of the atoms. The most commonly used explicit solvent model is represented by the TIP3P model.[111]

1.5.4 Metadynamics (MetaD)

Metadynamics (MetaD)[112,113,114] is a computational method aimed at improving the sampling of configurational space, exploring the properties of multidimensional free energy surfaces (FES) of systems by means of non-Markovian coarse-grained dynamics in the space defined by a few collective coordinates, called collective variables (CVs). The application of this technique takes place thanks to the presence of a history-dependent bias potential, which, over time, fills the minima in the FES, allowing the efficient exploration and accurate determination of the FES as a function of the CVs.

The relevance of this method is to make possible the estimate of the free energy surface of systems that are not tractable in reasonable computer time with regular unbiased molecular dynamics.

By applying a biased potential on the system in question, the latter is forced to leave the local minima and, therefore, to sample low-probability states.

To force the system to escape local minima, MetaD method periodically adds a small Gaussian hill to the potential energy of the current state space region (Figure 9): the local minimum is slowly filled by adding several Gaussian hills, and the system is therefore forced to explore different configurations.

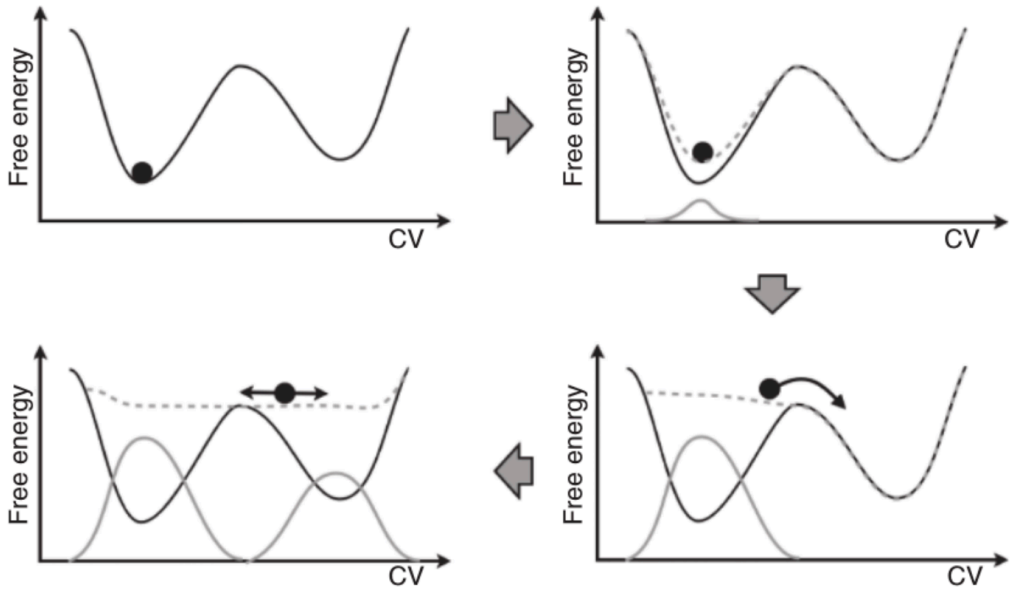


Figure 9. ***Schematic representation of MetaD method.*** First, the system evolves according to a normal dynamic (black ball), then a Gaussian potential is deposited (solid gray line). This lifts the system and changes the free energy landscape (dashed gray line) in which the dynamics evolve. Subsequently, the sum of the Gaussian potentials fills the first metastable state and the system moves into the second metastable basin. With the addition of the potentials, the second metastable basin is also filled and, finally, the system evolves into a flat landscape. The summation of the deposited bias (solid gray profile) provides a rough negative estimate of the free energy profile. Credits for the figure go to Ref.[115]

If the simulation is prolonged for a sufficient time (ms), all the minima will fill with accumulated Gaussian hills and the total potential energy will become flat.

The metadynamics bias potential can be written as:

$$V_G(\mathbf{s},t) = w \sum_{t'=\tau_g,2\tau_g,\ldots}^{t'<\tau} \exp\left(-\sum_{i=1}^{N} \frac{(s_i - s_i(t'))^2}{2\delta s_i^2}\right)$$

where τ is the total simulation time τ_g is the frequency at which the Gaussians are added and δs_i is the width of the Gaussian for each CV.

Although MetaD is an advanced technique used to accelerate rare events and to reconstruct free energy, it suffers from several limitations. A big problem is the filling speed which decreases exponentially with dimensionality. For this reason, no more than 3 CVs can be used in MetaD. Another limitation is represented by the influence

exerted by the choice of CVs on the reliability of MetaD. Therefore, in the MetaD calculations, being constrained to the choice of only 3 CVs, one could fall into the error of choosing a CV and neglect another that could have been more relevant for the purposes of the calculation.

1.5.5 Bias exchange (BE)

In order to escape these MetaD limitations, the bias exchange (BE)[116] was developed. In this method, a large set of CVs is chosen which should be relevant to the process under consideration. A number NR (number of replicas) of MD simulations (walkers) are run in parallel, biasing each walker with a metadynamics bias acting on just one or two collective variables. In BE, sampling is improved by exchanging the bias potentials between pairs of walkers at fixed time intervals. The swap is accepted with a probability

$$\min\left\{1, \exp\left[\left(V_G^a\left(r^a,t\right)/T + V_G^b\left(r^b,t\right)/T - V_G^a\left(r^b,t\right)/T - V_G^b\left(r^a,t\right)\right)/T\right]\right\}$$

where r^a and r^b are the coordinates of walker *a* and *b* and $V_G^{a(b)}$ (r, t) is the metadynamics potential acting on the walker *a(b)*. In this way, each trajectory evolves through the high dimensional free energy landscape in the space of the CVs sequentially biased by different low dimensional potentials acting on one or two CVs at each time.

The advantage of this method is that a large number of different variables can be biased, simultaneously, and a much larger dimensionality of space can be explored. The result of the calculation is not a multi-dimensional free energy, but several low-dimensional projections of the free energy surface along with all CVs.

CHAPTER II

Results and Discussion

2.1 Design of dual modulators against the nuclear retinoic acid-related orphan receptor γ (ROR-γ) and the membrane receptor G-protein bile acid receptor 1 (GPBAR1)

In the effort to identify new improved profile BAs receptor ligands with an improved profile to treat metabolic disorders, our investigation was focused on the discovery of new steroidal scaffold ligands that deactivate ROR-γ and, at the same time, activate GPBAR1.

In this study, using cholesterol as a template for dual ROR-γ/GPBAR1 modulation, a new generation of steroidal ligands embedded with multi-target activity was synthesized (Figure 10). The substitution on C24 of cholesterol with a *para*- or *meta*-substituted phenoxy group led us to the synthesis of compounds **1-12**, carried out by the research group of Professor Angela Zampella at the Department of Pharmacy of University of Naples-Federico II. Such compounds were analyzed through *in vitro* assays and, successively, the chemical and structural features of the ligand-receptor interaction were studied in terms of *in silico* analysis in order to find cholesterol derivatives as potent dual modulators of ROR-γ and GPBAR1 for the treatment of metabolic diseases.

Since the results of this project are neither published nor patented yet, I prefer to obscure the compounds' chemical structure over this thesis.

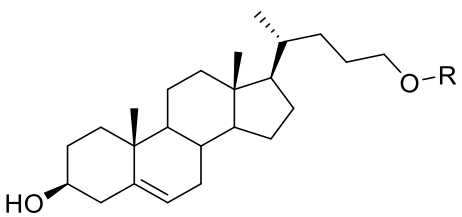


Figure 10. ***General chemical structure of derivatives compounds 1-12 identified in this study.***

2.1.1 Pharmacological activity

The activity of the synthetized derivatives **1−12** on GPBAR1 was calculated in terms of transactivation assay. The *in vitro* transactivation assay is based on stable transfection of a cell line with two plasmids: one encoding the receptor and the other, the response element (ARE) upstream of a reporter (REP) gene such as luciferase. When the transfected cell is exposed to a molecule endowed with agonistic activity, the receptor moves into the nucleus, dimerizes, binds to the ARE and activates the expression of REP which can be monitored. In this case, the efficacy was assessed in HEK-293T cells transfected with a cAMP responsive element (CRE) cloned upstream to the luciferase gene. For calculation of efficacy data, maximal transactivation of CRE caused by each compound (10 μM) was compared to maximal transactivation caused by the reference compound, CDCA (10 μM), which was set as 100%.

The binding affinity against ROR-γ was tested using the alpha screening assay. By carrying out the experiment in the presence of increasing concentrations of the inverse agonist compounds under examination, it is possible to quantify the efficacy of the compounds themselves and to construct a dose-response curve, calculating the IC_{50}.

The results shown in Table 1 demonstrated that the best match in terms of % dual-activity has been found for compounds **3** and **7**.

Table 1. ***Efficacy and Potency for Compounds 1−15.***[c]

Compound	*GPBAR1 EC_{50} (μM)* [a]	*ROR-γ IC_{50} (μM)*[b]
1	na	nt
2	na	0.04±0.010
3	0.22	1.19±0.73
4	na	nt

5	na	0.25±0.08
6	na	nt
7	1.28	0.11±0.03
8	na	1.60±0.51
9	na	nt
10	na	0.21±0.07
11	nt	0.23±0.19
12	nt	0.03±0.01

[a] These assays were performed in terms of transactivation, carried out by the research group of Professor Stefano Fiorucci at the Department of Medicine and Surgery of the University of Perugia (Perugia, Italy). [b] These assays were performed in terms of alphascreen assay, carried out by Doctor Chiara Cassiano at the Department of Pharmacy of the University of Naples-Federico II (Naples, Italy). [c] na, not active; nt, not tested.

In order to elucidate the molecular basis of this set of newly designed cholesterol-derived (Figure 10), compounds **1-12**, with potential dual activity, an inverse agonistic activity on ROR-γ and agonistic activity on GPBAR1 and on the basis of pharmacological assays, molecular modeling studies were carried out on the most promising dual-activity derivatives, compound **3** and compound **7**.

2.1.2 Computational studies

Firstly, molecular docking calculations were performed on compound **3** and compound **7** using the Glide software package[117,118] with the Standard Precision (SP) algorithm. For this study, the human crystallographic structure of the active conformation of the ROR-γ-LBD domain with PDB ID 3l0j[119] was employed. For GPBAR1 receptor, we used the 3D model developed in 2014 by D'Amore *et al.*[120] and already successfully employed by our group in other works.[121,122,123]

The docking analysis of compounds **3** and **7** predicted binding modes to ROR-γ showed that the binding modes of the steroidal scaffold were similar, also with respect to the co-crystallized ligand.[119]

Binding mode of 3 and 7 in ROR-γ. The peri-hydro-1,2-cyclopentane-phenanthrene scaffold of **3** is placed in the amphipathic pocket between helices (H) H4 and H5, where it is further stabilized by a set of hydrophobic interactions established by the steroidal moiety with the side chains of Leu287, Val361, Val376 and Phe378 (Figure 11A). The 3'-OH makes H-bond with Gln286 on H1. Instead, the flexible chain (-R, Figure 10) bends slightly, probably due to the presence of the H11. The alkyl linker makes hydrophobic interactions with Phe388, Ile397, Ile400, while the ethereal oxygen H-bonds Cys320. Finally, the aromatic ring protrudes towards H11, in the amphipathic pocket formed by H4, H8, H11, H12, and makes hydrophobic interaction with Val324 and Leu391 and a T-shaped with Trp317 (Figure 11A). The same interactions of compound **3** were shown for compound **7**, whose phenolic oxygen makes an additional H-bond with the backbone of Trp317 (Figure 11B).

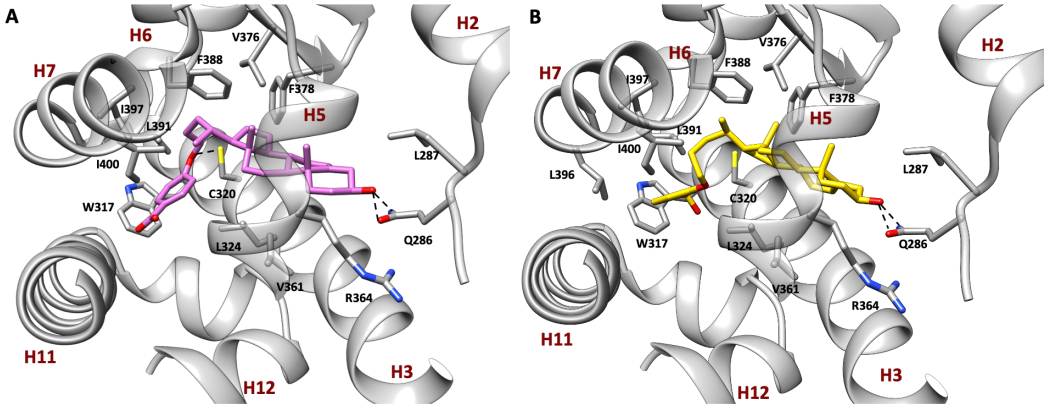


Figure 11. ***Binding mode of compounds (A) 3 and (B) 7 against ROR-γ.*** Ligands are represented as *pink* and *gold* sticks, respectively, whereas the interacting residues of the receptor are shown in *grey* and labelled. Oxygen atoms are depicted in *red* and nitrogens in *blue*. The receptors are represented as ribbons with their helices (H) labelled. Hydrogens are omitted for the sake of clarity and H-bonds are displayed as *black* dashed lines.

Binding mode of 3 and 7 in GPBAR1. Docking calculations showed that **3** and **7** binding mode to GPBAR1 (Figure 12) is similar to other bile acids, recently discovered, with agonistic activity.[120,121,124] Compound **3** steroidal scaffold is placed in the amphipathic pocket between transmembrane helices (TM) TM3, TM5 and TM6, where it makes hydrophobic interactions with Phe96$^{3.36}$ (superscripts refer to Ballesteros-Weinstein numbering)[137], Phe138$^{4.56}$ and Leu174$^{5.47}$. Moreover, the 3'-OH makes a H-bond with the backbone of Glu169$^{5.42}$, which is known to be involved in the binding of agonists and receptor activation.[120,125,126] The flexible side chain extends towards TM1 and TM2 making hydrophobic interaction between the alkyl linker, Leu71$^{2.59}$ and Leu266$^{7.39}$. Finally, the carboxyl group H-bonds Trp75$^{2.63}$ (Figure 12A). Also docking calculations performed in GPBAR1, showed that the binding mode of compound **7** was similar to **3** with the difference in the hydroxyl terminal group, which makes a H-bond with the backbone of Leu68$^{2.57}$ (Figure 12B). Moreover, the steroidal scaffold makes hydrophobic contacts also with Leu97$^{3.37}$, while the aromatic group with Leu71$^{2.59}$.

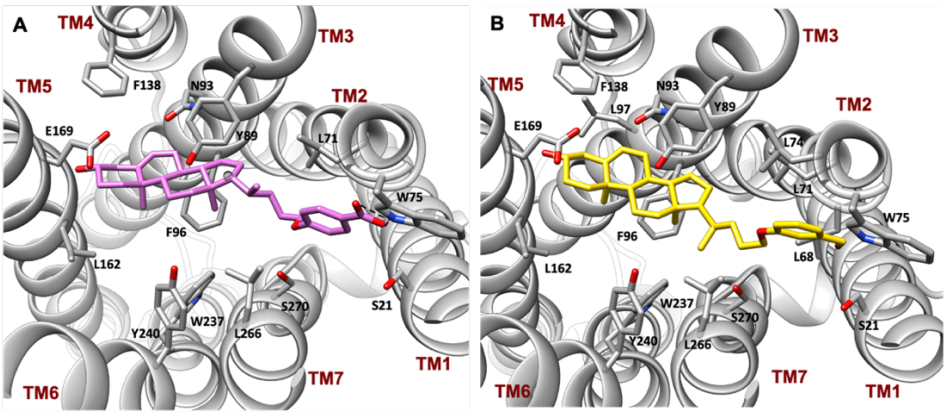


Figure 12. ***Binding mode of compounds (A) 3 and (B) 7 in GPBAR1.*** Ligands are represented as *pink* and *gold* sticks, respectively, whereas the interacting residues of the receptor are shown in *grey* and labelled. Oxygen atoms are depicted in *red* and nitrogens in *blue*. The receptors are represented as ribbons with their transmembrane helices (TM) labelled. Hydrogens are omitted for the sake of clarity and H-bonds are displayed as dashed *black* lines.

According to the literature data,[12] the mechanism of action of inverse agonist compounds against ROR-γ relies on the disruption of the H-bond between His479 and Tyr502, which stabilizes the active form of the receptor and the recruitment of the transcriptional coactivator SRC2. When an inverse agonist binds the receptor, it establishes a series of interactions with residues such as Leu324, Trp317, His479 and Tyr502. This causes a knock-on effect which stabilizes ROR-γ in the inactive form, therefore disallowing recruitment of the transcriptional coactivator and repressing target gene expression.

MD calculations. To further investigate the mechanism of action of these compounds on ROR-γ, classical MD calculations were carried out. In particular, each docking pose was subjected to 1 MD simulation in explicit solvent to deeply understand the molecular basis of ROR-γ inverse agonism and evaluate the structural stability of the two binding modes.

The result of 1 MD of ROR-*γ* shows that only one of the two docking poses is stable over the simulation time, as shown by the values of the average root mean square deviation (RMSD) computed for the ligand heavy atoms (Figures 13C and D). In particular, after a few ns of simulation compound **3** undergoes a conformational re-arrangement leading to an extended conformation of the flexible chain at C17 (-R, Figure 10), while the 3'-OH is still engaged in the H-bond with Gln286 side chain on H2. The steroidal scaffold is still placed in the amphipathic pocket between H4 and H5 where it is further stabilized by a set of hydrophobic interactions established with the side chains of Leu324, Val361, Val376 and Phe378. The aromatic ring is now located between H4 and H11, where it takes part in aromatic interaction with Trp317 and His479. The carboxyl group on the ligand side chain moves to form a H-bond with

Arg482 located on H11 (Figure 13A). Finally, this conformational change of compound **3** and, in particular, the stacking between the aromatic ring and the side chain of His479 produces a slight shift of H12 which alters the binding site of the co-activator (Figure 14A).

On the contrary, albeit during MD simulation compound **7** does not show any appreciable change of conformation, the MD analysis shows a change in its binding mode. In fact, at the end of the production run, the ligand binding mode is stabilized by different interactions: the 3'-OH group H-bonds Arg264, whereas the steroidal scaffold is stabilized by hydrophobic interaction with Leu324, Val361, Val376 and Phe388. The alkyl linker interacts with Leu391, Ile397 and Ile400, whereas the ethereal oxygen H-bonds Cys393. Finally, the aromatic ring is stabilized between H7 and H11 by aromatic interaction with Trp317, His479 and Phe486. A direct interaction of the side chain of compound **7** with Tyr502 causes a more marked weakening of the H-bond network between His479 and Tyr502 (Figure 13B), stabilizing the open form of H12 which hampers recruitment of the coactivator.

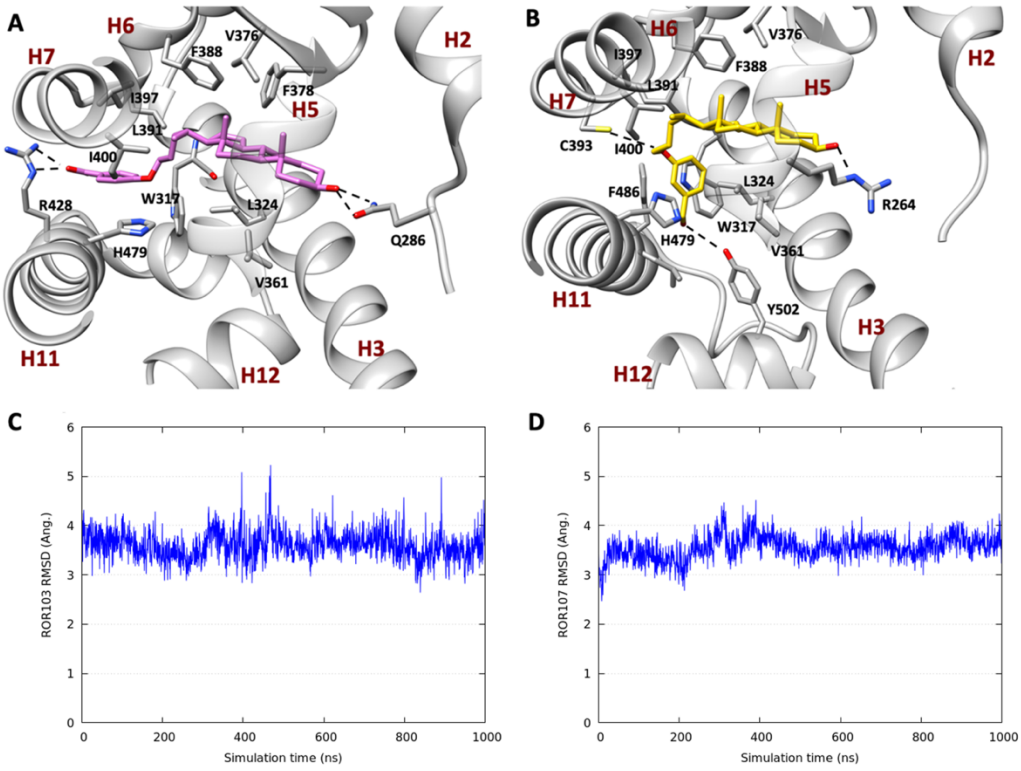


Figure 13. ***Molecular dynamics simulations of compounds 3 and 7 in ROR-γ.*** Representation of the last frame of MDs of (A) compound **3** and (B) compound **7** in the active conformation of ROR-γ. The ligands are represented as *pink* and *gold* sticks, respectively, whereas the interacting residues of the receptor are shown in *grey* and labelled. Oxygen atoms are depicted in *red* and nitrogens in *blue*. The receptors are represented as ribbons with their H labelled. Hydrogens are omitted for the sake of clarity and H-bonds are displayed as dashed *black* lines. Plots of average RMSD was calculated on the heavy atoms of (C) compound **3** and (D) compound **7**.

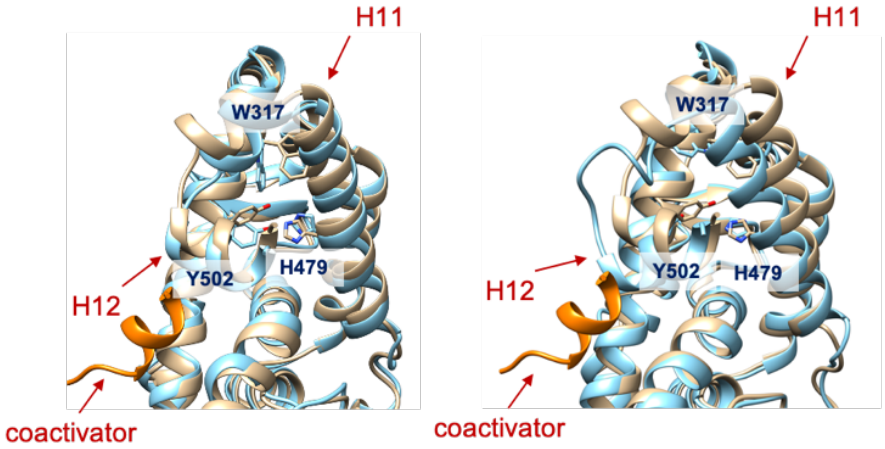


Figure 14. ***Compounds 3 and 7 mechanisms of action.*** Superimposition of the X-ray structure of ROR-γ (PDB ID 3l0j)[119] (*tan* cartoon) and the MDs last frame of A) compound **3** and B) compound **7** in complex with ROR-γ (*cyan* cartoon). The coactivator is represented in *orange* cartoon. Key residues are represented in sticks and labelled.

1 MD calculation is also performed starting from the docking poses of compounds **3** and **7** to GPBAR1 to fully take into account the receptor flexibility and, once again, to evaluate the structural stability of the two binding modes.

After 150 ns, compound **3** changes its binding mode resulting to reach a stable maintained until the end of the simulation, as reported in the RMSD plot (Figure 15C). The 3'-OH is still forming a H-bond with Glu169$^{5.42}$ backbone and the steroidal scaffold is located between TM3 and TM5, stabilized by hydrophobic interactions with Leu90$^{3.30}$, Phe96$^{3.36}$, Leu97$^{3.37}$, Leu174$^{5.47}$, Leu263$^{7.36}$, Leu266$^{7.39}$ and Trp267$^{7.40}$. The benzyl ring at C17 is also engaged in hydrophobic interactions with Leu18$^{1.36}$, Leu68$^{2.57}$ and Leu71$^{2.60}$ and it forms a t-shaped stacking with Trp63$^{2.52}$ (Figure 15A).

Compound **7** binding mode changes after 100ns of MD to achieve a stable binding mode, similar to the one achieved by compound **3** with differences in the interactions of the *meta*-alcohol benzylic group (Figure 15D). The binding mode is strongly stabilized by hydrophobic interactions with the side chain of some residue, such as Leu14$^{1.32}$, Trp63$^{2.52}$, Leu68$^{2.57}$, Leu71$^{2.60}$, Leu97$^{3.37}$, Leu174$^{5.47}$, Trp237$^{6.48}$, Leu263$^{7.36}$, Leu266$^{7.39}$, and by H-bonds between ligand's functional groups and GPBAR1 side chains. In particular, the 3'-OH makes a H-bond with the backbone of Glu169$^{5.42}$, the ethereal oxygen is engaged in H-bond with Ser21$^{1.39}$ and, finally, the alcohol benzylic hydrogen H-bonds Ser270$^{7.43}$ (Figure 15B).

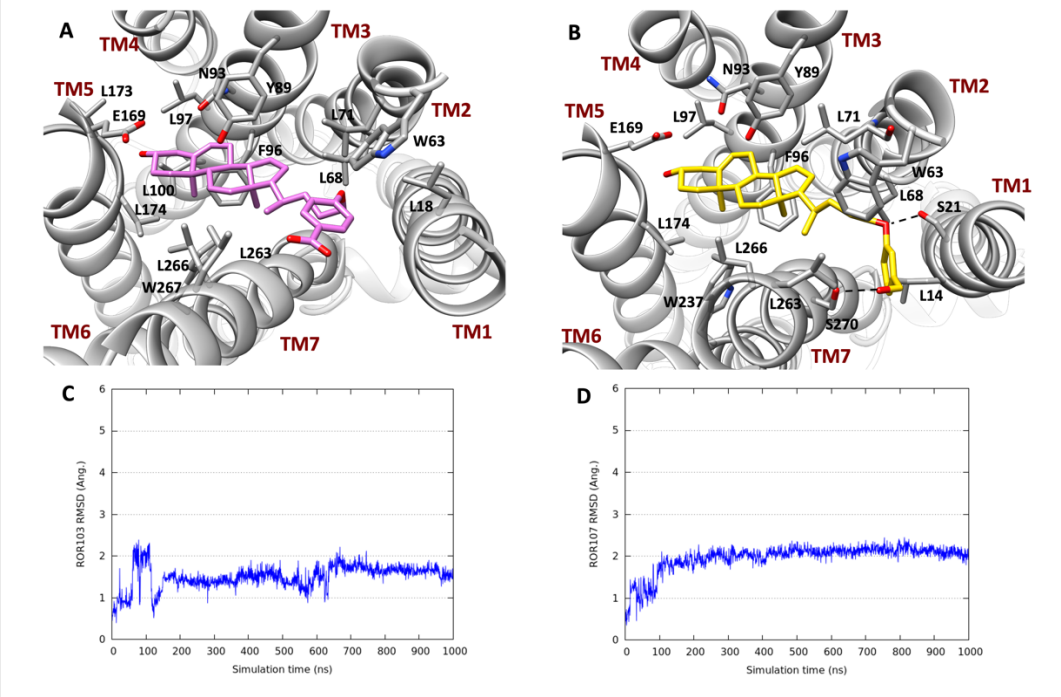


Figure 15. ***Molecular dynamics simulations of compounds 3 and 7 in GPBAR1 homology model.*** Representation of the last frame of MDs of compounds (A) **3** and (B) **7** in GPBAR1 homology model. The ligands are represented as *pink* and *gold* sticks, respectively, whereas the interacting residues of the receptor are shown in *grey* and labelled. Oxygen atoms are depicted in *red* and nitrogens in *blue*. The receptors are represented as ribbons with their H labelled. Hydrogens are omitted for the sake of clarity and H-bonds are displayed as dashed *black* lines. Plots of RMSD were calculated on the heavy atoms of (C) **3** and (D) **7.**

In conclusion, the substitution on C24 of the steroidal scaffold with a *para*- or *meta*-substituted phenoxy group led to the synthesis of compounds **1-12**.

Through combined *in vitro* and *in silico* studies, compounds **3** and **7** have been discovered as the most effective dual activity compounds so far identified, which showed promising activity for the treatment of colitis.

The first consideration to get is that *in vitro* assay reveals the *para*-hydroxyl substituted compound, **3** is more effective (IC_{50} 1.19 μM) than the *meta*-one, **7**, (IC_{50} 0.11 μM) in the modulation of ROR-*γ*. A similar result is also achieved in the capacity of such compounds to activate GPBAR1, with the *para*-substituted group having higher activity than the *meta* substituted one (**3** vs **7**) (EC_{50} 0.22 μM and 1.28 μM, respectively).

Moreover, the structure-activity relationship (SAR) analysis indicates that the introduction of sterically constrained functional groups (biphenyl) differently *meta*-substituted leads to good inverse agonist activity towards ROR-γ. In fact, although all the biphenyl derivatives are not active as agonists of GPBAR1, compounds **9** and **11**, the *meta* substituted biphenyl, show higher efficacy in inhibiting the ROR-γ receptor. This difference may be explained by the different capabilities of the LBD of these two receptors. The ROR-γ LBD is, in fact, better suited to accommodate bulky ligands, differently from GPBAR1 LBD. Furthermore, the pocket forming the GPBAR1 LBD, which hosts the ligands side chain represented by TM1 and TM2, is populated by different nonpolar residues, such as many Leucines.

2.2 Dual modulators against Cysteinyl Leukotriene Receptor 1 (CysLT$_1$R) and G-Protein coupled bile acid receptor 1 (GPBAR1)

Starting from the previous discovery of the multi-target profile of REV5901 to activate GPBAR1 and inhibit CysLT$_1$R,[43] was born the interest in exploring the quinoline scaffold of such compound aimed at the pharmacological development of new drugs, capable of simultaneously modulating CysLT$_1$R and GPBAR1 for potential therapeutic use in the treatment of colitis and other inflammatory processes. In this way, a new series of quinoline scaffold's derivatives have been developed, compounds **1-15**, starting from REV5901 as lead compound, which resulted in news CysLT$_1$R/GPBAR1 non-steroidal ligands (Figure 16).

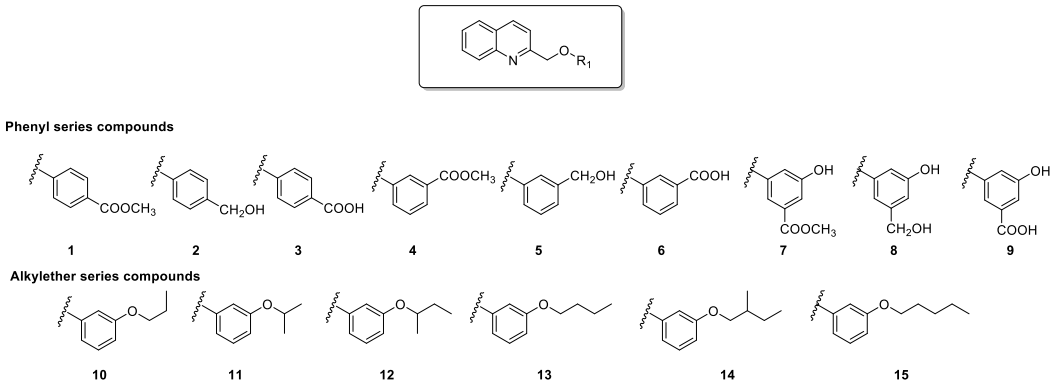


Figure 16. ***Chemical structure of derivatives 1–15 identified in this study.*** Credits for the figure go to Ref.[45]

2.2.1 Pharmacological activity

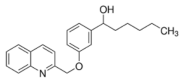
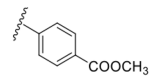
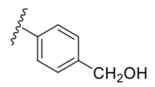
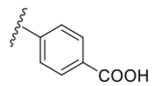
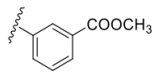
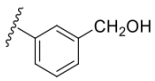
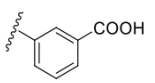
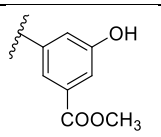
Compounds **1-15,** synthetized by the research group of Professor Angela Zampella at the Department of Pharmacy of University of Naples-Federico II, were, successively, tested for GPBAR1 agonistic activity in a luciferase reporter assay with GPBAR1 transfected HEK-293T cells, while evaluation of the compounds' antagonistic activity on the human CysLT$_1$R expressed in transfected CHO cells was

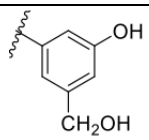
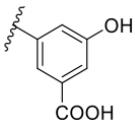
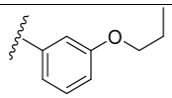
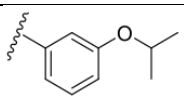
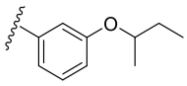
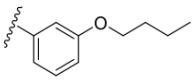
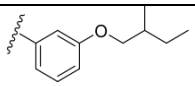
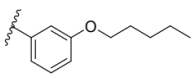
determined by measuring their effect on agonist-induced cytosolic Ca_2^+ ion mobilization using a fluorimetric detection method.[54]

The agonistic activity of the compounds on GPBAR1 was compared with that of the reference compound taurolithocolic acid (TLCA), which was fixed at 100%. Whereas the antagonist effect on $CysLT_1R$ was calculated as % inhibition of the control reference agonist response (LTD4) for each target.

As reported in Table 2, the best results in terms of efficacy and potency on receptors are found in **5** and **6**, as dual-activity compounds and in **14** as selective GPBAR1 agonist.

Table 2. ***Efficacy and Potency for compounds 1−15.***[f]

Compound	*Chemical structures*	*$CysLT_1R$*[a]	*IC_{50}*[b] *(μM)*	*GPBAR1*[c]	*EC_{50}*[d] *(μM)*
REV5901		116.80 ± 0.21	1.10 ± 0.50	136.74 ± 27.80	2.50 ± 1.20
1		48 ± 4.60	nd	20.70 ± 8.45	nd
2		85 ± 3.32	2.10 ± 1.50	12.63 ± 4.11	nd
3		-11 ± 4.38	nd	32.15 ± 3.29	16.50 ± 0.71
4		85 ± 0.91	3.90 ± 1.50	23 ± 4.76	nd
5		97 ± 0.78	1.20 ± 0.42	92.69 ± 0.73	7.40 ± 1.84
6		71 ± 2.05	2.80 ± 0.38	74.80 ± 3.96	3 ± 0.30
7		59 ± 3.25	nd	17.58 ± 1.37	nd

8		26 ± 14.30	nd	72.92 ± 1.57	23 ± 1.41
9		3 ± 7.42	nd	29.80 ± 0.89	20 ± 0.71
10		66 ± 4.66	5.11 ± 1.6	112.34 ± 12.21	1 ± 0.04
11		79 ± 5.09	nd	14.99 ± 1.81	nd
12		66 ± 0.07	9.63 ± 1.3	100.50 ± 13.91*	0.1 ± 0.05
13		4 ± 8.27	nd	138.88 ± 11.15	0.50 ± 0.22
14		15 ± 3.67	nd	106.43 ± 4.45	0.17 ± 0.07
15		22 ± 4.38	nd	137 ± 12.93	1.80 ± 0.07

[a,b]These assays were performed by Eurofins Cerep-Panlabs (France). The results are expressed as a percent inhibition of the control response to 0.1 nM LTD4. The standard reference antagonist is MK571. Results are mean of two experiments ±standard deviations (SD). [b]Results are mean of at least two experiments ±SD, and IC50 was determined for efficacy >60%. [c,d] These assays were performed by the research group of Professor Stefano Fiorucci at the Department of Medicine and Surgery of the University of Perugia (Perugia, Italy) [c]Eff (%) is the maximum efficacy of the compound (10 μM) relative to TLCA (10 μM) as 100 in transactivation of a cAMP-responsive element (CRE) on HEK-293T cells; results are mean of two experiments ±SD. [d]Results are mean of at least two experiments ±SD, and EC50 was determined for efficacy >25%. [e]Efficacy calculated with 1 μM of compound. [f]nd, not determined.

The anti-inflammatory activity of these compounds was investigated by *in vitro* assays. In this case, RAW264.7 mouse macrophages were primed with lipopolysaccharide (LPS) and co-incubated with or without REV5901, compounds **5**, **6** and **14** at 0.1, 1.5 and 10 μM, respectively. All compounds reduced the production

of LPS-induced proinflammatory cytokines (Tnfα and Il-1β), but only REV5901, **5** and **14** increased the expression of the anti-inflammatory Il-10 gene (Figure 17).

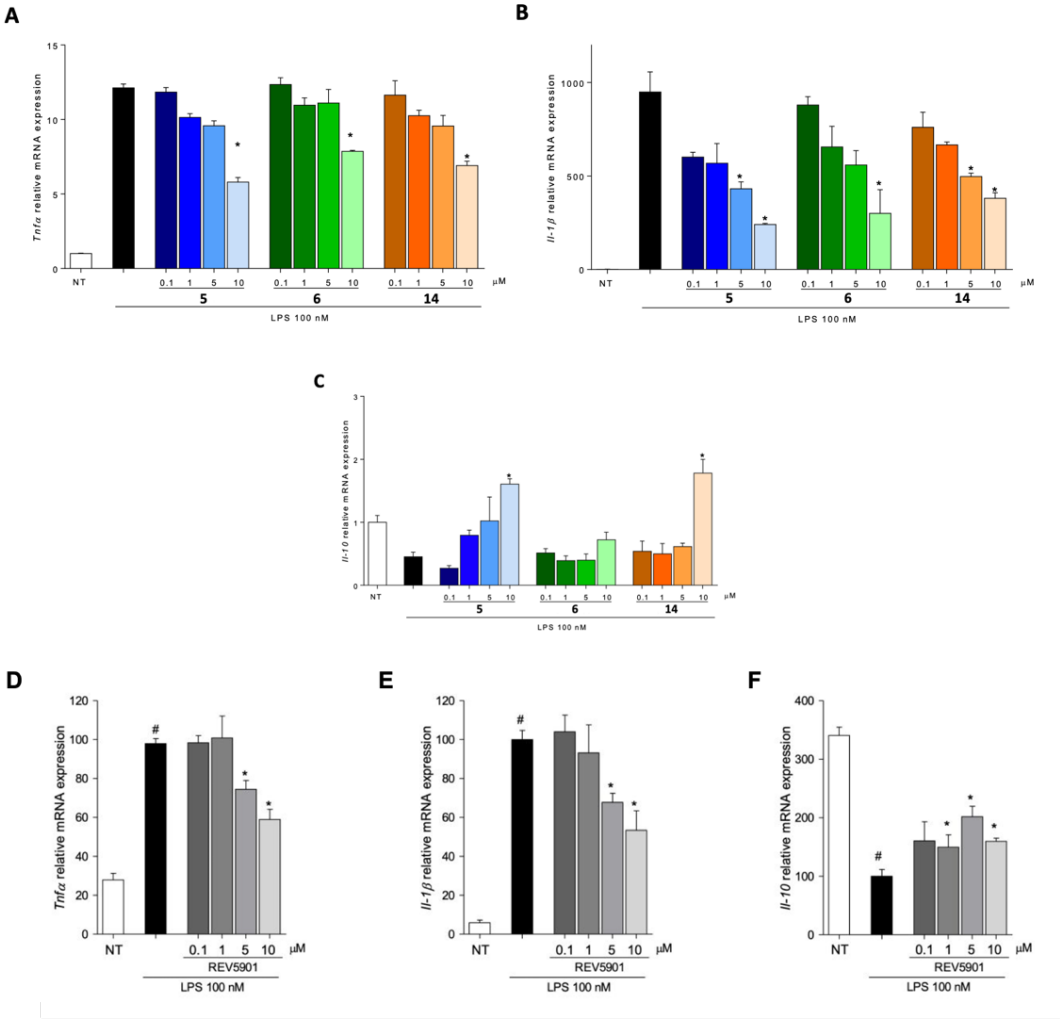


Figure 17. ***Cell proliferation assay.*** RAW2647. cells were classically activated with LPS (100 nM) and exposed or not to REV5901, compounds **5**, **6** or **14** at the concentration of 0.1, 1.5 and 10 μM for 16 h. Quantitative real-time PCR analysis of expression of pro-inflammatory genes Tnf-α (**A** and **D**) and Il-1β (**B** and **E**), and anti-inflammatory genes Il-10 (**C** and **F**). Data are derived from 6 replicates from 2 independent experiments. Results represent the mean ± SEM. *p < 0.05 vs LPS group. Credits for the figure go to Ref.[45]

The modulation of the proliferation of LTD4 induced by RAW264.7 for compounds **5**, **6** and **14** was also investigated using REV5901 as a control. As shown in Figure 18, LTD4 increases RAW264.7 cell proliferation, while **5**, **6** and **14**, like REV5901, reverse this effect, confirming their CysLTR1 antagonistic activity.

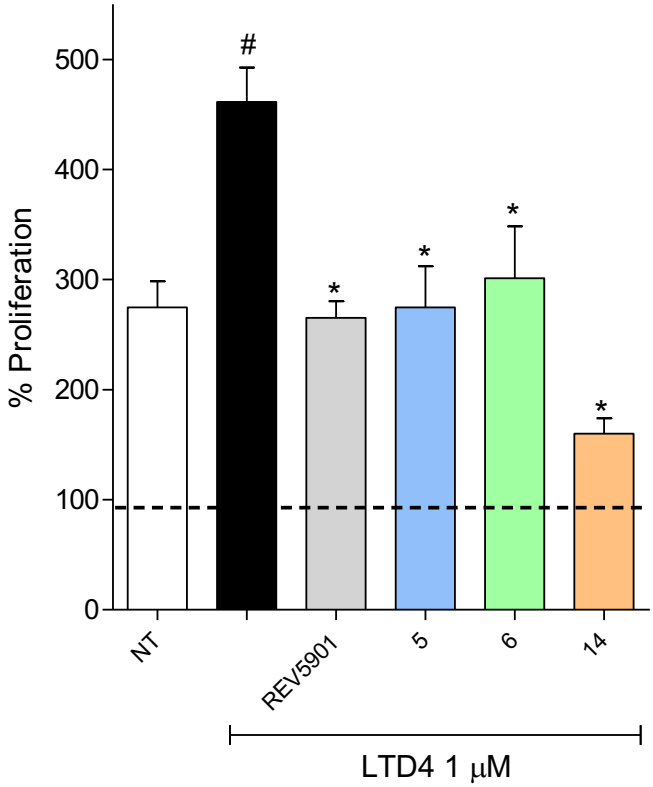


Figure 18. ***Cell proliferation assay.*** RAW264.7 were exposed to LTD4 (1 μM) for 48 h alone or in combination with compounds at concentration of 10 μM. Cell counting was performed using Trypan Blu staining. Data are derived from 6 replicates from 2 independent experiments. Results represent the mean ± SEM. #$p < 0.05$ vs NT group and *$p < 0.05$ vs LTD4 group. Anova-way analysis of variance was used for statistical comparisons. Credits for the figure go to Ref.[45]

Based on previously published data regarding the attenuation of the influx of immune cells from the circulation into inflamed tissues induced by the activation of GPBAR1[40,127,128,129] and considering that leukotrienes play an important role in the inflammatory process,[130,131,132,133] the efficacy of the new compounds was tested in an adhesion test, using human aortic endothelial cells (HAEC cells) and a human monocytic cell line (U937 cells) (Figure 19).

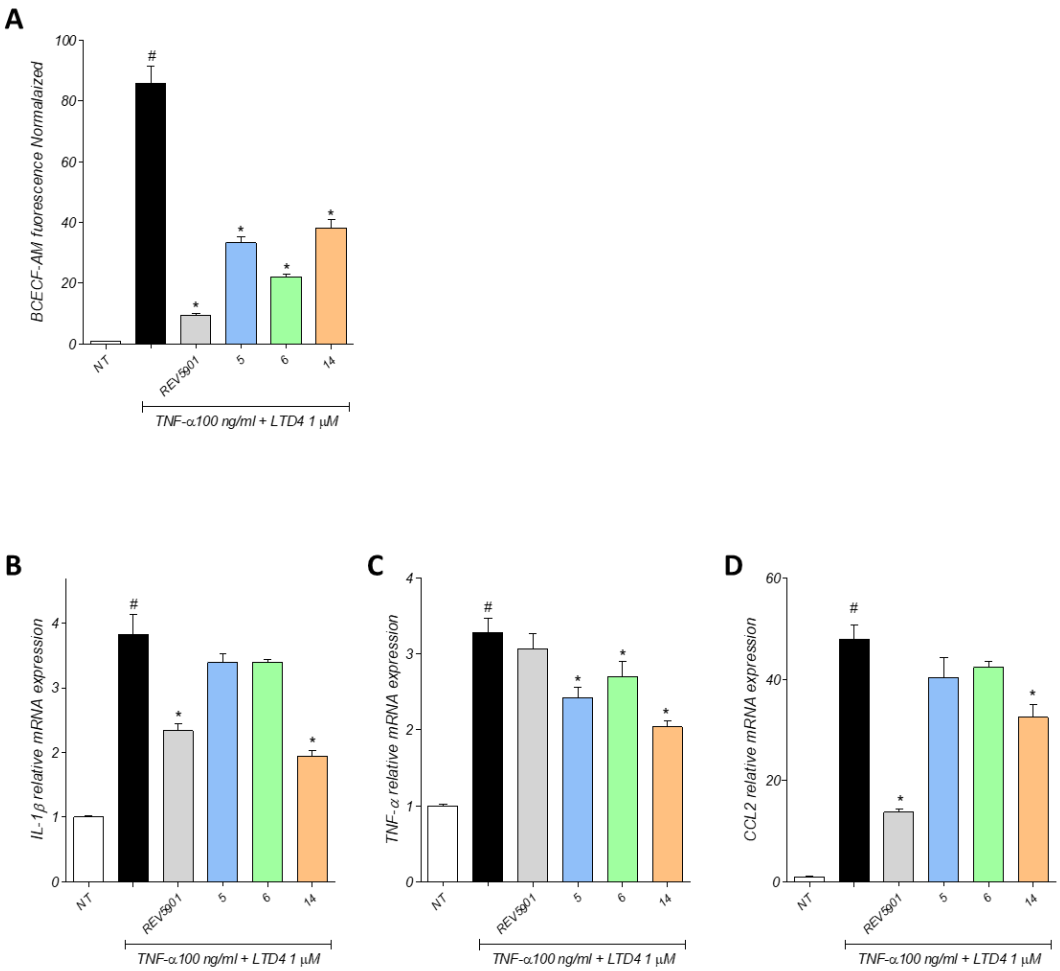


Figure 19. ***Cell adhesion assay.*** HAEC cells were plated on a 24-well plate and activated with TNFα (100 ng/mL) and LTD4 (1 μM) for 24 h alone or in combination with compounds **5**, **6**, **14** and REV5901 at 10 μM. U937 cells were treated under the same conditions. (**A**) For adhesion assay, U937 cells were fluorescently labelled with BCECF-AM and were incubated for 120 min with HAEC cells. Nonadherent monocytes were removed by gentle washing and fluorescence intensity was measured (485-nm excitation and 520–560-nm emission) using a microplate reader. (**B**) Quantitative real-time PCR analysis of expression of pro-inflammatory genes Il-1β (**B**) and Tnf-α (**C**) and chemokine Ccl2 (**D**) in U937 cells. These data are normalized to Gapdh mRNA. Data are derived from 8 replicates from 2 independent experiments. Results represent the mean ± SEM. #$p < 0.05$ vs NT group and *$p < 0.05$ vs TNF-α + LTD4 group. Anova-way analysis of variance was used for statistical comparisons. Credits for the figure go to Ref.[45]

Exposure of cells to TNF-α + LTD4 has been shown to increase the adhesion of monocytes to HAEC by approximately 100-fold, while REV5901 and all tested compounds **5**, **6** and **14** significantly reduced adhesion of the cells. U937 cells to HAEC (Figure 19A). Furthermore, considering that cell-cell adhesion in this assay

determines the activation of monocytes, the expression of additional mediators, such as Il-1β and TNF-α and CCL2 produced by U937 cells, was evaluated (Figures 19B-D), demonstrating that if cells are exposed to TNF-α + LTD4, the expression of these mediators increases and, furthermore, all tested compounds can reverse this effect. In particular, compounds **5** and **6** reduce TNF-α expression to a greater extent than REV5901 (Figure 19C) but are less effective on Il-1β and CCL2 (Figures 19B-D), while it is interesting to note that compound **14** which was found to be the most active compound in GPBAR1, it reduces the expression of these inflammatory genes to a greater extent (Figures 19B-D).

2.2.2 Computational studies

Molecular docking calculations were performed to investigate the binding modes of compounds **1-15** to $CysLT_1R$ and GPBAR1 using the Glide software package.[117,118] As regards the docking simulations in $CysLT_1R$, the crystallographic structure with PDB ID 6rz4[134] has been employed, while for GPBAR1, we used both the cryo-EM structures (PDB ID 7cfn and 7bw0)[135,136] and the 3D model developed *in house*[120] that has already been successfully employed by my research group in numerous drug design studies.[121,122,123]

Docking calculations of **1-15** in the GPBAR1 model showed remarkably convergent binding modes. The quinoline group is positioned in the amphipathic pocket between transmembrane helices (TM) TM3 and TM5, interacting with residues known to participate in GPBAR1 activation, like $Tyr89^{3.29}$, $Asn93^{3.33}$, $Phe96^{3.36}$ and $Trp237^{6.48}$ (superscripts refer to Ballesteros-Weinstein numbering)[137] (Figure 20A).[43,120,121,122,123,125] On the contrary, docking calculations in the cryo-EM structures did not lead to convergent results, therefore, they were not further considered in the study.

In $CysLT_1R$, the quinoline fraction of compounds **1-15** is placed in the pocket formed by TM3, TM4 and TM5, directed towards the lateral entrance of the receptor embedded in the bilayer (Figure 20B). It is important to note that the docking results show a propensity for the quinoline scaffold to bind both $CysLT_1R$ and GPBAR1 promiscuously. Therefore, on the basis of pharmacological assays, we decided to investigate, through more advanced methods, the binding mode of compounds **5** and **6**, the most powerful double ligands $CysLT_1R$/GPBAR1 and of compound **14**, which is a selective GPBAR1 ligand, with the aim of clarifying the structural basis for the double modulation of $CysLT_1R$/GPBAR1.

Binding mode of 5 in GPBAR1. Compound **5** results the most potent dual-target ligand of the series by pharmacological assays. The most populated binding mode in GPBAR1 shows the quinoline scaffold anchored between TM3 and TM5, forming a H-bond with $Asn93^{3.33}$ and hydrophobic interactions with $Phe96^{3.36}$, $Leu97^{3.37}$, $Leu100^{3.40}$, $Leu173^{5.46}$ and $Leu174^{5.47}$ (Figure 20A). Furthermore, the phenyl group makes polar and apolar contacts with $Leu71^{2.60}$, $Tyr89^{3.29}$, $Pro92^{3.32}$, $Glu169^{5.42}$, $Trp237^{6.48}$ and $Leu266^{7.39}$. Finally, the methyl-hydroxyl group extends towards TM1 and TM7, pointing towards $Leu68^{2.57}$ and forming an additional H-bond interaction with $Ser270^{7.43}$ (Figure 20A). The binding mode of the docking has been validated using atomic MD calculations which take into account the flexibility of the receptor and the solvent effect, which are important factors in the study of the ligand/receptor interaction and typically overlooked by the docking calculations.[138]

Along 1 μs of the MD simulation, the binding mode - very similar to the docking pose - that results stable and preserved until the end of the calculation, as shown by the values of the root mean square deviation (RMSD) computed on the ligand heavy atoms (Figure 22C). In particular, the binding mode represented by the cluster centroid

in figure 22A shows the quinoline scaffold located between TM3 and TM5, where it interacts with residues such as Tyr89$^{3.29}$, Leu97$^{3.37}$, Glu169$^{5.43}$, Leu173$^{5.46}$ and Leu244$^{6.55}$. Similarly to the docking pose, the quinoline moiety H-bonds with Ans93$^{3.33}$, whereas it forms a π-π stacking interaction with Phe96$^{3.36}$. The phenyl group of **5** engages a T-shaped π stacking interaction with Trp237$^{6.48}$ and, pointing towards TM1, TM2 and TM7, it also interacts with Leu68$^{2.57}$, Leu71$^{2.60}$, Thr74$^{2.63}$, Pro92$^{3.32}$ and Leu266$^{7.39}$. This pose is further stabilized by the two H-bonds formed between the ligand terminal hydroxyl group and Ser270$^{7.43}$, and by the ligand ethereal oxygen and Tyr240$^{6.51}$.

Binding mode of 5 in CysLT$_1$R. In CysLT$_1$R, the quinoline group of **5** is oriented orthogonal to TM3 and TM5, located in the pocket formed by Tyr108$^{3.37}$, Ser155$^{4.57}$, Phe158$^{4.60}$, Val186$^{5.35}$, Ser193$^{5.42}$ and Leu257$^{6.59}$ (Figure 20B). The ligand binding mode is stabilized by the cation-π interaction formed by the quinoline moiety with Arg253$^{6.55}$ and the H-bond between the ligand's ethereal oxygen and Tyr104$^{3.33}$, which, together with Tyr249$^{6.51}$, also form a T-shaped stacking interaction with the phenyl ring of **5**.

It is important to underline the similarity between the binding mode of **5** and the crystallographic binding pose of the known CysLT$_1$R antagonist pranlukast (Figure 21).[134] In fact, it is possible to note that pranlukast interacts through its tetrazole group with Arg79$^{2.60}$ via a direct and a water-mediated interaction. This interaction is emulated by the terminal hydroxyl group of compound **5** pointing towards Arg79$^{2.60}$, Val277$^{7.35}$ and Leu281$^{7.39}$ (Figure 20B). Considering that water molecules are not explicitly considered in docking simulations, MD and free energy calculations were performed to investigate the role of the solvent and the conformational flexibility of the ligand and receptor.

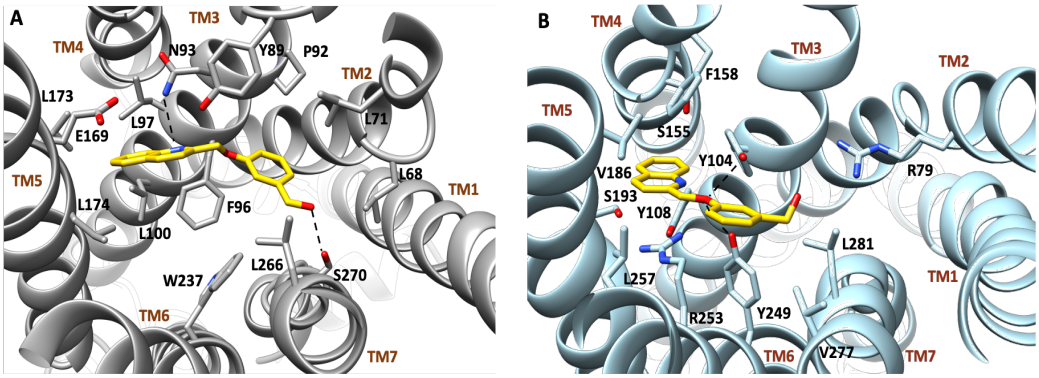


Figure 20. ***Binding modes of 5 in (A) GPBAR1 and (B) CysLT$_1$R identified via docking calculations.*** The ligand is represented as *gold* sticks, whereas the interacting residues of the receptors are shown in *grey* (GPBAR1) or *cyan* (CysLT$_1$R) and labelled. Oxygen atoms are depicted in *red* and nitrogen atoms in *blue*. The receptors are represented as *grey* (GPBAR1) or *cyan* (CysLT$_1$R) ribbons with their TMs labelled. Hydrogens are omitted for the sake of clarity and H-bonds are displayed as *black* dashed lines. Credits for the figure go to Ref.[45]

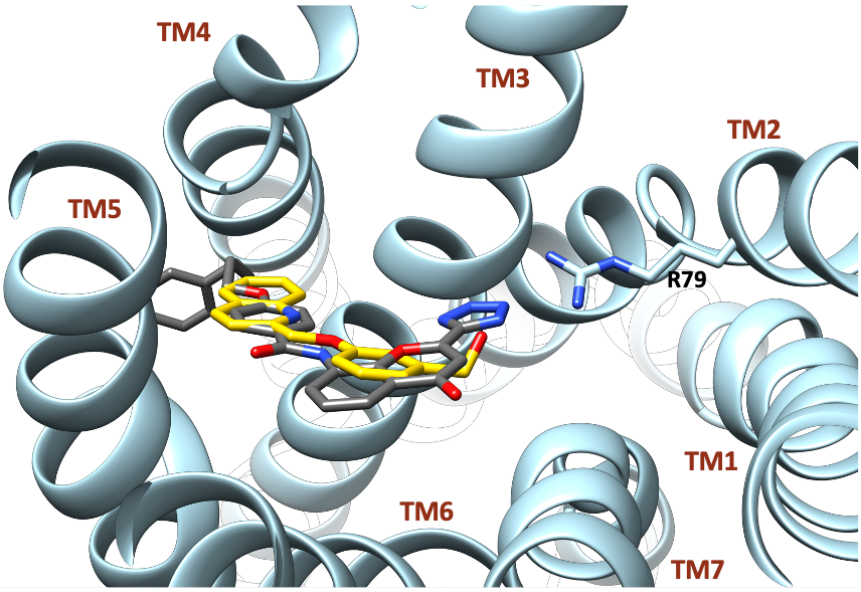


Figure 21. ***Comparison between the binding mode of 5 (gold sticks) and pranlukast (grey sticks) to CysLT$_1$R.*** Oxygen atoms are depicted in *red* and nitrogen atoms in *blue*. The receptor is represented as cyan ribbons with its TM and H labelled. Arg79$^{2.60}$ is shown in *cyan* sticks. Hydrogens are omitted for the sake of clarity. Credits for the figure go to Ref.[45]

During 1 μs MD simulation, the binding mode of compound **5** in CysLT$_1$R is stable up to 900 ns. However, the last 100 ns shows a slight flexibility in the methyl-hydroxyl

terminal group as shown in the RMSD plot (Figure 22D), allowing a water molecule to mediate optimally a H-bond between **5** hydroxyl group and Arg79$^{2.60}$, as shown in the binding mode of the centroid of the most populated cluster retrieved from the MD simulation (Figure 22B). Here, the quinoline scaffold is placed between TM4 and TM5, where it forms contact with Tyr108$^{3.37}$, Pro157$^{4.59}$, Val186$^{5.35}$, Leu189$^{5.38}$ and a T-shaped π stacking interaction with Phe158$^{4.60}$. On the other hand, the phenyl ring interacts with Tyr104$^{3.33}$, Leu257$^{6.59}$ and engages in a cation- π interaction with Arg253$^{6.55}$. Differently from the binding mode obtained by the MD in GPBAR1 which was found to be comparable to that of the docking calculations, at this time, we note that the MD binding mode slightly differs from the docking pose, particularly in the positioning of the quinoline ring between TM4 and TM5 and the reorientation of the methyl-hydroxyl tail that makes room for the water molecule mediating the H-bond interaction with Arg79$^{2.60}$ (Figures 20B and 22B). As already highlighted above, a similar interaction can be observed in the binding mode of the antagonist pranlukast (Figure 21).

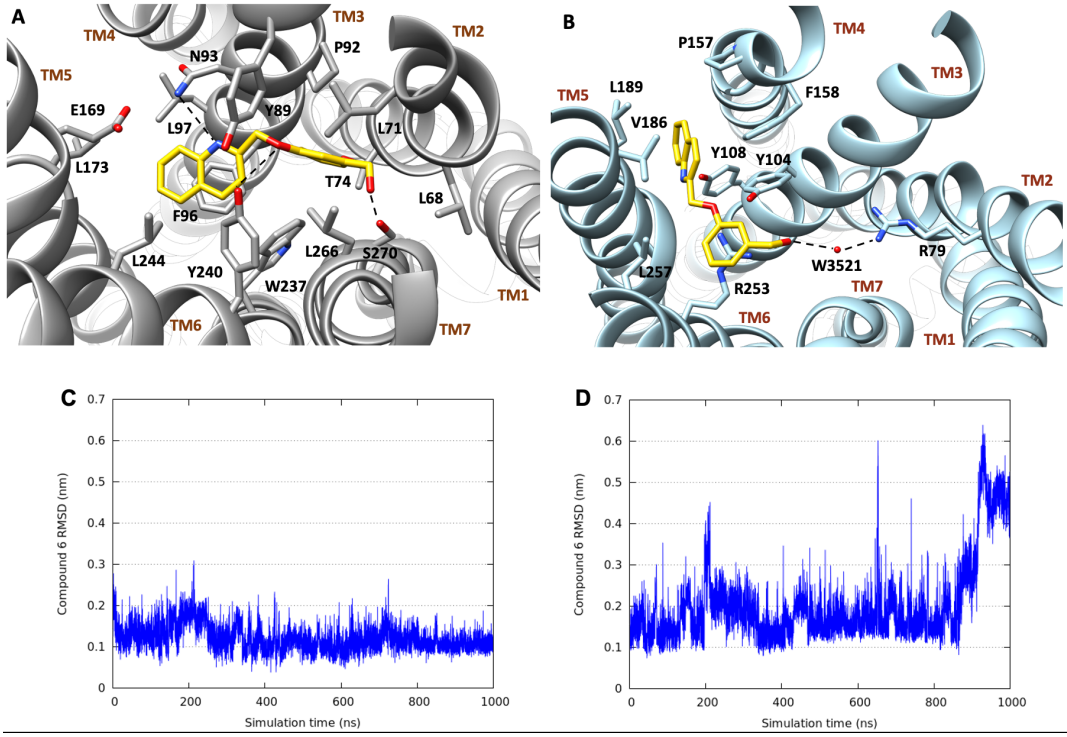


Figure 22. ***Results from compound 5 molecular dynamics calculations.*** **A-B**) Centroids of the most populated clusters of **5** in (**A**) GPBAR1 and (**B**) $CysLT_1R$ MD simulations. The ligand is represented as *gold* sticks, whereas the interacting residues of the receptor are shown in *grey* (GPBAR1) or *cyan* ($CysLT_1R$) and labelled. Oxygen atoms are depicted in *red* and nitrogen atoms in *blue*. The receptors are represented as *grey* (GPBAR1) or *cyan* ($CysLT_1R$) ribbons with their TMs labelled. Hydrogen atoms are omitted for the sake of clarity and H-bonds and salt bridges are displayed as *black* dashed lines; **C-D**) Average RMSD of the heavy atoms of **5** in GPBAR1 (**C**) and $CysLT_1R$ (**D**) along the MD simulations. Prior to the RMSD calculations, trajectory frames were aligned on the same atoms Credits for the figure go to Ref.[45]

Free-energy calculations. Starting from the difference between MD and docking poses, in order to deeply investigate the binding mode of compound **5** in $CysLT_1R$, we performed MetaD calculations,[112,113] which is an advanced technique used successfully by us and other research groups to disclose ligand binding mode in different DNA and protein systems including GPCR, based on free energy calculations.[138,139,140,141,142,143] This advanced computational technique consists in adding a Gaussian potential on user-defined degrees of freedom - named Collective Variables (CVs) - of the system under investigation. Thus, the system is thus allowed to explore every energetic minimum. At the end of the calculation, the most stable

binding mode of the ligand is identified at the lowest energy minimum. In this study, we defined as CV the distance between the center of mass of the ligand's quinoline ring and the C_{β} of Arg79$^{2.60}$, in order to identify the most stable binding mode of **5** to the $CysLT_1R$ binding pocket, founded in the lowest energy minimum at 1.8 nm. The ligand MetaD pose is strikingly similar to the MD one with a remarkably low RMSD value of 0.08 nm computed for the ligand heavy atoms (Figure 23B). Considering that during the MetaD simulation the ligand is free to explore all the possible binding modes in the binding pocket, the result obtained from the MetaD confirms the MD binding mode.

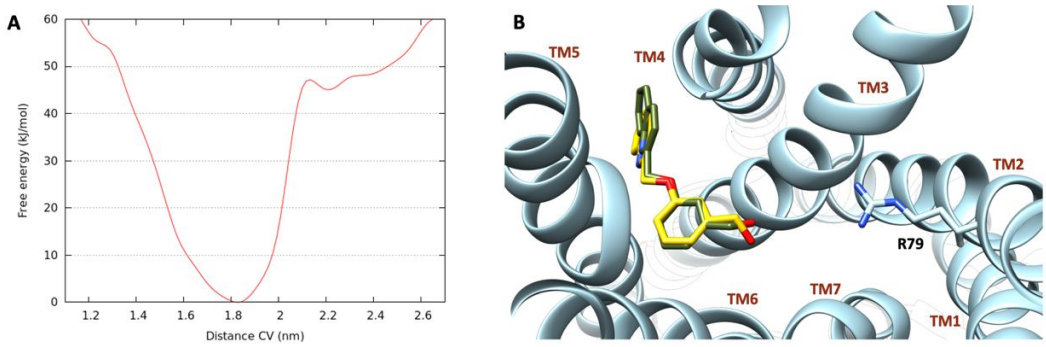


Figure 23. ***Results from compound 5 MetaD calculations in CysLT$_1$R.*** **A**) Free energy profile of **5** in $CysLT_1R$ binding pocket; **B**) Comparison between the energetically most stable binding pose obtained from MetaD (*gold* sticks), and the centroid of the MD most populated cluster (*dark green* sticks). Arg79$^{2.60}$ is shown as *cyan* sticks and labelled. Oxygen atoms are depicted in *red* and *nitrogen* atoms in blue. The receptor is represented as *cyan* ribbons with its TMs labelled. Hydrogens are omitted for the sake of clarity. Credits for the figure go to Ref.[45]

Since compound **6** is the second most powerful dual ligand in the series, also its binding mode in $CysLT_1R$ and GPBAR1 was studied.

Binding mode of 6 in GPBAR1. The protein conformation obtained from the MD simulation on the **5**/GPBAR1 complex was employed for the docking calculations, considering that, in this receptor state, the conformation of the binding site residues is already optimized to host compound **5**, which is structurally similar to **6**, thus

improving the reliability of the docking calculations. In the best-ranked – in terms of docking score and clusterization- docking pose, **6** interacts with GPBAR1 very similarly to **5** (Figure 20A). In particular, the quinoline scaffold is located in the amphipathic cleft formed by TM3, TM5 and TM6 where it engages favorable interactions with residues Tyr89$^{3.29}$, Leu97$^{3.37}$, Glu169$^{5.43}$, Leu173$^{5.46}$, Tyr240$^{6.51}$, Val241$^{6.52}$, Leu244$^{6.55}$. As with **5**, the quinoline moiety H-bonds Asn93$^{3.33}$ and forms π-π stacking interactions with Phe96$^{3.36}$. Furthermore, the ligand's phenyl ring engages a T-shaped π stacking interaction with Trp237$^{6.48}$ and points towards TM2 and TM7 forming apolar interactions with Leu68$^{2.57}$, Leu71$^{2.60}$, Thr74$^{2.63}$, Pro92$^{3.32}$, and Leu266$^{7.39}$. Finally, the carboxyl group establishes a H-bond with Ser270$^{7.43}$ (Figure 24A).

Binding mode of 6 in CysLT$_1$R. In the same approach employed for docking calculations in GPBAR1, the protein conformation retrieved from the MD simulation on the **5**/CysLT$_1$R was used for docking of **6** to CysLT$_1$R. In CysLT$_1$R, compound **6** occupies the binding site similarly to **5**, however, emerging few differences. The quinoline ring is located close to TM4 and TM5, forming apolar contacts with residues like Phe158$^{4.60}$, Val186$^{5.35}$ and Leu257$^{6.59}$. Furthermore, **6** forms a T-shaped π and cation-π interaction with Tyr104$^{3.33}$ and Arg253$^{6.55}$, respectively. At variance with **5**, compound **6** points the phenyl group towards TM3 and TM2 contacting with Leu103$^{3.32}$ and Leu281$^{7.39}$, with the carboxyl group interacting through salt bridge with Arg79$^{2.60}$ (Figure 24B).

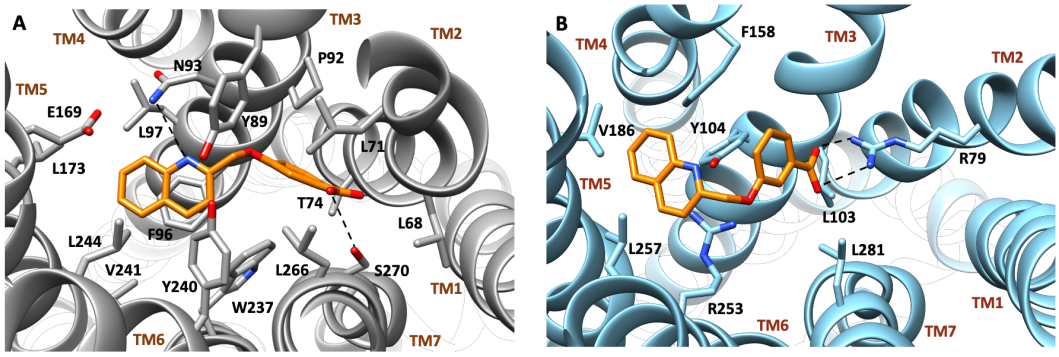


Figure 24. ***Binding mode of 6 in (A) GPBAR1 and (B) CysLT$_1$R identified via docking calculations.*** The ligand is represented as *orange* sticks, whereas the interacting residues of the receptor are shown in *grey* (GPBAR1) or *cyan* (CysLT$_1$R) and labelled. Oxygen atoms are depicted in *red* and nitrogen atoms in *blue*. The receptors are represented as *grey* (GPBAR1) or *cyan* (CysLT$_1$R) ribbons with their TMs labelled. Hydrogens are omitted for the sake of clarity and H-bonds and salt bridges are displayed as *black* dashed lines. Credits for the figure go to Ref.[45]

Binding mode of 14 in GPBAR1. Finally, we investigated the binding mechanism of **14**, the most potent selective GPBAR1 ligand synthetized in this study. As previously found for compounds **5** and **6**, the quinoline scaffold of **14** is placed between TM3 and TM5 where H-bonds Asn93$^{3.33}$ and establishes hydrophobic interactions with Leu97$^{3.37}$, Leu100$^{3.40}$, Leu173$^{5.46}$ and Leu174$^{5.47}$. The phenyl group interacts with Pro92$^{3.32}$ and makes a T-shaped π interaction with Phe96$^{3.36}$. Finally, the alkyl group extends towards TM1 and TM7 making additional hydrophobic contacts with Leu68$^{2.57}$, Leu71$^{2.60}$ and Leu266$^{7.39}$ (Figure 25).

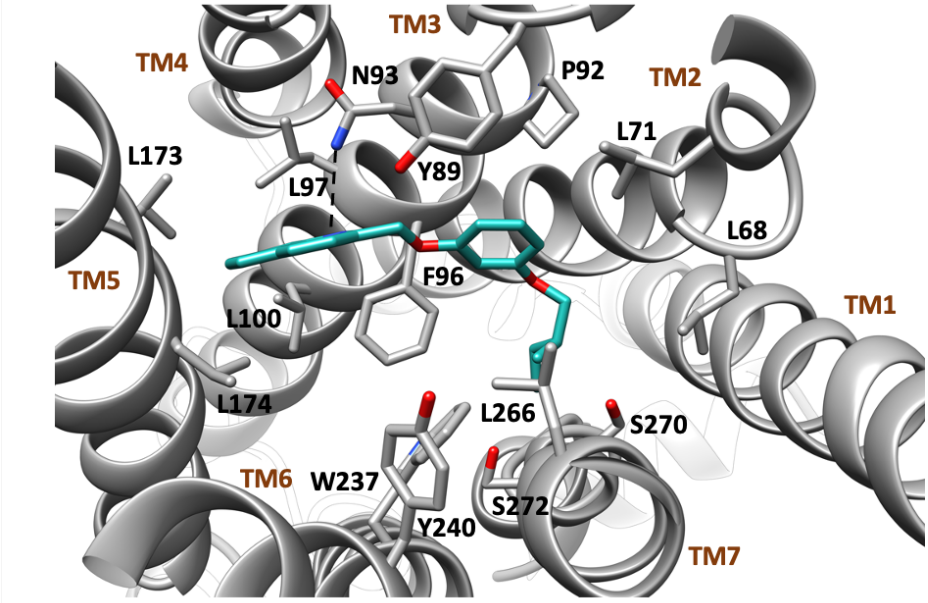


Figure 25. ***Binding mode of 14 in GPBAR1 identified via docking calculations.*** The ligand is represented as *light sea green* sticks, whereas the interacting residues of the receptor are shown in *grey* and labelled. Oxygen atoms are depicted in *red* and nitrogen atoms in *blue*. The receptor is represented as *grey* ribbons with its TMs labelled. Hydrogens are omitted for the sake of clarity and H-bonds are displayed as black dashed lines. Credits for the figure go to Ref.[45]

It is of considerable importance to note that the binding poses of **5**, **6** and **14** confirm the data previously reported for REV5901 (Figure 26).[43] In particular, the interactions with residues known to play an important role in the activation of GPBAR1 are here conserved, such as the H-bond with Asn93$^{3.33}$ and the hydrophobic pattern involving Tyr89$^{3.29}$, Phe96$^{3.36}$ and Trp237$^{6.48}$ (Figure 26). [43,120,121,123,124,125]

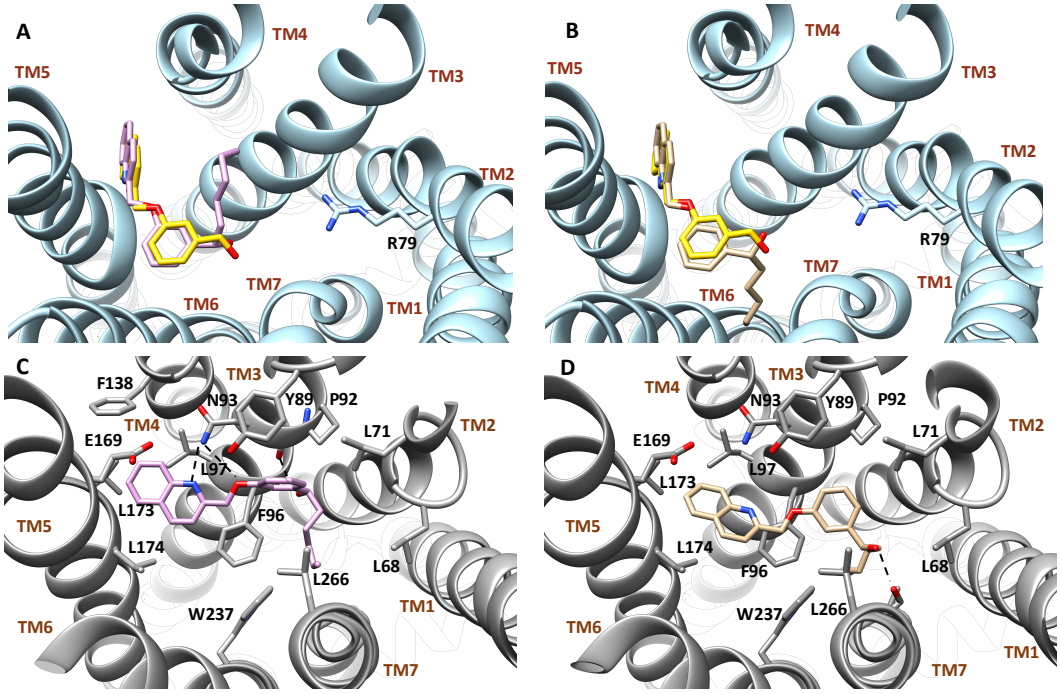


Figure 26. ***Binding mode of REV5901 to CysLT$_1$R and GPBAR1 identified via docking calculations.*** **A-B**) Comparison between the binding mode in CysLT$_1$R of **REV5901** in the S (A, *plum* stick) and R configuration (B, *tan* stick) and the binding mode of **compound 5** (*gold* sticks); **C-D**) Binding mode in GPBAR1 of **REV5901** in the R (C, *plum* stick) and S configuration (D, *tan* stick). Oxygen atoms are depicted in *red* and nitrogens in *blue*. The receptors are represented as *grey* (GPBAR1) or *cyan* (CysLT$_1$R) ribbons with their TMs labelled. CysLT$_1$R's Arg79$^{2.60}$ is shown in *cyan* sticks. Hydrogens are omitted for the sake of clarity. Credits for the figure go to Ref.[45]

Thus, it is possible to assert that in the series of compounds **1**-**9**, the methoxycarbonyl group is less effective than the carboxyl in activating GPBAR1 in both meta and para positions, while it is more active against CysLT$_1$R (**1**, **4** vs **3**, **6**). The hydroxymethyl group in para or meta shows the same efficacy and potency for CysLT$_1$R (**5** vs **2**), while, in GPBAR1, the efficacy significantly increases passing from para to meta (**5** vs **2**), up to the point of reversing the efficacy of the compounds. hydroxymethyl substituted versus carboxylic ones (**2** vs **3** and **5** vs **6**). The phenoxyquinoline scaffold substituted with a hydroxyl group in the meta position and several substitutions in the para position (**7**, **8**, **9**) showed a reduction in efficacy against CysLT$_1$R compared to mono-substituted compounds. Also in GPBAR1, **7** and **9**

showed a low efficacy, while a good efficacy was found, but a lower power for **8** compared to **5** and **6**.

In the alkyl ether series (**10**-**15**), it was noted that the increase in the length of the alkyl chain led to a reduction in efficacy against $CysLT_1R$ (**13**, **14**, **15** vs **10**, **11**, **12**), while no significant changes on GPBAR1, with the exception of **11**, which did not show good efficacy against the aforementioned receptor.

In conclusion, this study demonstrates that the presence of a polar or negatively charged group suitably spaced by an aromatic ring in the ligand represents a structural requirement to interact with $CysLT_1R$'s $Arg79^{2.60}$ and obtain the double activity of $CysLT_1R$/GPBAR1. Instead, the presence of a hydrophobic alkyl chain allows to obtain a selective activity on GPBAR1 on $CysLT_1R$.

Moreover, in this study, by exploiting a better synthetic accessibility, useful for future lead optimization studies, dual activity compounds were obtained and the binding modality to $CysLT_1R$ was rationalized for the first time through the combined application of experimental and *in silico* methodologies. This discovery may guide future studies for the identification of new drugs active on $CysLT_1R$ and GPBAR1.

This study has been published in Journal of Medicinal Chemistry (ref: Fiorillo, B.; Sepe, V.; Conflitti, P.; Roselli, R.; Biagioli, M.; Marchianò, S.; De Luca, P.; Baronissi, G.; Rapacciuolo, P.; Cassiano, C.; Catalanotti, B.; Zampella, A.; Limongelli, V.; Fiorucci, S. Structural Basis for Developing Multitarget Compounds Acting on Cysteinyl Leukotriene Receptor 1 and G-Protein-Coupled Bile Acid Receptor 1. *J Med Chem.* **2021**, *64*, 16512-16529.)

2.3 Natural and semi-synthetic steroidal agent acting on Spike protein Receptor Binding Domain (RBD) of SARS-CoV-2

The lack of the treatment for SARS-CoV-2 diseases has incited the scientific community to search for possible approaches and, in this regard, the goal of our study was to find small molecules, additional to those already discovered,[144,145,146,147,148,61,149] capable of binding residues at the interface between the SARS-CoV-2 S protein RBD and the ACE2 receptor to prevent the entry of SARS-CoV-2 into target cells. To this end, a VS campaign was conducted to identify compounds capable of interacting with Spike RBD and the binding modes of the most interesting compounds were validated by MD studies. Finally*, in vitro* binding assays were conducted to evaluate the ability of these compounds to inhibit the binding of protein S to the ACE2 receptor.

2.3.1 Computational studies

An FDA-approved drug library approved was screened using the RBD Cryo-EM 3D structure retrieved from the Protein Data Bank (PDB ID 6vsb; Chain A, residues Asn331-Ala520)[150] in order to find compounds capable of inhibiting the ACE2/Spike interaction by targeting the RBD of the S1 domain of SARS-CoV-2 (Figure 27). The protein in the prefusion conformation

Missing regions in the Cryo-EM structure were built using the SwissModel webserver.[151] Subsequently, a pocket search was performed on the prepared protein through the Fpocket web server.[152] The server's result recognized ≈ 300 putative pockets on the whole trimeric structure of the S protein. We, therefore, restricted the research to the identification of pockets within the RBD only on the basis of three main factors: (i) the potential druggability, through which it could be possible to interfere

directly or through an allosteric mechanism, with the interaction with ACE2; (ii) the flexibility degree of the pockets, i.e., excluding from the search of the pockets those defined, even partially, by highly flexible loops, whose coordinates were not defined in the experimental structure; (iii) sequence conservation with respect to SARS-CoV RBD (Figure 27A). Thus, six pockets were selected from the Fpocket and numbered according to the Fpocket ranking (Figures 27A and C).

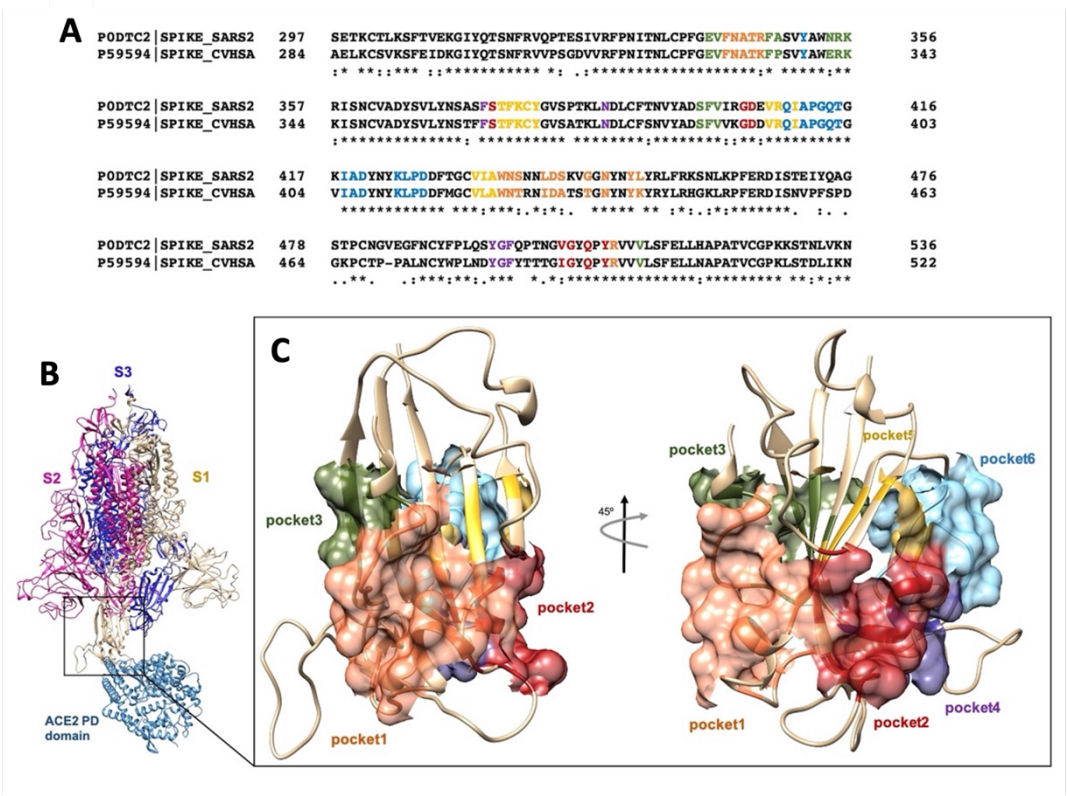


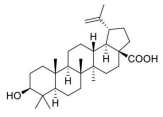
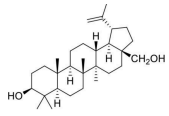
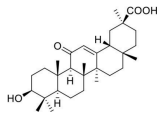
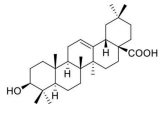
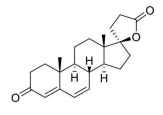
Figure 27. ***SARS-CoV-2 RBD pocket analysis.*** A) Clustal Omega alignment of RBD regions of SARS-CoV and SARS-CoV-2 Spike protein. Residues bearing to different pockets are colored respectively *yellow* (Pocket 1), *green* (Pocket 2), *light blue* (Pocket 3), *magenta* (Pocket 4), *red* (pocket 5), and *dark slate blue* (Pocket 6). B) Representation in cartoon of SARS-2 Spike protein in complex with the PD domain of ACE2. Complex obtained through the superposition of the PDB structures with PDB ID 6vsb and 6m0j, respectively.[150,67] C) Surface representation of the six selected pockets used for the screening. The Spike RBD is represented in *tan* cartoon. Credits for the figure go to Ref.[153]

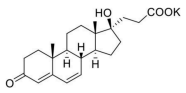
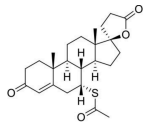
On these selected pockets 2,906 FDA-approved drugs retrieved from the DrugBank and the Selleckchem websites[154, 155] were subjected to a virtual screening using the AutoDock4.2.6 program and the Raccoon2 graphical user interface.[156,157] The binding

affinity was then studied on the basis of the ADscore of AutoDock4 software and selecting the results that showed an ADscore lower than -6 kcal/mol.

The study led to the identification of several compounds characterized by steroid and triterpenoid scaffolds, such as betulinic acid and the corresponding alcohol (betulin), glycyrrhetinic acid, canrenone and the corresponding open form on the γ-lactone ring as the potassium salt (canrenoate of potassium), spironolactone and oleanolic acid, showing a strong binding selectivity towards pocket 1 of the RBD (Table 3).

Table 3. ***Results of the screening of FDA approved drugs on the RBD region of the Spike protein of SARS-CoV-2 with the Autodock 4.2.6 program.***

Compound	*Chemical structures*	*ADscore (kcal/mol)*	*Pocket*
Betulinic acid		-8.1	1
Betulin		-7.4	1
Glycirrhetinic acid		-8.6	1
Oleanoic acid		-8.2	1
Canrenone		-7.9	1

Potassium canrenoate		-6.9	1
Spironolactone		-6.2	1

The pocket located on the β sheet in the central core of the RBD, pocket 1, is the least conserved of those examined. In fact, there are five conservative mutations (R346K, S438T, L440I, S442A) and two non-conservative (G445T and L451K) from SARS-CoV-2 to SARS-CoV.

Glycyrrhetinic acid turned out to be the best compound according to ADscore. The binding to the pocket is stabilized by both polar and not polar interactions. The triterpenoid scaffold is placed in the amphipathic pocket of the RBD, defined by Phe274, Trp436, Leu441 and Arg509. In addition, the binding mode is further stabilized by ionic contacts between the carboxyl group at C17 and the side chain of Arg509 and by two H-bonds: the first one between the carbonyl group at C11 of glycyrrhetinic acid and Asn440 and the second one between the 3-hydroxyl group and the side chain of Ser375. Different from the binding mode of glycyrrhetinic acid, betulinic acid and oleanolic acid shared a similar binding mode with the main difference in the carboxylic groups oriented toward the solvent. Finally, potassium canrenoate showed a different positioning of the steroid system within the binding site, having the A and B rings located between Trp436 and Leu441 and the carboxylic group H-bonding Ser375 (Figure 28).

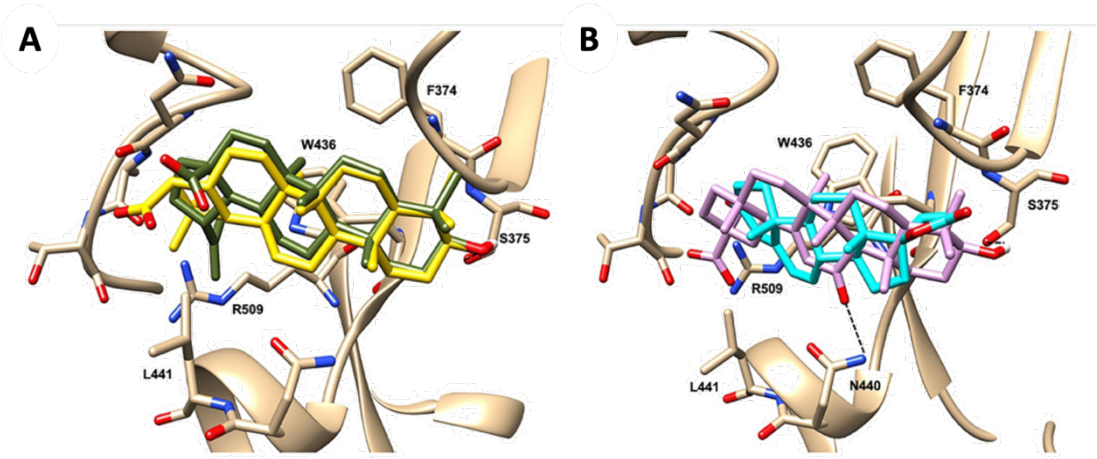


Figure 28. ***Graphical representation of the binding mode of the best compounds resulting from the screening in pocket 1.*** The RBD region is represented as tan cartoon. in transparent surface colored by residues hydrophobicity. Binding mode of: **(A)** Betulinic acid (*dark olive-green* stick) and oleanolic acid (*gold* stick). **(B)** Glycyrrhetinic acid (*plum* stick) and potassium canrenoate (*cyan* stick). For clarity reasons hydrogen atoms are omitted and only interacting aminoacids are displayed in sticks and labelled. Credits for the figure go to Ref.[153]

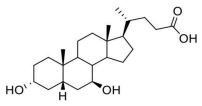
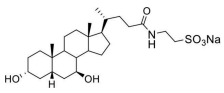
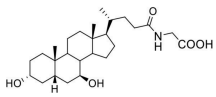
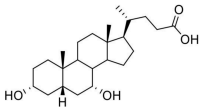
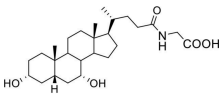
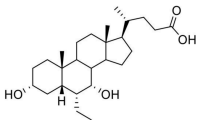
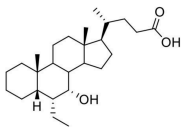
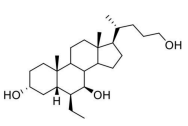
Furthermore, it is known in the literature that the aforementioned triterpenoids have been identified as natural ligands for two bile acid-activated receptors, FXR and GPBAR1.[1,158,159] In particular, oleanolic, betulinic and ursolic acids act as selective and potent agonists of GPBAR1, while glycyrrhetinic acid, the main metabolic component of licorice, and its corresponding saponin, glycyrrhizic acid, have been shown to act as dual agonists of FXR and GPBAR1 in the transactivation assay,[160] which also promote GLP-1 secretion in type 1 diabetic rats.[161]

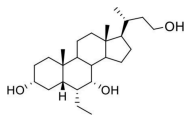
Due to the above, we further investigated whether other ligands of these receptors could bind the above mentioned RBD's pockets.

Taking into account the structural similarity between bile acids and these ligands and the ability to bind the same receptor systems, we performed a molecular docking study of natural bile acids and their semisynthetic derivatives currently available in

therapy or in preclinical and clinical development[159] and we tested their ability to bind the aforementioned pockets in the RBDs of the SARS-CoV-2 S protein (Table 4).

Table 4. ***Results of the screening of FDA approved drugs on the RBD region of the Spike protein of SARS-CoV-2 with the Autodock 4.2.6 program.***

Compound[a]	*Chemical structures*	*ADscore (kcal/mol)*	*Pocket*
UDCA		-7.0	5
T-UDCA		-7.0	5
G-UDCA		-7.3	5
CDCA		-7.3	5
G-CDCA		-7.6	5
OCA		-7.6	5
BAR704		-7.2	5
BAR501		-6.9	5

BAR502		-7.3	5

[a]Semi-synthetic compounds were synthetized by the research group of Professor Angela Zampella at the Department of Pharmacy of the University of Naples-Federico II

Natural bile acids and their semi-synthetic derivatives' docking poses are characterized by higher affinity ADscores when binding pocket 5 (Table 4 and Figures 29A–C), which is located in the central β-sheet core but on a different side than pocket 1. Comparing this pocket in SARS-CoV-2 and SARS-CoV, it was very conserved, showing only one mutation, I434L.

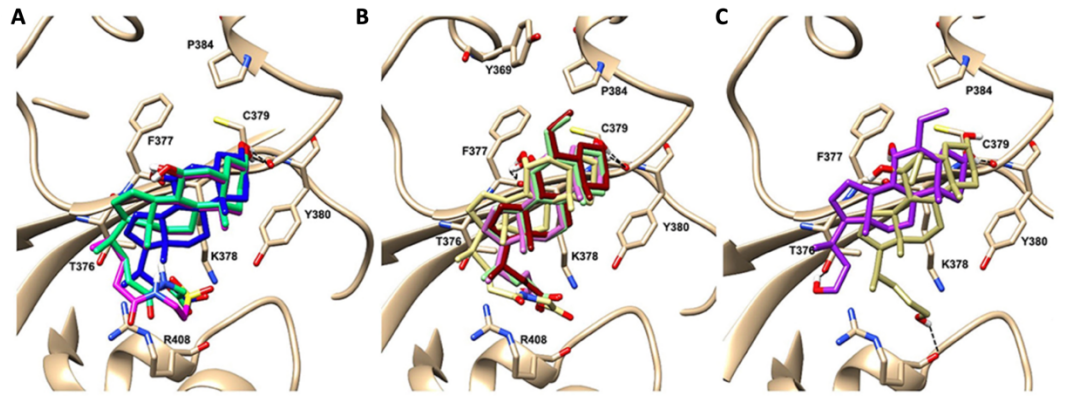


Figure 29. ***Graphical representation of the binding mode of the best compounds resulting from the screening in pocket 5.*** The RBD region is represented in *tan* cartoon. Binding mode of: (A) UDCA (*blue* stick), T-UDCA (*magenta* stick) and G-UDCA (*spring-green* stick); (B) CDCA (*orchid* stick), OCA (*light-green* stick), BAR704 (*dark-red* stick) and G-CDCA (*khaki* stick); (C) BAR501 (*gold* stick) and BAR502 (*purple* stick). For clarity reasons hydrogen atoms are omitted and only interacting aminoacids are displayed in sticks and labelled. Credits for the figure go to Ref.[153]

The UDCA binding mode shows the steroidal scaffold is placed in a hydrophobic pocket defined by Thr376, Phe377, Lys378, Tyr380 and Phe384. The carboxyl group

on the side chain points towards Lys378 and Arg408, while the 3β-hydroxyl group interacts with the Cys379 backbone.

The corresponding derivatives conjugated with glycine and taurine, G-UDCA and T-UDCA, respectively, showed the same interactions between their negatively charged side groups and Lys378 and Arg408, as well as the interaction between the 3β-hydroxyl group and the Cys379 backbone is maintained. The steroidal scaffold is slightly shifted towards Thr376 due to the increased side chain length. Furthermore, in the case of G-UDCA, the shift causes an additional π-interaction between the glycine amide region and the guanidine moiety of Arg408: this results in a better score for G-UDCA and a reduction in the case of T-UDCA, likely due to a lower interaction between the taurine fraction and the residues present within the binding pocket.

CDCA showed a similar binding mode, but, even more, stabilized by the presence of an additional H-bond with the backbone of Phe377 due to the modification in the configuration of the C7 hydroxyl group (α-oriented in CDCA and β-oriented in UDCA). Also G-CDCA established the same H-bonds network of CDCA, while the steroidal core slightly shifted as happened for G-UDCA. It is interesting to note that AD scores of G-UDCA and G-CDCA indicated that the H-bond between the hydroxyl group at C7 and Phe377 does not contribute significantly to the binding mode.

Instead, the introduction of the ethyl group in position C6 present in OCA and in BAR704 improves the internal energy of the ligand compared to CDCA (-0.27 for CDCA vs. -0.59 and -0.60 kcal/mol for OCA and BAR704, respectively), even if, albeit in close proximity of Pro384 and Tyr369, the 6-ethyl group did not show any particular interaction with residues within the RBD region (Figure 29B).

BAR501, a neutral UDCA derivative, having a hydroxyl terminal group and an β-ethyl group at C6, showed a very similar binding mode compared to UDCA, with the

side chain hydroxyl group H-bonding Arg408. Finally, BAR502, characterized by one less carbon on the side chain, presents the steroidal scaffold positioned in the same way as G-CDCA, thus allowing the C23 hydroxyl group H-bonding with the hydroxyl group of Thr376.

To further validate our hypothesis about the allosteric inhibitory potential of the identified pockets, we firstly performed 500 ns of MD simulations of the RBD domain and, then, a dynamical network and community map analysis on the obtained trajectory.

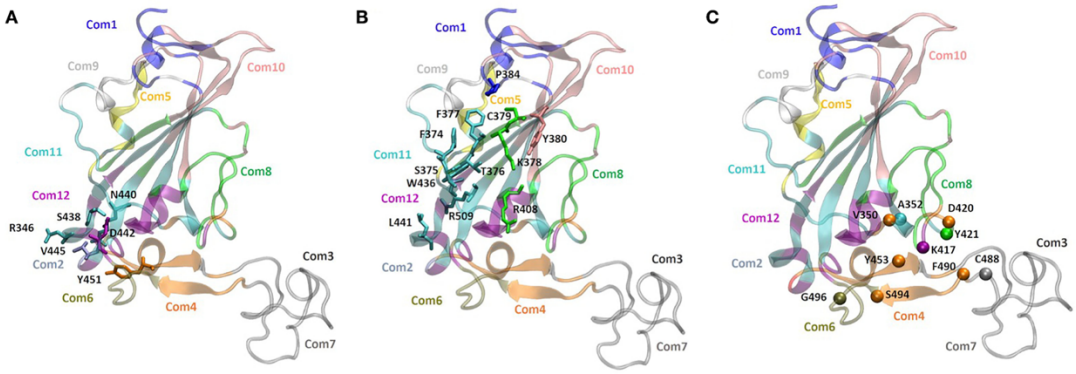


Figure 30. ***Dynamical network analysis.*** Community network representation of the RBD domain and community residue members of (A) pocket 1 (Arg346, Ser438, Asn440, Asp442, Val445, and Tyr451), (B) pocket 5 (Phe374, Ser375, Thr376, Phe377, Lys378, Cys379, Tyr380, Pro384, Arg408, Trp436, Leu441, and Arg509). (C) Highest score edge connectivity residues retrieved on the basis of the betweenness matrix. Spheres indicate the Cα atoms of residues that occur in a majority of shortest paths connecting nodes in different communities. Credits for the figure go to Ref.[153]

Table 5. ***Community map distribution of the RBD domain, retrieved after 500 ns-long MD simulation.***

Community	*N. of members*	*Residues*	*Color code cartoon*
Com1	14	N334; C361; V382; P384; T385; L387; D389; V524-K529;	Blue
Com2	1	V445	Ice-blue

Com3	1	G476	Dark-gray
Com4	18	V350; G416; D420; G446-R454; F456; F490-S494;	Orange
Com5	7	C336; E340; F342; A344	Yellow
Com6	9	Y495-G502; G504	Tan
Com7	33	L455;R457-A475; S477-Y489	Light-gray
Com8	20	T376; K378; C379; R408; I410-T415; I418; A419; Y421; Y423-P426; D428; T430; V511	Green
Com9	9	A363-Y369; S371; S383	White
Com10	30	L335; R355-N360; V362; Y380; G381; K386; L390-V395; D427-F429; L513-T523	Pink
Com11	38	V341; N343;T345; R346; Y351-N354; N370; A372-S375; F377; Y396-F400; N422; G431-A435; N437; N439-L441; S443; P507-V510; V512	Cyan
Com12	16	V401-V407; Q409; K417; W436; S438; D442; K444; V503; Y505; Q506	Purple

The network analysis found 12 communities (Com1-Com12) (Figure 30 and Table 5). In each community, there are residues of the RBD whose movement influences that of the other. The nodes correspond to the Cα atoms. Those who belong to the same community are highly interconnected. However, nodes defined as "critical" can also indicate connections between different communities characterized by a metric called betweenness (Figure 30C). From the network analysis we performed, 12 communities emerged, distributed as follows: the RBM region resulted in a split into three communities (Com4, Com6, and Com7), with Com4 including the short β-sheet, while Com6 and Com7 include residues of the binding loops Gly496-Tyr505 and Phe456-Phe490 (Table 5), respectively. Com11 included the majority of the residues of Pocket 1 and pocket 5 residues. Few residues are included in other communities: in particular, the residue Tyr451 of pocket 1 resulted in Com4 and residues Ser438 and Asp442 in Com12, while pocket 5 residues Thr376, Lys378, Cys379, Arg408 in Com8. Finally, Tyr380 was part of Com10. In order to highlight the hypothesis about the allosteric communication among the different communities, we analyzed the edge betweenness (Figure 30C), searching for the shortest paths between pairs of nodes belonging to two different communities. We found that communities including residues of pocket 1 and pocket 5 communicate, through Com4, with Com6 and Com7. In particular, Com8, Com10, Com11, and Com12, which include most of the residues present in pockets 1 and 5, were connected to Com4, which was strongly connected to Com6 and weakly to Com7, thus confirming a strong potential allosteric communication among the pockets at the receptor interface.

2.3.2 In vitro binding assays

In vitro binding assays were performed by the research group of Professor Stefano Fiorucci at the Department of Medicine and Surgery of the University of Perugia.

To evaluate if compounds of Table 3 can inhibit the binding of protein S to the ACE2 receptor, a Spike/ACE2 inhibitor screening test kit was employed, finding that, if Spike RBD is incubated with betulinic acid, glycyrrhetinic acid, oleanolic acid and potassium canrenoate, concentration-dependent reduction in the binding of S Spike RBD to the ACE2 receptor occurs. Furthermore, while glycyrrhetinic acid and potassium canreonate reversed binding at a concentration of 10 μM, betulinic acid and oleanolic acid showed significant inhibition at a concentration of 0.1 and 1 μM, respectively (n = 3 replicates) (Figure 31).

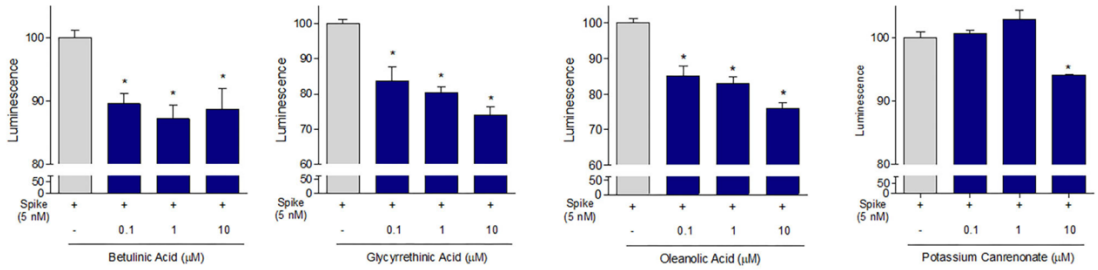


Figure 31. ***ACE2:SARS-CoV-2 Spike Inhibitor Screening assay.*** Betulinc acid, glycyrrethinic acid, oleanolic acid and potassium canrenonate were tested at different concentration (0.1, 1, and 10 μM), to evaluate their ability to inhibit the binding of Spike protein (5 nM) to immobilized ACE2. Luminescence was measured using a Fluo-Star Omega fluorescent microplate reader. Luminescence values of Spike 5 nM were arbitrarily set to 100%. Results are expressed as mean ± standard error. *$p < 0.05$ vs. Spike 5 nM. Data are the mean ± SE, n = 3. Credits for the figure go to Ref.[153]

The efficacy in reducing the SARS-CoV-2-ACE2 interaction of natural compounds of Table 4 was also tested, finding that UDCA and its taurine conjugate, T-UDCA, cause a mild dose-dependent inhibition of binding of RBD protein S to the ACE2 receptor (Figures 32A and B). In contrast, the major human metabolite of UDCA, G-UDCA, causes about 20% concentration-dependent inhibition of RBD binding to the

ACE2 receptor, like CDCA and its metabolite, G-CDCA (Figures 32C-E). Finally, the combination of UDCA and G-CDCA was studied and what we observed is a slight additive effect in inhibiting the interaction, confirming that UDCA itself has very limited inhibitory activity (Figure 32E).

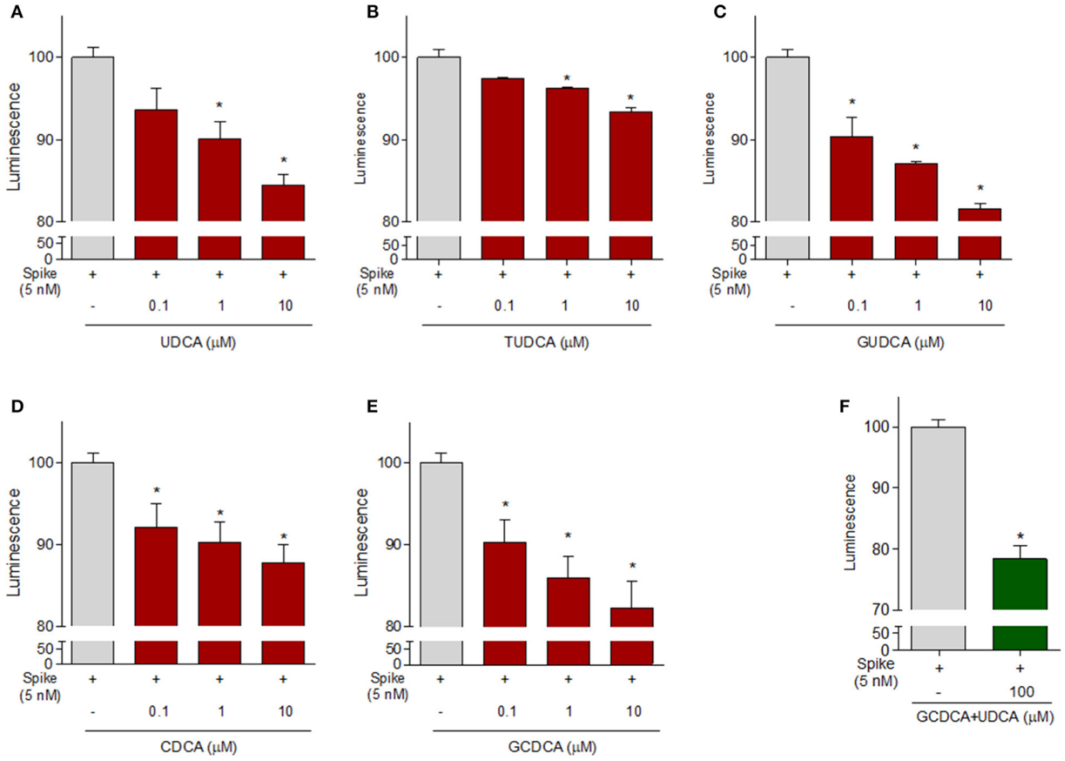


Figure 32. ***ACE2:SARS-CoV-2 Spike Inhibitor Screening assay.*** Natural bile acids (A) UDCA, (B) TUDCA, (C) GUDCA, (D) CDCA, (E) GCDCA (0.1, 1 and 10µM) and (F) a combination of GCDCA + UDCA (100µM), were tested to evaluate their ability to inhibit the binding of Spike protein (5 nM) to immobilized ACE2, by using the ACE2:SARS-CoV-2 Spike Inhibitor Screening assay Kit. Luminescence was measured using a Fluo-Star Omega fluorescent microplate reader. Luminescence values of Spike 5 nM were arbitrarily set to 100%. Results are expressed as mean ± standard error. *$p < 0.05$ vs. Spike 5 nM. Data are the mean ± SE, n = 3. Credits for the figure go to Ref.[153]

Subsequently, we investigated whether semisynthetic bile acid derivatives, OCA, BAR704, BAR501, and BAR502, possess comparable or superior inhibitory capabilities to G-CDCA. OCA decreased the Spike RBD-ACE2 interaction by about

20%, whereas, in contrast, BAR704, a derivative of CDCA, reduced binding by ~40% at the 10 µM dose. Finally, BAR501 and BAR502, derivatives of UDCA and CDCA, respectively, showed weak efficacy in reducing the binding of Spike RBD to ACE2 (Figure 33).

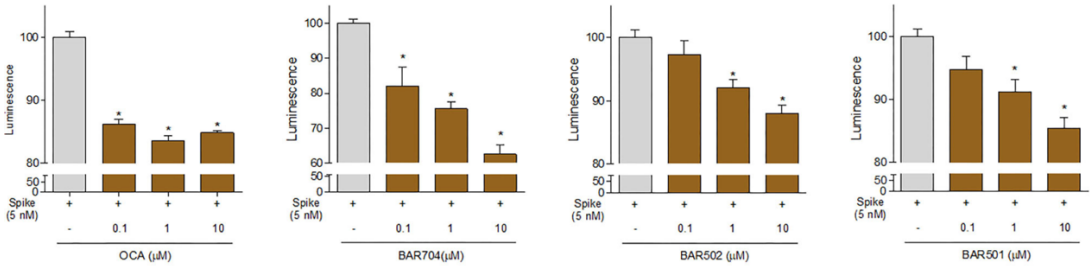


Figure 33. ***ACE2:SARS-CoV-2 Spike Inhibitor Screening assay***. The semi-synthetic bile acid receptor agonists OCA, BAR704, BAR502, and BAR501, were tested at different concentration (0.1, 1, and 10 µM) to evaluate their ability to inhibit the binding of Spike protein (5 nM) to immobilized ACE2, by using the ACE2:SARS-CoV-2 Spike Inhibitor Screening assay Kit. Luminescence was measured using a Fluo-Star Omega fluorescent microplate reader. Luminescence values of Spike 5 nM were arbitrarily set to 100%. Results are expressed as mean ± standard error. *$p < 0.05$ vs. Spike 5 nM. Data are the mean ± SE, n = 3. Credits for the figure go to Ref.[153]

Furthermore, by pre-incubating Spike RBD alone with 10 µM of the selected compound, many of them showed a greater ability to reduce the interaction between Spike and ACE2 compared to the standard incubation performed in the same experiment (Figure 34). In particular, oleanolic and glycyrrhetinic acid reduced the binding of Spike-RBD to ACE2 by 40% when pre-incubated with RBD, as did UDCA, T-UDCA, CDCA, G-CDCA, OCA, BAR502 (inhibitory power of ~ 45-50%), while BAR704 reduced the interaction up to 50%. Conversely, betulinic acid and potassium canrenoate showed no difference from previous results (Figure 34).

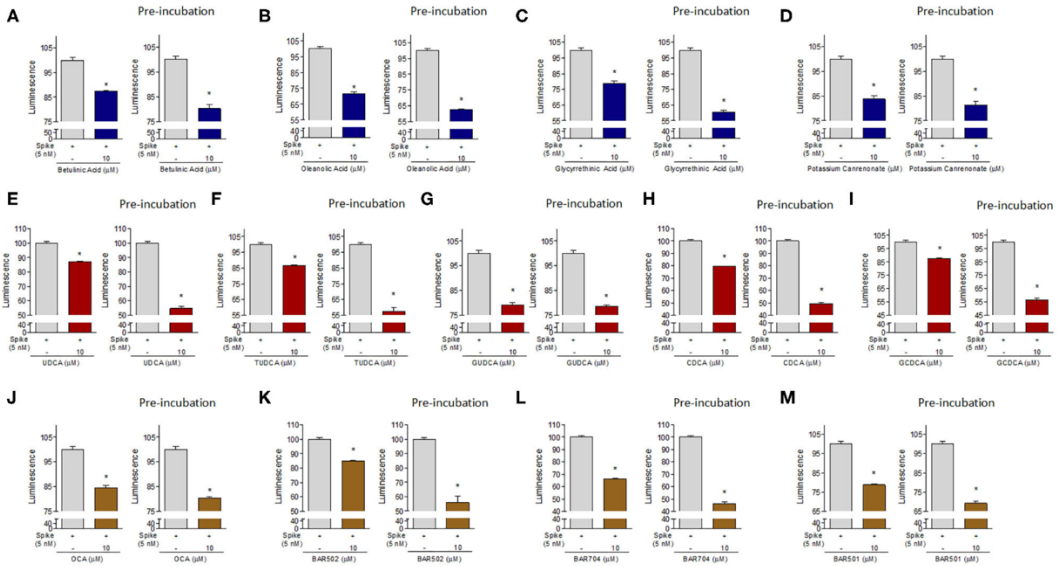


Figure 34. ***ACE2:SARS-CoV-2 Spike Inhibitor Screening assay.*** The selected compounds were tested at 10 µM to evaluate their ability to inhibit the binding of Spike protein (5 nM) to immobilized ACE2, according to the ACE2:SARS-CoV-2 Spike Inhibitor Screening assay Kit instructions or with a modified protocol in which we have performed a pre-incubation of these compounds with Spike-RBD (2 h). Tested compounds were: (A) Betulinic Acid, (B) Oleanolic Acid, (C) Glycyrrethinic Acid, (D) Potassium Canrenoate, (E) UDCA, (F) TUDCA, (G) GUDCA, (H) CDCA, (I) GCDCA, (J) OCA, (K) BAR502, (L) BAR704, (M) BAR501. Luminescence was measured using a Fluo-Star Omega fluorescent microplate reader. Luminescence values of Spike 5 nM were arbitrarily set to 100%. Results are expressed as mean ± standard error. *$p < 0.05$ vs. Spike 5 nM. Data are the mean ± SE, n = 3. Credits for the figure go to Ref.[153]

This discovery confirmed the results of the molecular docking as it appears that the reduction of the Spike-ACE2 interaction is due to the binding of the tested compounds with the Spike-RBD residues.

To confirm that the interaction with ACE2 was prevented after binding of ligands in the pockets located in the central core of Spike RBD's β sheet, we performed a series of experiments using the remains of plasma samples from five donors who recovered from COVID-19. We found that, despite all donors having a different titer of antibodies to SARS-CoV-2, the dilutions tested effectively inhibited the binding of Spike RBD to ACE2 by more than 95%, confirming that all results in our study

correctly identify the binding of SARS-CoV-2 RBD to ACE2, but the levels of inhibition were significantly lower than those that could be reached by anti-SARS-CoV-2 antibodies (Table 6).

Table 6. ***Quantitative Analysis of the Anti-SARS-CoV-2 IgG Antibodies for calculating Percentage of inhibition of the Spike:ACE2 binding.*** Serum efficacy has been calculated in ACE2:SARS-CoV-2 Spike Inhibitor Screening Assay Kit as percent of inhibition of Spike RBD binding to ACE2 binding obtained using SPIKE at 5 nM, arbitrarily set as 100%. Credits for the table go to Ref.[153]

		% of Binding Inhibition		
Patient ID	**Antibody Title**	**5 µL of Serum**	**10 µL of Serum**	**20 µL of Serum**
1	96.6 AU/mL	98.6	99.5	99.6
2	170 AU/mL	99.3	99.4	99.3
3	89.4 AU/mL	98.1	99.3	99.4
4	125 AU/mL	98.8	99.3	99.4
5	146 AU/mL	95.7	96.9	97.3

Therefore, in this study, we first identified Spike RBD as a potential pharmacological target and identified six potentially druggable pockets on the surface of the central core of the β sheet of Spike RBD that could be exploited in order to prevent virus binding to ACE2. Furthermore, through a virtual screening campaign of the FDA-approved drug library, we identified steroid compounds as potential ligands of pockets 1 and 5, demonstrating, through molecular docking studies, that steroidal scaffolds with A/B rings junction in trans configuration (Table 3) prefer pocket 1, while compounds with the A/B junction in cis configuration (Table 4) prefer pocket 5. The identification of pockets 1 and 5 through *in silico* studies was confirmed by *in vitro* tests. Furthermore, glycyrrhetinic acid and oleanolic acid showed good efficacy

in silico - in terms of docking ADscore - and *in vitro* for their ability to inhibit the Spike/ACE2 interaction. Furthermore, from the results obtained it is clear that the main determinant for the effectiveness of the inhibition is hydrophobicity. In fact, oleanolic acid, for example, does not make any charge interaction inside the pocket, making it the most effective inhibitor of the series. The importance of hydrophobicity is also suggested by the difference in binding mode and inhibition efficacy of CDCA and OCA with their 6-ethyl derivative, BAR704. As suggested by our studies, the 6α-ethyl group determines a critical effect in the inhibition activity, due to the possibility of engaging multiple hydrophobic contacts, while the contribution of the 3β-hydroxyl group is negligible. Furthermore, the internal energy contribution of the AD score, which is significantly higher for 6-ethyl derivatives, represents a measure of the conformational energy of the bound versus the unbound state of the ligand, thus indicating that the ethyl group facilitates the assumption of the bioactive conformation. Comparison of binding modes for G-CDCA and G-UDCA also supported the hypothesis that the activity is related to the network of hydrophobic interactions rather than the presence of a hydrogen bond. In fact, the binding mode of G-UDCA is very similar to that of G-CDCA. Finally, the improved inhibitory efficacy of BAR501 compared to UDCA confirmed once again that the inhibitory effect is not given by a side chain loading group. Finally, from the analysis of the binding mode of BAR501, it is possible to state that the stereochemistry of the ethyl group in C6 is not relevant from the pharmacophoric point of view since the 6β-ethyl group is still able to potentially interact with Pro384 and Tyr369.

Finally, it is important to underline that some of these agents are already known for their antiviral properties.[162] For example, oleanolic acid has been reported as an inhibitor of broad-spectrum influenza virus entry,[163] while it has been shown that betulinic acid can reduce inflammation and pulmonary edema induced by the influenza

virus[164] and potassium canrenoate, the main metabolite of spironolactone in alive, it intervenes in the renin-angiotensin-aldosterone system and is therefore used in hypertension. Finally, BAR501, a selective GPBAR1 agonist,[165] has shown promise in reducing inflammation and treating immune dysfunction by shifting the polarization of macrophages of the colon from the inflammatory M1 phenotype to the M2 anti-inflammatory phenotype, increasing the expression of Il-10 gene transcription in the intestine and increased secretion of Il-10 by macrophages.[38]

Furthermore, the study conducted on plasma obtained from donors recovered from COVID-19 shows that the small molecules that bind the hydrophobic pockets are less effective than antibodies to inhibit the Spike RBD/ACE2 interaction up to 99%. Thus, our pharmacological approach may be ineffective in the presence of a high viral load, but it could pave the way for further optimization of the binding mode in order to identify further potential interactions.

This study has been published in Frontiers in Chemistry (ref: Carino, A.; Moraca, F.; Fiorillo, B.; Marchianò, S.; Sepe, V.; Biagioli, M.; Finamore, C.; Bozza, S.; Francisci, D.; Distrutti, E.; Catalanotti, B.; Zampella, A.; Fiorucci, S. Hijacking SARS-CoV-2/ACE2 Receptor Interaction by Natural and Semi-synthetic Steroidal Agents Acting on Functional Pockets on the Receptor Binding Domain. *Front Chem.* **2020**, *8*, 572885.)

2.4 Bile acids derivatives as potent angiotensin converting enzyme 2 (ACE2) activators

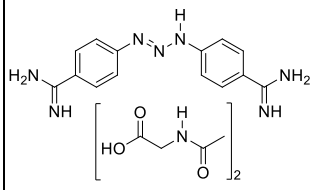
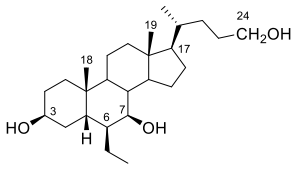
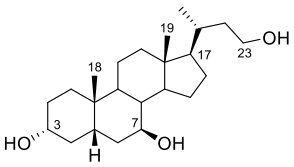
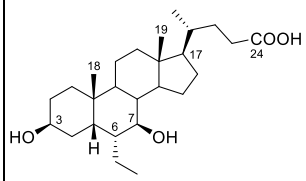
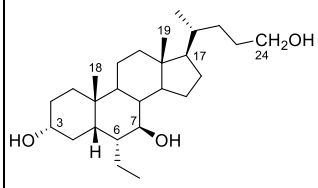
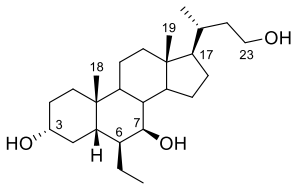
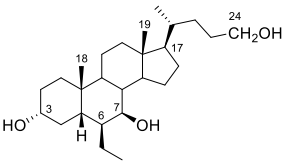
Driven by our previous studies on the inhibition of the Spike/ACE2 interaction by endogenous bile acids,[153] my research group and I have also investigated the activation of ACE2 by UDCA derivatives since no safe and druggable ACE2 activators are known. To this end, structure-based drug discovery approaches combined with experimental studies were employed with the aim to identify drug candidates that could bind with high affinity and selectivity to the target. The ligands' mechanism of action was also investigated.

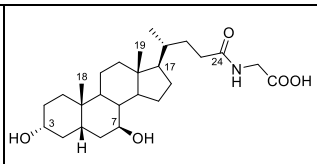
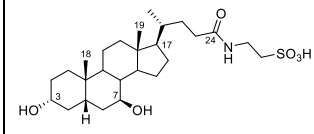
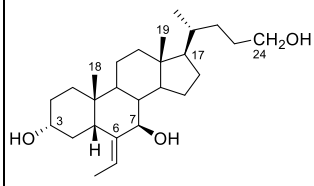
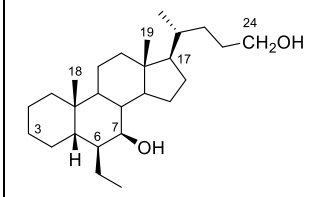
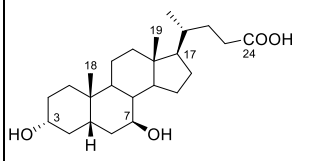
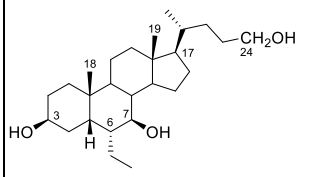
2.4.1 Computational studies

A study of VS of an *in house* library of 67 natural compounds and semisynthetic bile acid derivatives, enriched with previously identified ACE2 activators, hydroxyzine, minithixen, xanthenone and DIZE, was conducted to search for novel and druggable ACE2 activators. Molecular docking calculations were performed on the human X-ray structure of ACE2 in the open conformation (PDB ID 1r42)[166] retrieved from the Protein Data Bank in order to study ligand/protein binding interaction.[167] Docking calculations were carried out in the hinge-bending region, which resulted in the ACE2 activator binding site.[90,91]

Table 7 shows the VS results with ADscores for different BA derivatives.

Table 7. ***AutoDock4 docking scores (ADscore), enzymatic activity, and inhibition of ACE2/Spike interaction of the tested bile acids derivatives.***

Compound[a]	*Structure*	*ADscore (Kcal/mol)*	*ACE2 activity*[b]	*Inhibition of Spike/ACE2 Binding (%)*[c]
Diminazene aceturate (DIZE)		-6.44	141.04 ± 2.73	N.A.
BAR708		-8.98	141.28 ± 7.3	3.8 ± 1.9
BAR107		-8.61	140.13 ± 6.94	10.8 ± 6.4
BAR712		-8.43	116.37 ± 3.75	32.9 ± 3.92
BAR501-6alfa		-9.08	114.33 ± 1.25	40.8 ± 5.84
BARn501		-8.25	114.24 ± 1.19	39.9 ± 7.25
BAR501		-9.34	109.83 ± 12.99	30.8 ± 1.98[c]

Glyco-ursodeoxycholic acid (G-UDCA)		-8.58	108.34 ± 10.7	21.3 ± 0.63[c]
Tauro-ursodeoxycholic acid (T-UDCA)		-8.11	105.77 ± 6.98	42.4 ± 3.83[c]
BAR503		-9.42	105.01 ± 8.7	28.4 ± 8.02
BAR702		-9.16	102.74 ± 7.7	28.2 ± 6.5
Ursodeoxycholic acid (UDCA)		-8.37	102.23 ± 9.1	45.3 ± 2.23[c]
BAR707		-9.07	98.82 ± 11.95	18.9 ± 7.46

[a] Semi-synthetic compounds were synthetized by the research group of Professor Angela Zampella at the Department of Pharmacy of University of Naples-Federico II [b]Effect on ACE2 activity of compounds tested at 10 μM, referred to the activity in absence of any compound (100). Results are expressed as mean ± standard error. *p < 0.05 vs. Data are the mean ± SE, n = 3; [c] Inhibition of Spike-RBD/ACE2 Binding for each compound tested at 10 μM, expressed as % ± SE.

All compounds show similar binding modes within the ACE2 hinge-bending region, contacting both residues from Sub I (Lys94, Leu95, Glu98, and Glu102–Helix

α3) and Sub II (Tyr202, Asp206 from Helix α7, Val209, Asn210 from Helix 3_{10} H3, Pro565 and Trp566 from Helix α19) (Figure 35).

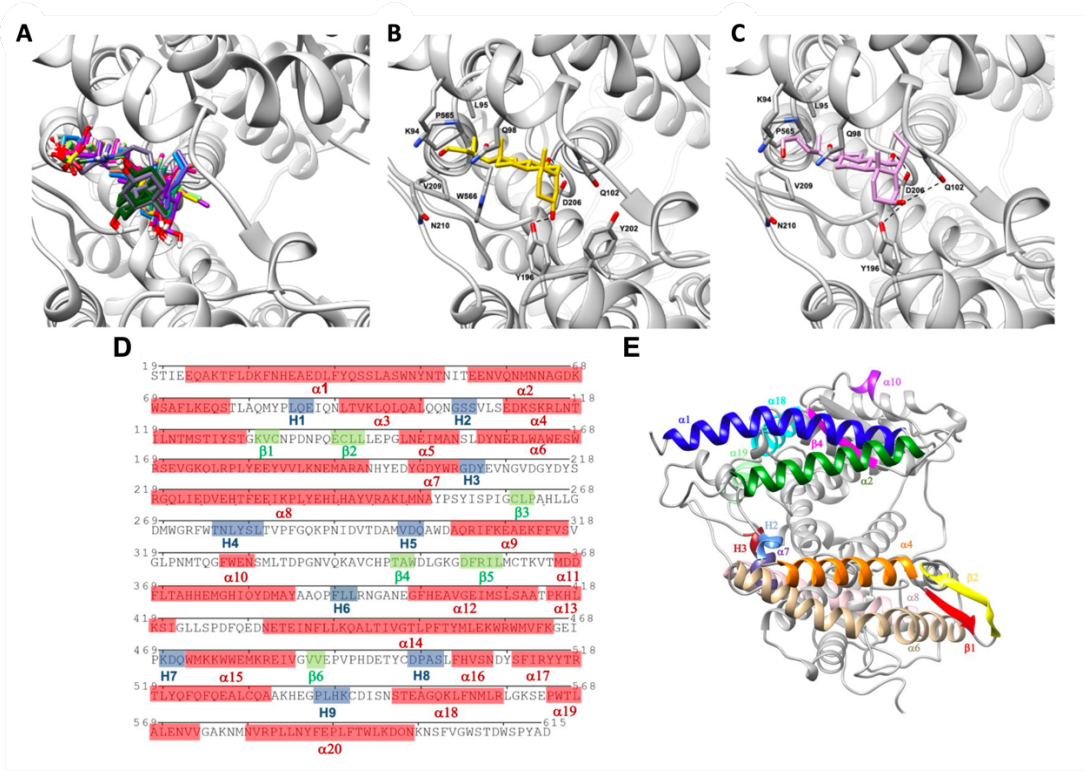


Figure 35. ***Graphical representation of the binding mode of the best docking poses.*** (A) Superimposition of the best docking pose for all the compounds in reported Table 1: (B) and (C) details of the best docking pose of BAR107 (*gold* sticks) and BAR708 (*light-violet* sticks), respectively. The interacting residues of the receptor are shown in *grey* sticks and labelled. Oxygen atoms are depicted in *red* and nitrogens in *blue*. Protein receptors are represented as grey cartoon. Hydrogens are omitted for the sake of clarity, while H-bonds are displayed as *black* dashed lines. (D) Numbering of the secondary structure of ACE2; (E) Labelling of the more relevant secondary structures in 3D structure (PDB ID 1r42).[166] Credits for the figure go to Ref.[168]

In order to identify the mechanism of action related to ligand-induced ACE2 activation, an MD study of the ACE2 metalloproteinase domain (PD domain) was performed. For this purpose, two human X-ray structures of ACE2 retrieved from the Protein Data Bank corresponding to the native forms (apo) and inhibitor-bound forms of the ACE2 PD (PDB ID 1r42 and 1r4l, respectively)[166] were employed, considering that the receptor is in a conformational equilibrium between an open and active state (native) and a completely closed and inactive state (inhibitor-bound). These two states

differ in the relative position of two non-continuous subdomains, the N-terminal subdomain I (Sub I; residues 19-102, 290-397, and 417-430) and the C-terminal subdomain II (Sub II; residues 103 –289, 398–416 and 431–615) (Figure 36A).

To determine whether ACE2 spontaneously passes from one conformational state to another, three independent MD simulations of 500 ns each of ACE2 apo (PDB ID 1r42)[166] were performed for a total simulation time of 1.5 μs.

The conformations visited by ACE2 obtained from the unified MD trajectory were analyzed by hierarchical clustering. The cluster analysis showed that the most populated cluster family (32% of the population) corresponds to the closed form of the receptor, similar to that assumed by the ACE2/inhibitor complex (PDB ID 1r4l)[166]. Instead, the second most populated cluster (20% of the population) was characterized by a completely open conformation similar to the X-ray structure of ACE2 apo, while the third cluster (18% of the population) assumed an intermediate conformation (Figures 36B and 37B).

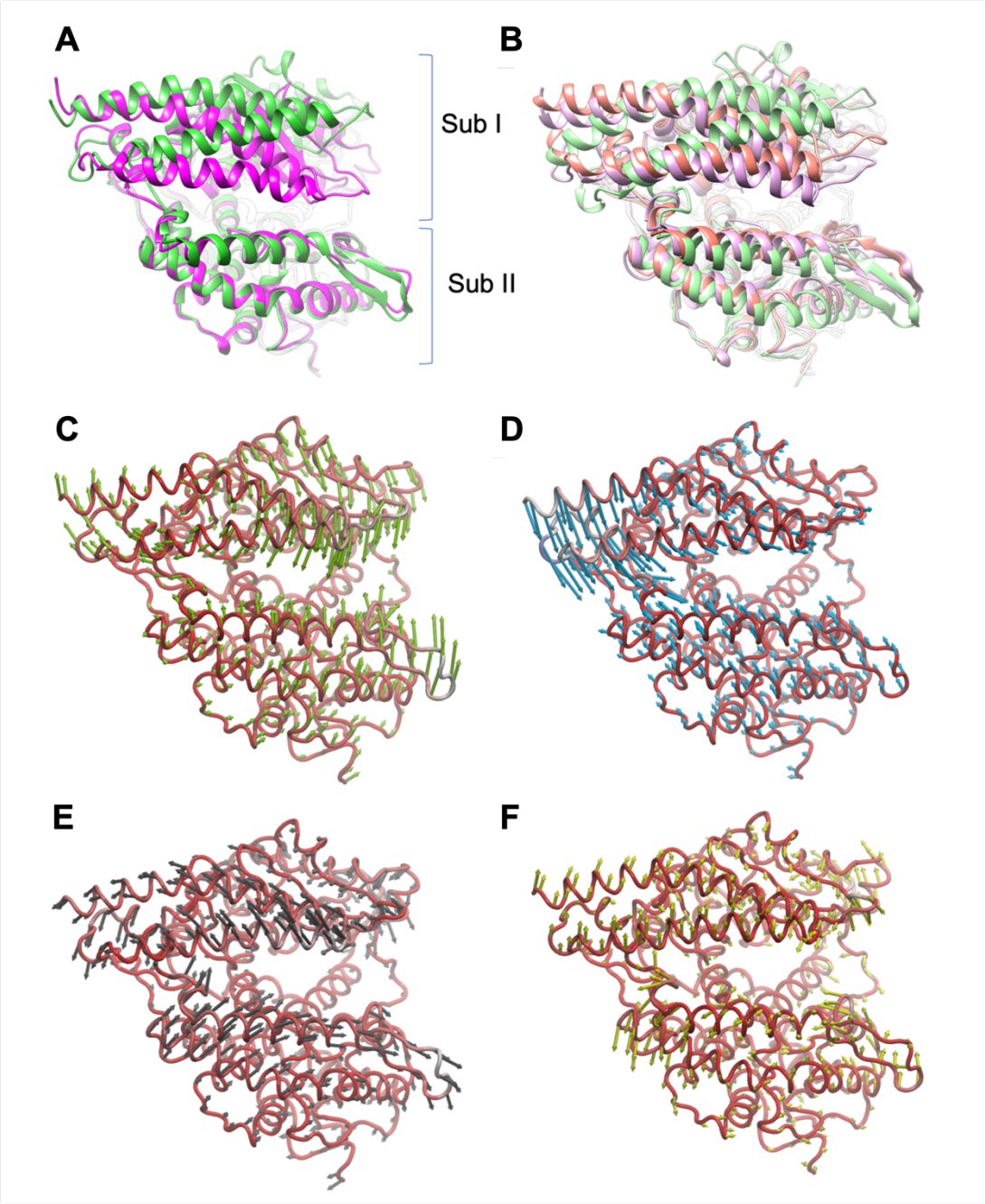
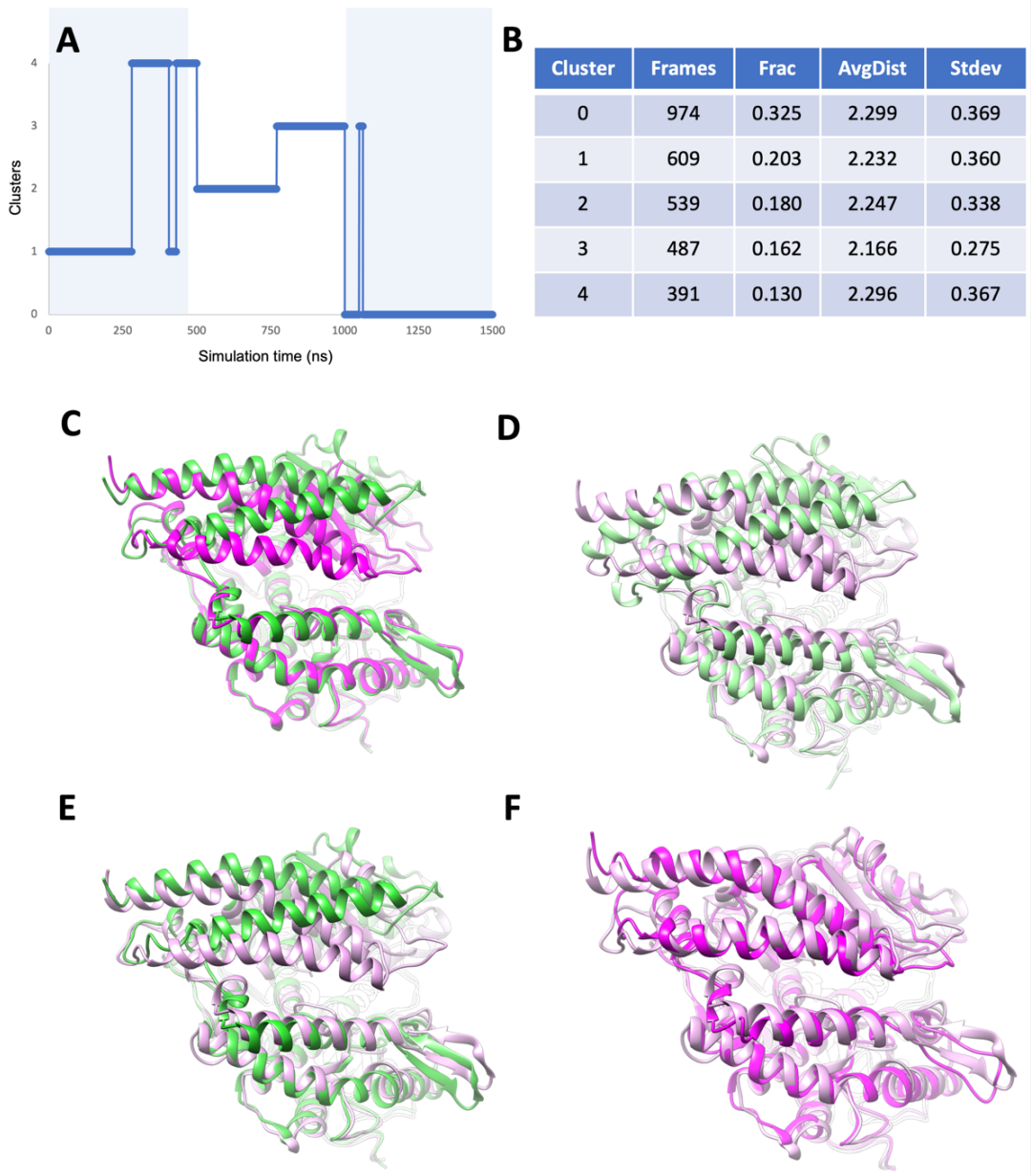


Figure 36. ***Dynamic behavior of native ACE2.*** Superimposition on the Sub II protein backbone atoms between: (A) the X-ray structures of the open apo ACE2 (PDB ID 1r42;[166] *green* cartoon) and the closed state of ACE2 complexed with the potent inhibitor MLN-4760 (PDB ID 1r4l;[166] *magenta* cartoon); (B) the most populated three clusters of apo ACE2 obtained after 1.5 μs MD simulation: Cluster0 (32% population, *light-violet* cartoon), Cluster1 (20% population, *light-green* cartoon), and Cluster 2 (18% population, *light-red* cartoon); (C–F) porcupine plots of the first four eigenvectors (PC1–4) identified from the PCA analysis after 1.5 μ MD simulations of apo ACE2. Protein backbones are represented as *red* ribbons, while arrows indicate the direction of the prominent motions and the length represents the magnitude of the corresponding eigenvalue. Credits for the figure go to Ref.[168]



Cluster	Frames	Frac	AvgDist	Stdev
0	974	0.325	2.299	0.369
1	609	0.203	2.232	0.360
2	539	0.180	2.247	0.338
3	487	0.162	2.166	0.275
4	391	0.130	2.296	0.367

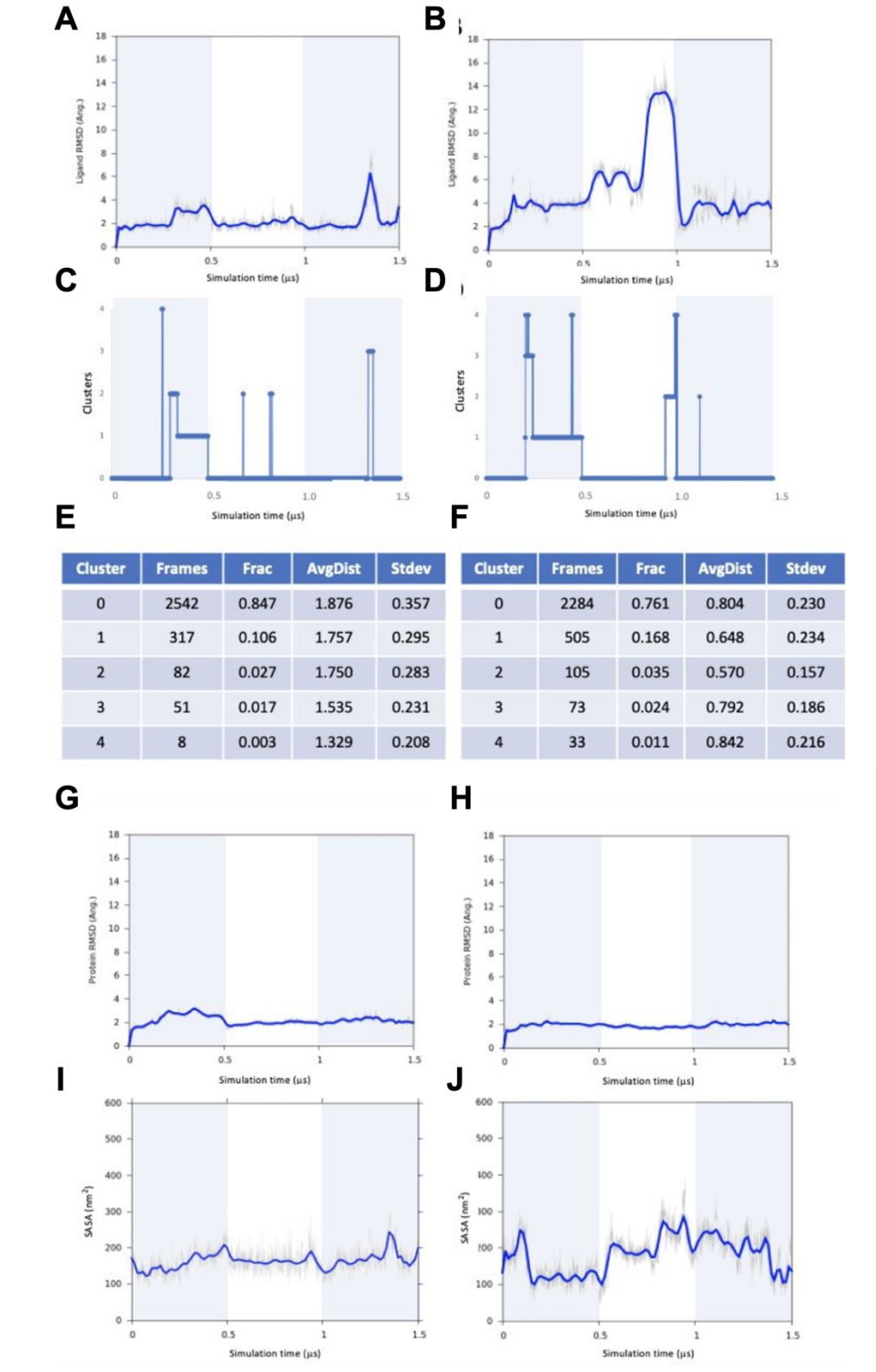
Figure 37. ***Dynamic states of native ACE2.*** (A) Cluster analysis of the MD simulations of apo ACE2; (B) plot representation of cluster distribution versus MD simulation time of the apo ACE2 receptor. (C-E) Superimposition on the Sub II protein backbone between: (C) the X-ray structures of the open apo ACE2 (PDB ID 1r42;[166] *green* cartoon) and the closed state ACE2 in complex with the potent inhibitor MLN-4760 (PDB ID 1r4l;[166] *magenta* cartoon); (D) the cluster1 (*light-green* cartoon) and the cluster0 (*light-violet* cartoon) of the apo ACE2 obtained after 1.5 μs of MD; (E) the X-ray structure of the open state apo ACE2 (PDB ID 1r42;[166] *green* cartoon) and the cluster0 of the closed form obtained after MD (*light-violet* cartoon); (F) the X-ray structure of the inhibitor-bound closed form (PDB ID 1r4l;[166] *magenta* cartoon), and the cluster0 of the closed obtained after MD (*light- violet* cartoon). Credits for the figure go to Ref.[168]

Subsequently, the trajectories were also studied through principal component analysis (PCA), which allows to identify the largest conformational changes of the enzyme, through the analysis of essential movements, associated with the longest time scale.

The result of the PCA showed that the first two main components are represented by the movement of Sub I towards Sub II. In particular, in the first component (PC1; Figure 36C), the movement of the sheets of the helix α4 and β1 − β2 of Sub II towards the helix α2 and the ring between α10 and β4 was observed, whereas in the second component (PC2) the helices α1 and α2 moved towards the 3_{10} H2 helix (PC2; Figure 36D). The third component, on the other hand, described a sliding motion of Sub II towards Sub I (PC3; Figure 36E). Taken together, these movements describe the conformational balance between the open and closed forms of the receptor. Finally, it is interesting to note that, according to PC1 and PC2, the critical region that allows the structural flexibility of ACE2 is the one surrounded by α4, by the C-terminal side of α6 and by the helix 3_{10} H3, region that is within hinge-bending region and corresponding to the BA binding site identified through the docking calculations previously performed.

The best docking pose of the most active compounds, BAR107 and BAR708, resulting from the docking calculations (Figures 35B and C) in complex with the ACE2 receptor in the native state (PDB ID 1r42)[166] was subjected to MD calculation. Once again, three independent MD runs of 500 ns each were performed for a total simulation time of 1.5 μs. The trajectories were then merged for the analysis of the dynamic behavior of the enzyme. As demonstrated by the RMSD and SASA plots calculated over time (Figure 38), the binding modes of BAR107 and BAR708 are stable, although BAR107 showed a slight binding mode change in the last 50 ns of the second MD run, in which a reversal of the binding orientation of the ligand at the

binding site is observed. However, this conformational change was found to belong to a less populated cluster, therefore not relevant (Figures 38D and F) and is probably due to the plasticity and solvent-exposed nature of the binding site at ACE2 hinge-bending region. This hypothesis was confirmed by the Solvent Accessible Surface Analysis (SASA). Indeed, the SASA of BAR708 (Figure 38I) shows slight fluctuations compared to the SASA of BAR107 (Figure 38J) in all three MD runs, thus showing a lower tendency for BAR708 to move towards an environment more exposed to the water solvents.



Cluster	Frames	Frac	AvgDist	Stdev
0	2542	0.847	1.876	0.357
1	317	0.106	1.757	0.295
2	82	0.027	1.750	0.283
3	51	0.017	1.535	0.231
4	8	0.003	1.329	0.208

Cluster	Frames	Frac	AvgDist	Stdev
0	2284	0.761	0.804	0.230
1	505	0.168	0.648	0.234
2	105	0.035	0.570	0.157
3	73	0.024	0.792	0.186
4	33	0.011	0.842	0.216

Figure 38. ***Molecular dynamics simulations.*** MD evolution time (μs) of the ligand RMSD (Å) of BAR708 (A), BAR107 (B) in ACE2. Plot representation of the cluster population of BAR708 (C) and BAR107 (D) during 1.5 μs of MD simulation. MD evolution time (μs) of the RMSD (Å) of ACE2 complexed with BAR708 (G) and BAR107 (H). MD evolution time (μs) of the Solvent Accessible Surface Area (SASA) of BAR708 (I) and BAR107 (J) in ACE2 receptor. Credits for the figure go to Ref.[168]

Then, we investigated whether the binding of BAR708 and BAR107 influenced the conformational behavior of the apo ACE2 receptor during the MD simulations, finding that BAR708 strongly influences enzymatic dynamics preventing the complete closure of Sub I to Sub II, as reported in the most populated cluster (85% of the population) which represents an open receptor structure, very similar to the open X-ray structure of apo ACE2 (Figure 39A).

The PCA conducted on the MD BAR708/ACE2 simulation shows the same essential motions found in the analysis made on the ACE2 apo, with a reduction of the vector length and in a different order. In fact, in this case, it is PC2 that represents the sliding movement between Sub I and Sub II (Figure 39D), corresponding to PC3 in apo ACE2 PCA, while component PC3 describes the movement of the helices α1 and α2 towards the helix 3_{10} H2 (Figure 39E), corresponding to PC2 in apo ACE2 PCA. Therefore, from these results it is possible to state that the binding of BAR708 in the flexion region of the hinge prevents the spontaneous closure of Sub I on Sub II of ACE2 and thus stabilizes the open conformation of the enzyme, providing evidence of a molecular mechanism activation of ACE2.

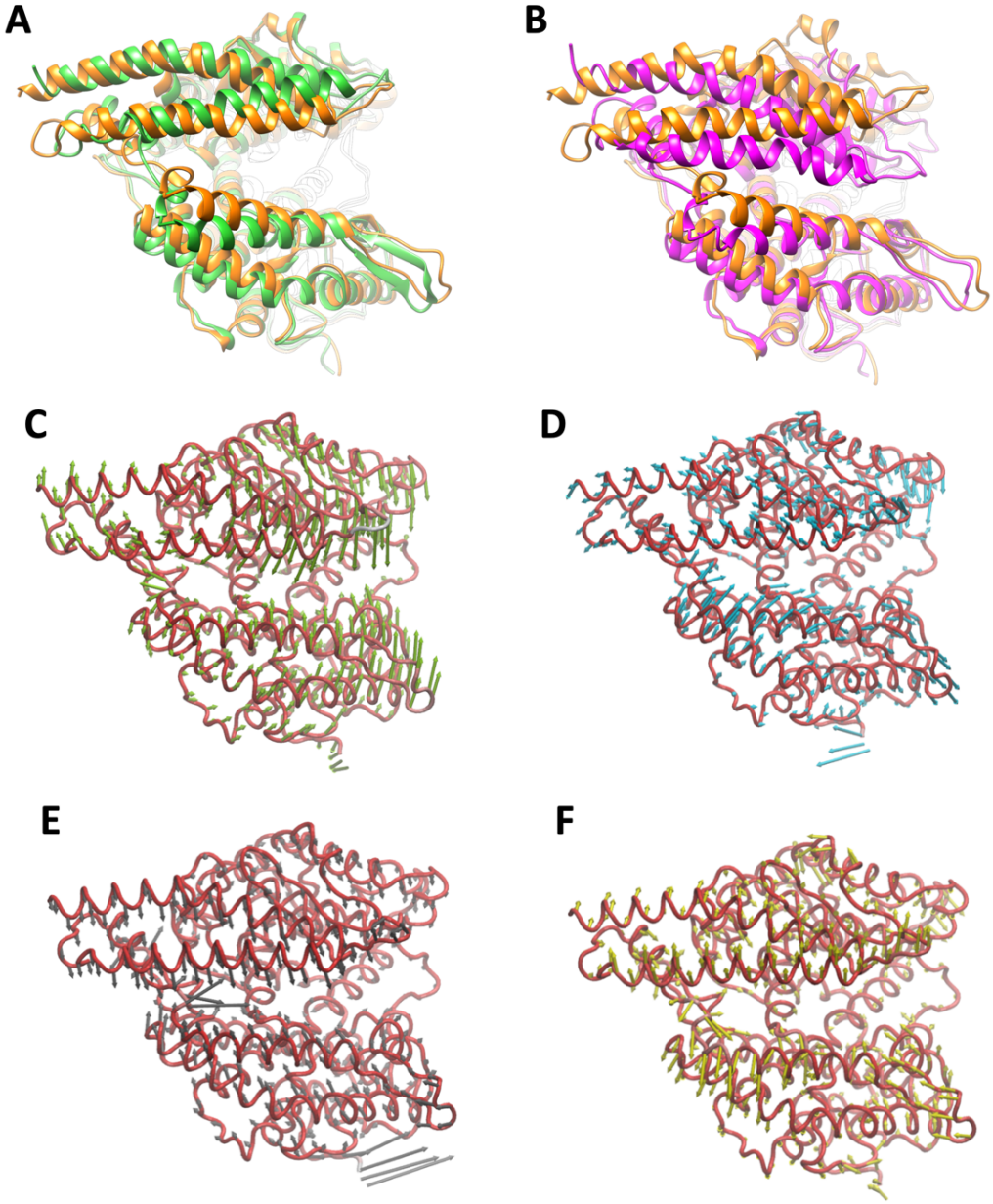


Figure 39. ***Dynamic states of ACE2 complexed with the activator BAR708.*** (A, B) Superimposition on the Sub II protein backbone between (A) the most populated cluster over 1.5 μs MD of the BAR708/ACE2 complex (*orange* cartoon) and the X-ray structure of the open apo ACE2 (PDB ID 1r42; *green* cartoon)[166] and (B) the most populated cluster over 1.5 μs MD of the BAR708/ACE2 complex (*orange* cartoon) and the closed ACE2 complexed with the potent inhibitor MLN-4760 (PDB ID 1r4l;[166] *magenta* cartoon). (C–F) Correlated motions from the PCA analysis during 1.5 μs MD simulations of the BAR708/ACE2 complex, represented by porcupine plots of the first four vectors (PC1–4). Protein backbones are represented as ribbons, the arrows indicate the direction of the motion, and the length represented the magnitude of the corresponding eigenvalue. Credits for the figure go to Ref.[168]

The analysis of the MD simulations of BAR107 also confirmed the previous hypothesis since the main cluster (76% of the population) obtained from the MD trajectory of the BAR107/ACE2 complex (Figure 38F) corresponds to an open receptor conformation (Figures 40A and B). Furthermore, also the PCA of the first four major components (PC1–4) of the BAR107/ACE2 MD simulations (Figures 40C–F) shows a reduction in Sub I closure on Sub II, as observed in BAR708/ACE2.

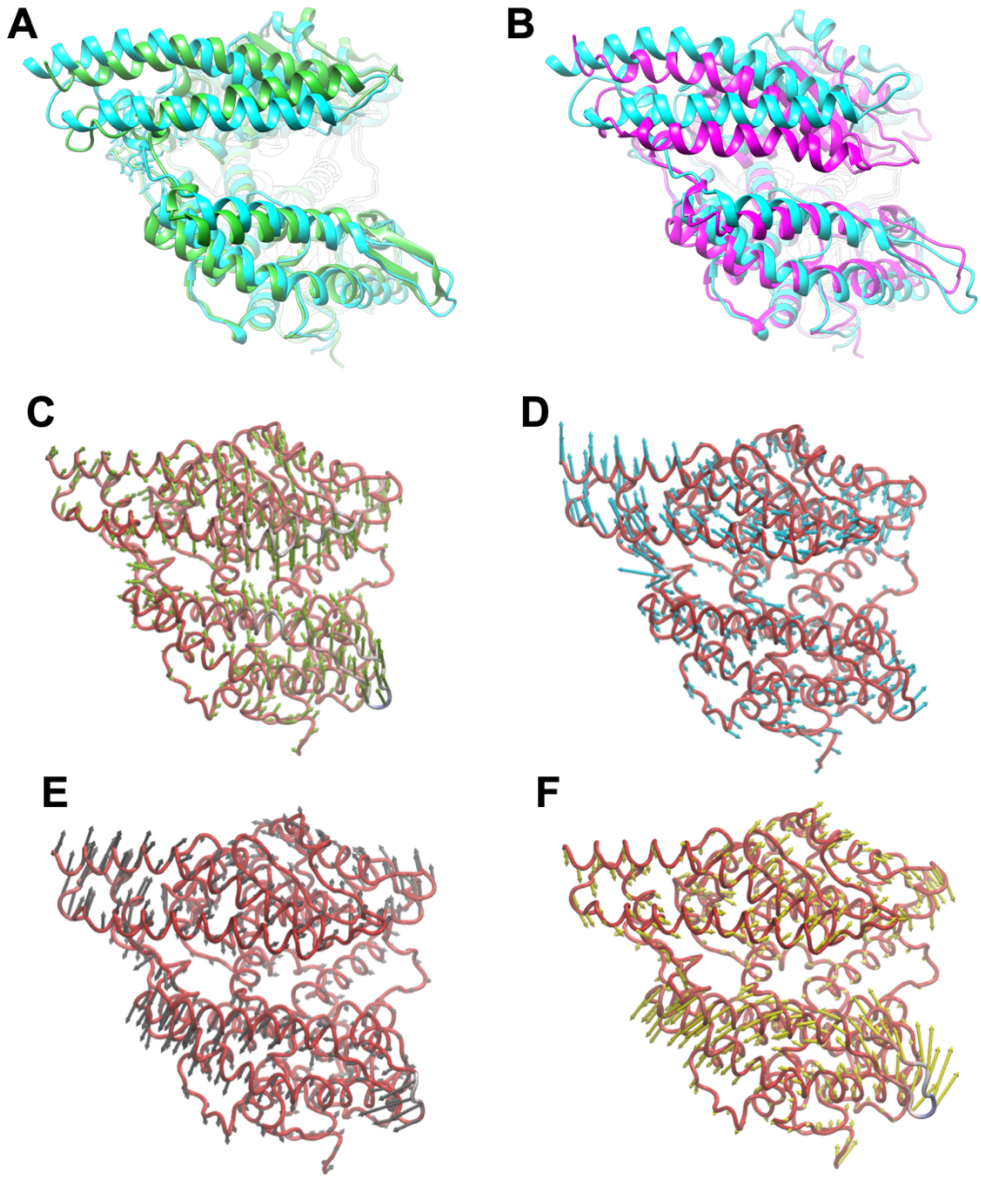


Figure 40. ***Dynamic states of ACE2 in complex with the activator BAR107.*** (A) superimposition between the most populated cluster0 (76%) over 1.5 μs MD of the ACE2/BAR107 complex (*cyan* cartoon) and the X-ray structure of the open apo form ACE2 (PDB ID 1r42;[166] *green* cartoon); (B) the most populated cluster0 (76%) over 1.5 μs MD of the ACE2/BAR107 complex (*cyan* cartoon) on the closed state ACE2 in complex with the potent inhibitor MLN-4760 (PDB ID 1r4l;[166] *magenta* cartoon); (C-F) Correlated motions obtained from the PCA analysis of 1.5 μs MD simulations of the ACE2/BAR107 complex, represented by porcupine plots of the first 4 vectors. Protein backbones are represented as *red* ribbons, the arrows indicate the direction of the motion, and the length represented the magnitude of the corresponding eigenvalue. Credits for the figure go to Ref.[168]

To analyze the general dynamic behavior of ACE2, the residual fluctuation (root mean square fluctuation, RMSF) in each different system was also calculated. As shown in the plot in Figure 41, the binding of activators ACE2, BAR107 and BAR708 (brown and magenta lines), significantly reduced the amplitude of residual fluctuation in three regions, differently from the apo ACE2 RMSF (black line) which instead, shows higher fluctuations.

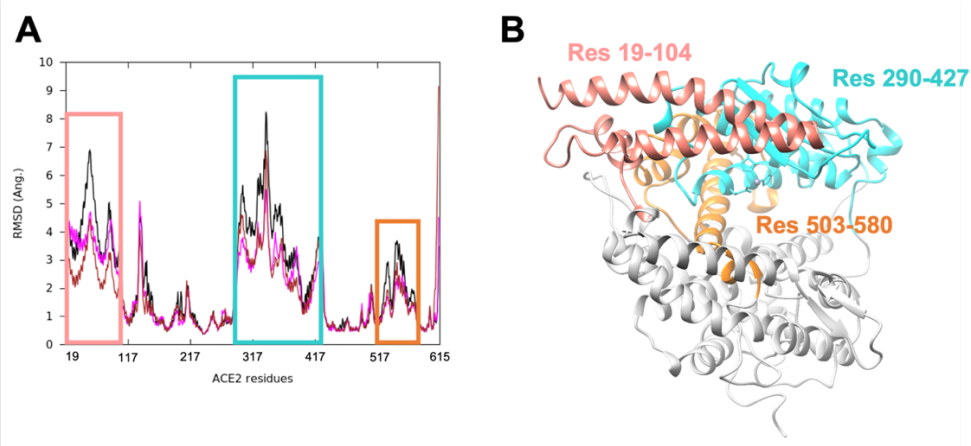


Figure 41. ***(A) RMSF plot of the ACE2 residues in different systems***: ACE2 apo form (*black* line), ACE2 complexed with BAR708 (*brown* line) and BAR107 (*magenta* line). ***(B) Open native state of ACE2 highlighting the most fluctuating residues from RMSF analysis.*** RMSF was calculated on Cα atoms aligning on the α-helices of Sub II. Credits for the figure go to Ref.[168]

MD binding mode showed that both BAR107 and BAR708 bind a specific binding pocket stacked between the helix α3, the loop between α7 and α8, and the loop connecting α18 and α19, interacting differently with the residues of the binding pocket depending on the intrinsic chemical properties and the position of the substituents on the steroid scaffold (Figure 42).

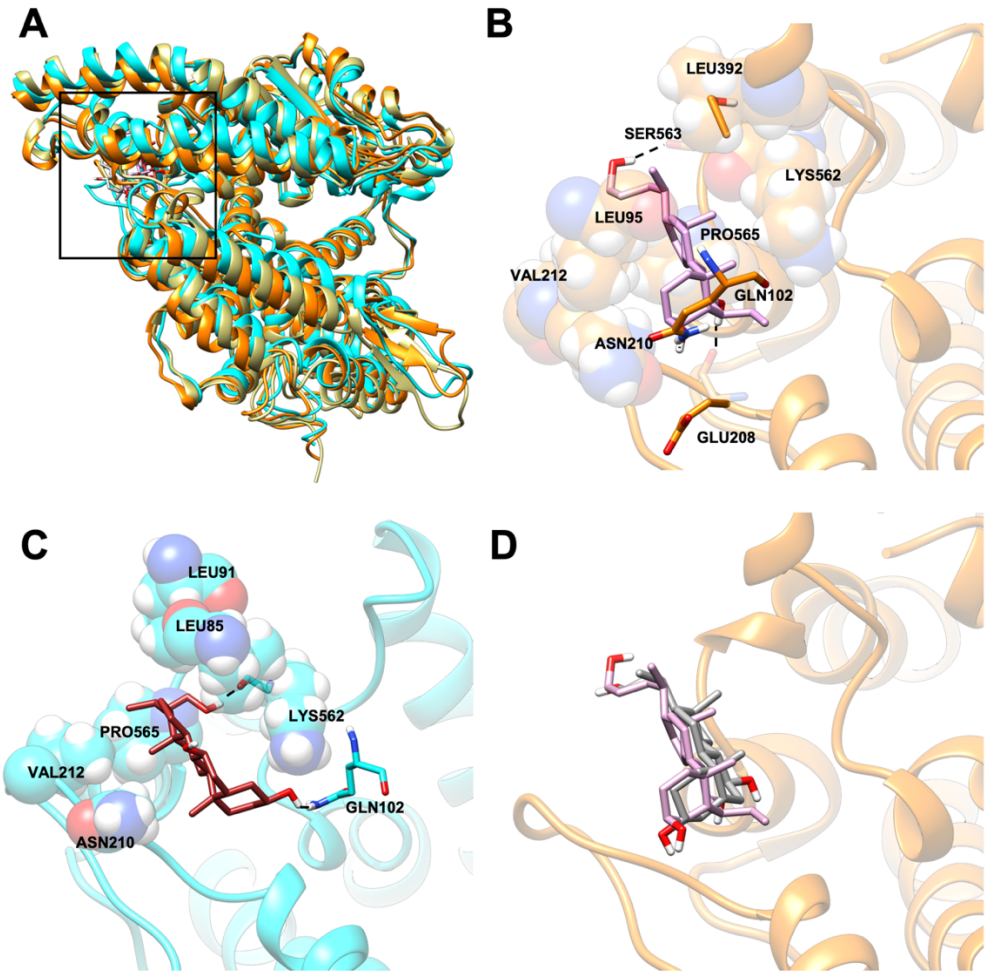


Figure 42. ***Molecular dynamics derived clusters of BAR708/ACE2 complex.*** (A) Overall representation of the most populated MD-derived clusters of BAR708/ACE2 complex (*orange* cartoon) and BAR107/ACE2 complex (*cyan* cartoon). The black square indicates the hinge-bending region targeted as the agonist binding site. (B) Cluster0 (85%) binding mode of BAR708/ACE2 complex (protein is represented in the *orange* cartoon, while ligand in the *light-violet* stick). (C) Cluster0 (76%) binding mode of BAR107/ACE2 complex (protein is represented in the *cyan* cartoon, ligand in the *brown* stick). (D) Superimposition between the cluster1 of BAR107/ACE2 complex and the cluster0 BAR708/ACE2 complex (protein represented in the *orange* cartoon, BAR708 in the *light-violet* stick, and BAR107 in the dark-gray stick). Credits for the figure go to Ref.[168]

As mentioned above, the binding mode of BAR708 in ACE2 is quite stable during the three MD simulations performed for a total of 1.5 μs (Figures 38 and 39). In the most populated cluster (85% of the population), the steroidal scaffold is located in an amphipathic pocket, making contact with the hydrophobic side chains of Val212, Leu392 and with the methylene chain of Lys562, while the 3β-OH H-bonds the side chain of Gln102, while the 7β-OH H-bonds with the backbone of Glu208. Furthermore, the side chain of the ligand at C17 makes hydrophobic interactions with the side chains of Leu395 and Pro565. Finally, the hydroxyl group at the C24 side chain makes an additional H-bond with the side chain of Ser253.

Even the most populated cluster (75% of the population) obtained from the MD of the BAR107/ACE2 complex, shows the BAR107 steroidal scaffold positioned in the same amphipathic pocket as BAR708, between the α3 helix and the ring between α7 and α8 (Figure 42C). However, differently from BAR708 binding mode, BAR107 is rotated along the axis of the steroid scaffold, maintaining the H-bond between 3α-OH and the carbonyl group of Gln102 side chain, the hydrophobic contact between the steroidal moiety and the methylenic chain of Lys562 and the hydrophobic interaction of the C24 side chain with Pro565. The steroidal scaffold establishes further hydrophobic contacts with Leu95, while the methyl group at C21 interacts with Val212 and the alkyl chain ligand at C17 engages in contacts with Leu91. Furthermore, 3α-OH H-bonds the Gln102 side chain, as seen for BAR708, while the C19 methyl group points towards Glu208 and Asn210. Finally, the hydroxyl group on the C24 side chain makes an H-bond with the Lys562 backbone carbonyl. Conversely, it is important to note that the binding mode represented by the centroid of the second most populated cluster (19% of the population) is very similar to that of BAR708 (Figure 42D) and does not cause changes in the enzyme protein conformation, as shown by the very low RMSD value calculated for the backbone of the protein (1.8 Å) (Figure 38H).

2.4.2 In vitro enzymatic assays

In vitro binding assays were performed by the research group of Professor Stefano Fiorucci at the Department of Medicine and Surgery of the University of Perugia.

Compounds with the best ADscores have been experimentally investigated for the activity towards ACE2, using the ACE2 Inhibitor Screening Assay Kit and DIZE as the reference ACE2 activator.

The result shows that none of the tested compounds inhibited ACE2 activity, while BAR708, BAR712, BARn501, BAR501-6α and BAR107 significantly increased ACE2 activity (Table 7, Figure 43). Additionally, BAR708 and BAR107 were shown to be as effective as DIZE in activating ACE2, confirming the docking results on the possibility that UDCA derivatives are able to bind and activate ACE2.

Finally, still on the line of our previous discovery on the ability of BA derivatives to inhibit the ACE2/SARS-CoV-2 RBD interaction,[153] we investigated whether the compounds in Table 7 could interfere with the interaction between Spike-RBD and ACE2 using one Spike/ACE2 Inhibitor Screening Assay Kit.[153] By incubating Spike-RBD with UDCA derivatives, the interaction between Spike-RBD and ACE2 is reduced in a concentration-dependent manner (Table 7, Figure 43), as in the case of BARn501 and BAR501-6α which reduce it by ∼40%, while other derivatives by ∼30%, confirming, once again, our previous discovery of the ability of BA and derivatives to inhibit the binding of Spike-RBD to the ACE2 *in vitro*.

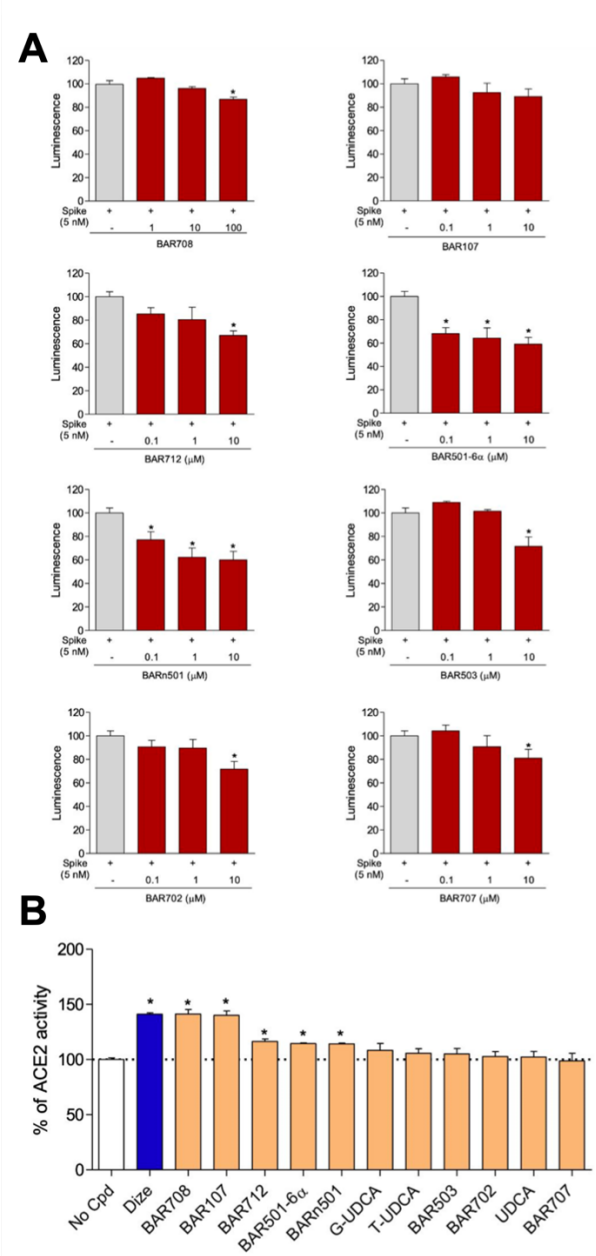


Figure 43. ***(A) The ACE2:SARS-CoV-2 Spike Inhibitor Screening assay***. BAR708, BAR107, BAR712, BAR501-6α, BARn501, BAR503, BAR702 and BAR707 were tested at different concentration (0.1, 1, and 10μM), to evaluate their ability to inhibit the binding of Spike protein (5 nM) to immobilized ACE2. Luminescence was measured using a Fluo-Star Omega fluorescent microplate reader. Luminescence values of Spike 5 nM were arbitrarily set to 100%. Results are expressed as mean ± standard error. *p < 0.05 vs Spike 5 nM. ***(B) ACE2 activity assay.*** Compounds were tested on a cell-free enzymatic assay to screen activators of ACE2 activity. Dize was used as positive control. The assay is designed to measure the exopeptidase activity of ACE2, it utilizes the ability of an active ACE2 to cleave a synthetic fluorogenic substrate to release a free fluorophore. The released fluorophore is quantified using a fluorescence microplate reader. Fluorescence values of activity in absence of any compound were arbitrarily set to 100%. Results are expressed as mean ± standard error. *p < 0.05 vs No Cpd. Credits for the figure go to Ref.[168]

In conclusion, in this study, the discovery of UDCA derivatives that activate ACE2 and its mechanism of action, identified through atomistic studies MD and slower degrees of freedom of ligand/ACE2 systems (i.e., PCA) is reported. A database of natural compounds and semisynthetic bile acid derivatives targeting the hinge-bending region was used for a virtual screening campaign with the final aim to find novel bile acid derivatives capable of activating ACE2. The results obtained from the *in silico* study were confirmed by *in vitro* enzyme assays. Among the UDCA derivatives with promising ACE2 activating activity, BAR107 and BAR708, showed an ACE2 activation comparable to that of DIZE, the most studied ACE2 activator.[92,93,96,95,94,97]

First, through the analysis of 1.5 μs of MD simulations conducted on the crystal structure of the native ACE2 apo (PDB ID 1r42),[166] it was found that the protein structure undergoes conformational changes involving the movement of a region (Subdomain I) towards another (Subdomain II), leading to a completely closed structure similar to the conformation of ACE2 in complex with the potent inhibitor MLN-4760 (PDB ID 1r4l).[166] The PCA study conducted on this MD trajectory clarified that the opening/closing movement is regulated by the segment of the hinge-bending region which includes the helix α3, the ring between α7 and α8 and the ring connecting α18 and α19.

Successively, the effects induced by the binding of BAR107 and BAR708 on the conformational change of ACE2 were studied by analyzing the MD trajectories of 1.5 μs. performed for each complex. PCA, RMSF and clustering revealed that these ligands reduce the full closure of Subdomain I on Subdomain II, stabilizing the open conformational state and thereby inducing activation of the enzyme.

On the other hand, the binding mode of the two compounds shows different ligand orientations within the same binding pocket, although they share some interactions. This is related to the presence of different chemical substituents on the steroidal

scaffold which results in different binding modes, but similar effects on the conformational freedom of ACE2.

Furthermore, considering that the series of BA derivatives analyzed in this work differ in the stereochemistry of the hydroxyl group in position 3, the presence and stereochemistry of an ethyl group in position 6, for the length of the side chain at C17 of the tetracyclic nucleus and for the functional terminal group on the side chain, we found that the specific stereochemistry of individual substituents does not play a key role.

Moreover, based on our previous discovery on the ability of UDCA and other BA derivatives to influence interactions between the SARS-CoV-2 RBD spike protein and ACE2 *in vitro*,[153] we investigated the capability of this series of derivatives UDCA through a SARS-CoV-2/ACE2 interaction assay, finding that two new BA derivatives, BARn501 and BAR501–6α, inhibit the RBD/ACE2 interaction by approximately 40%, while they are mild ACE2 activators. On the other hand, BAR107 and BAR708, the best ACE2 activators in this series, showed a very weak *in vitro* inhibition of RBD/ACE2 interaction. Conversely, the reference compound DIZE did not exert any effect on the RBD/ACE2 interaction.

Hence, from this study emerged two best compounds of the series, BAR107 and BAR708, able to activate ACE2 more than DIZE. Furthermore, the ligands' mechanism of action was also delineated, finding that the new discovered compounds involve a peculiar conformational change of the enzyme. These two molecules can be used in the prevention and treatment of infection and inflammation due to SARS-CoV-2, as well as other pathologies related to the dysfunction of the RAS/Mas pathway.

Furthermore, since BAR107 and BAR708 are active both on ACE2 and on other bile acid receptors - in fact, they are classified as GPBAR1 agonists exerting anti-

inflammatory effects- the future prospects are to study a possible crosstalk between the RAS/Mas pathway and that of activation of the bile acid receptors.

This study has been published in Journal of Chemical Information and Modeling (ref: Fiorillo, B.; Marchianò, S.; Moraca, F.; Sepe, V.; Carino, A.; Rapacciuolo, P.; Biagioli, M.; Limongelli, V.; Zampella, A.; Catalanotti, B.; Fiorucci, S. Discovery of Bile Acid Derivatives as Potent ACE2 Activators by Virtual Screening and Essential Dynamics. *J Chem Inf Model.* **2022**, *62*, 196-209.)

CHAPTER III

Side projects

In this section, I will present the computational studies I carried out in other scientific projects not related to my main doctoral project on multi-potent compounds targeting bile acid receptors.

3.1 The TBC1D31/praja2 complex controls primary ciliogenesis through PKA-directed OFD1 ubiquitylation

3.1.1 Introduction

During my PhD, I was involved in a project, in collaboration with the research group of Professor Antonio Feliciello, of the Department of Molecular Medicine and Medical Biotechnologies at the University of Naples-Federico II in Naples, aiming to disclose the mechanism of the ciliopathies, which are due to genetic defects of the primary cilia.[169]

The cilium is an organelle and protuberance that protrudes from the cell body present in eukaryotic cells. There are two types of cilia: motile and non-motile cilia. Non-motile cilia, also called primary cilia, act as sensory organelles. They are based on microtubules and mediate the signaling pathways from the environment or from other cells. Although primary cilia were discovered in 1898, they were largely ignored for a century and considered a vestigial organelle without important function. Recent findings regarding its physiological roles in chemosensation, signal transduction, and cell growth control and development, have revealed its importance in cell function.[170,171,172,173,174,175]

Mutations involving the cilia biogenesis process often cause genetic developmental disorders, also known as ciliopathies.

Among the various types of congenital ciliopathy is type I digital orofacial syndrome (OFDI), X-linked and caused by a mutation that inactivates the OFD1 gene, a protein that plays a key role in cilium biogenesis: in particular, the OFD1 degradation is necessary for the cilium formation. In fact, OFD1 syndrome resides in a centrosomal complex that links GPCR signaling to ubiquitylation and degradation of OFD1,

controlling cilium morphology and dynamics and vertebrate development. The mutation in the OFD1 gene generates malformations of the oral cavity, face, fingers and, furthermore, it may be the cause of polycystic kidney disease and implications at the level of the central nervous system.[176,177]

It has been declared that in serum-deprived cells, it is necessary to remove OFD1 from the centriolar satellites by autophagy for ciliogenesis to begin.[178]

This work has been principally focused on the mechanism of OFD1 degradation. For this reason, it has been identified a novel complex, assembled at centrosomes and centriolar satellites that positively regulates ciliogenesis and signaling pathways, formed by TBC1D31, the E3 ubiquitin ligase praja2, protein kinase A (PKA), and OFD1. It is important to note that TBC1D31 is essential for ciliogenesis (Figure 44).

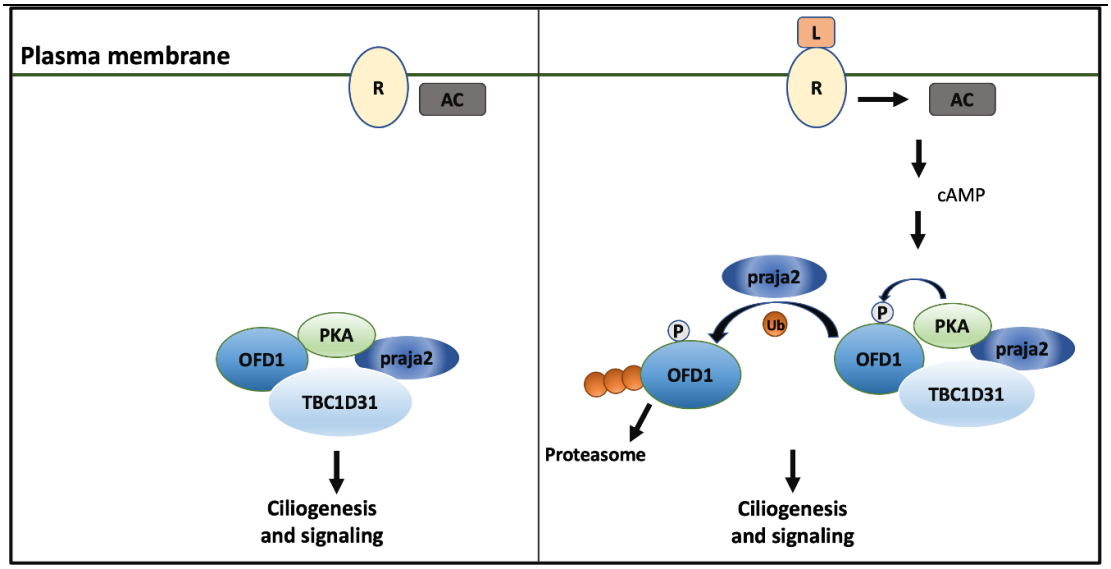


Figure 44. ***Schematic representation of the complex assembled at centrosomes by TBC1D31, including the E3 ubiquitin ligase praja2, protein kinase A (PKA), and OFD1.***

Among the various pathways involved in the deregulation of the signaling pathways linked to the ciliary compartment, there is that of cAMP, a known second messenger that plays a fundamental role in the regulation of metabolism, cell growth,

development, differentiation and synaptic activities.[179,180] In particular, the increase in cAMP intracellular concentration induces the activation of PKA, which is found in the cell in the form of a tetrameric holoenzyme composed of two regulatory (R) and two catalytic (PKAc, C) subunits. When a molecule binds to a GPCR, activating it, there is the activation of adenylate cyclase (AC) with an increase in intracellular cAMP which causes the dissociation of the PKA holoenzyme and the consequent release of active PKAc subunits. PKAc performs important regulatory functions of biological processes thanks to its ability to phosphorylate cell substrates.[181,182,183] The cAMP signals at sites distal from signal generation can be regulated by the activity of A-kinase anchor proteins (AKAP) which are capable of compartmentalizing PKA at certain intracellular sites.[184,185,186,187,188,189,190] AKAP praja2 binds and targets the PKA holoenzyme to the cell membrane, perinuclear region and cell organelles. The co-localization of praja2-PKA complexes with PKA effector molecules results in the propagation of locally generated cAMP to distant target sites.[191] Indeed, praja2 acts as an E3 ubiquitin ligase, controlling the ubiquity and stability of PKA.[192,193,194,195,196,197] PKA plays a key role in cilia biology and, therefore, components of the cAMP cascade as residents and regulators of the ciliary compartment have been analyzed.[198,199,200] In particular, it has been found that PKA activation within the cilium inhibits the Sonic Hedgehog pathway, a regulator of embryonic development.[201,202] Furthermore, in 2018 PKA signaling has been related to the ubiquitin-proteasome system in the ciliary compartment by *Porpora et al.*,[203] arguing that PKA phosphorylation of NIMA-related kinase NEK10 promotes its ubiquitylation by the E3 CHIP/Stub1 ligase. NEK10 once ubiquitylated, is degraded by the proteasome, leading to primary resorption of the cilia.[203] However, the role of this PKA-ubiquitin signaling system in cilia biogenesis still remains unknown.

The mechanism underlying ciliogenesis is thus characterized by PKA-induced phosphorylation, after the activation of the G protein-coupled receptor (GPCR) and the increase of cAMP intracellular concentration, of OFD1 at Ser735, thus promoting OFD1 proteolysis through the praja2-UPS circuitry (Figure 44).[181,183] This is possible considering that a non-phosphorylatable OFD1 mutant greatly influences the morphology and dynamics of the cilium. Finally, through computational studies, which include docking calculations and MD simulation, and *in vitro* assays, we revealed the molecular basis of praja2 and TBC1D31 interaction.

3.1.2 Results and Discussion

Thanks to the collaboration with the pharmacological group of Prof. Antonio Feliciello, the protein–protein interaction network of physical interactors of OFD1, praja2 and TBC1D31 was clarified.

First, the C-terminus residue of TBC1D31$_{940\text{-}970}$ was identified, using the C-terminus of praja2 as bait, by screening two yeast hybrids.

Then, through co-immunoprecipitation (CoIp) experiments, the interaction between praja2 and TBC1D31 in cell lysates was confirmed (Figure 45A). Using deletion mutagenesis assays and CoIp assays, the segment praja2$_{530\text{-}630}$ that binds to TBC1D31 was identified (Figures 45B and C); while, through GST pull-down experiments, the TBC1D31$_{940\text{-}970}$ fragment has been identified as residues interacting with praja2 (Fig. 45D). Furthermore, since praja2 is known to bind PKA to specific intracellular sites,[191] the presence of PKA in the praja2/TBC1D31 complex was tested, again through CoIp assays, finding the PKAc-praja2-TBC1D31 complex in the immunoprecipitates (Figure 45E). In situ immunostaining analysis confirmed both TBC1D31 is localized to the centrosome and the centriolar satellites and the presence of a praja2/TBC1D31 complex within the same intracellular compartment (Figure 45F, upper and lower

panels, respectively). Furthermore, since TBC1D31 acts as an anchor for praja2, the genetic silencing of TBC1D31 drastically reduced the localization of praja2 to the centrosome and the centriolar satellites (Figure 45G). In contrast, praja2 silencing did not have a significant impact on the intracellular localization of TBC1D31 (Figures 45I and J), once again supporting the idea that TB1D31 acts as an anchor for praja2 (Figure 45H).

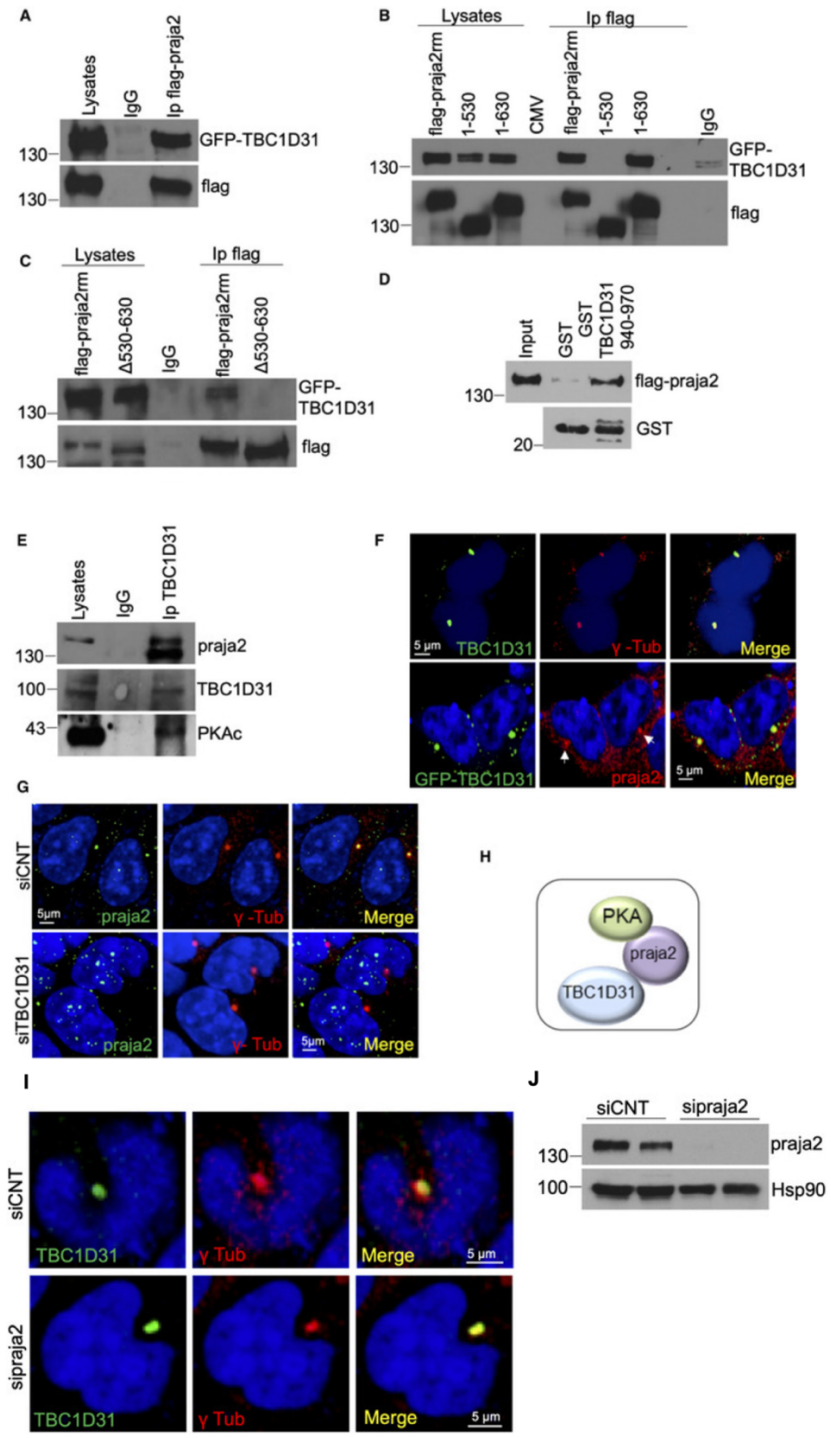


Figure 45. ***TBC1D31 binds and targets praja2 to the centrosome.*** A) Co-immunoprecipitation of flag-praja2 and GFP-TBC1D31 from lysates of HEK293 cells. The immunoprecipitation (Ip) was performed using an anti-flag antibody or control IgG. (B and C) Same as in (A), with the exception that cells expressing flag-praja2rm or praja2 deletion mutants (praja2$_{1–530,}$ praja2$_{1–630}$ and praja2$_{\Delta530–630}$) were included in the analysis. (D) Lysates expressing flag-praja2 were subjected to pull down assay with GST and GST-TBC1D31$_{940–970}$polypeptides. (E) Co-immunoprecipitation of endogenous TBC1D31/praja2/PKAc complex from cell lysates. (F) Staining of

HEK293 cells with anti-TBC1D31, anti-γ-tubulin and anti-praja2 antibodies. Nuclei were stained with DRAQ5 (blue). Where indicated, cells were transfected with GFP-TBC1D31. Arrows indicate the pool of praja2 colocalizing with TBC1D31 staining at the centrosome. (G) Cells transfected with control siRNA (siCNT) or siRNA targeting TBC1D31 (siTBC1D31) were stained for praja2, anti-γ-tubulin and DRAQ5. (H) Schematic picture of TBC1D31/praja2/PKA complex. (I) Cells transiently transfected with control siRNA or siRNA targeting endogenous praja2 were stained for TBC1D31, γ-tubulin and DRAQ5. (J) Immunoblot analysis of praja2 and Hsp90 in siRNA-silenced cells. Credits for the figure go to Ref.[204]

Next, we investigated the interaction between praja2 and TBC1D3 *in vitro*, using a fusion protein carrying praja2$_{530\text{-}630}$ segment fused to the C-terminus of the glutathione S-transferase polypeptide (GST) co-precipitated GFP-TBC1D31 from cell lysates. (Figure 46A). To minimize residues of the praja2 segment that binds TBC1D31, microscale thermophoresis binding experiments were performed *in vitro* using synthesized peptides of different lengths, but which included the praja2$_{530\text{-}630}$ domain. As shown in Figure 46B, praja2$_{530–570}$ and praja2$_{550–610}$ peptides bind the C-terminus domain of TBC1D31 with micromolar affinity (K_D 37 μM and K_D 80 μM, respectively), whereas no binding was observed with the praja2$_{590–630}$ peptide. This finding suggested that the praja2550–570 segment binds TBC1D31, confirmed by the fact that the elimination of praja2 residues 550–570 (Δ550-570) drastically reduced the binding to GFP-TBC1D31 (Figure 46C).

Docking and MD studies were performed to identify the molecular basis of the praja2 and TBC1D31 interaction.

First, the 3D structures of the praja2$_{550–570}$ and TBC1D31$_{941–970}$ (Figures 46B and D, respectively) segments were generated using the threading approach implemented in the I-TASSER website. The structure of TBC1D31$_{941–970}$ generated by the web-server resulting as an α helix with a kink at the level of Gln941 was confirmed by the CD spectra which showed that the TBC1D31$_{941–970}$ domain assumed a partial helical structure (Figures 46E and 50B).

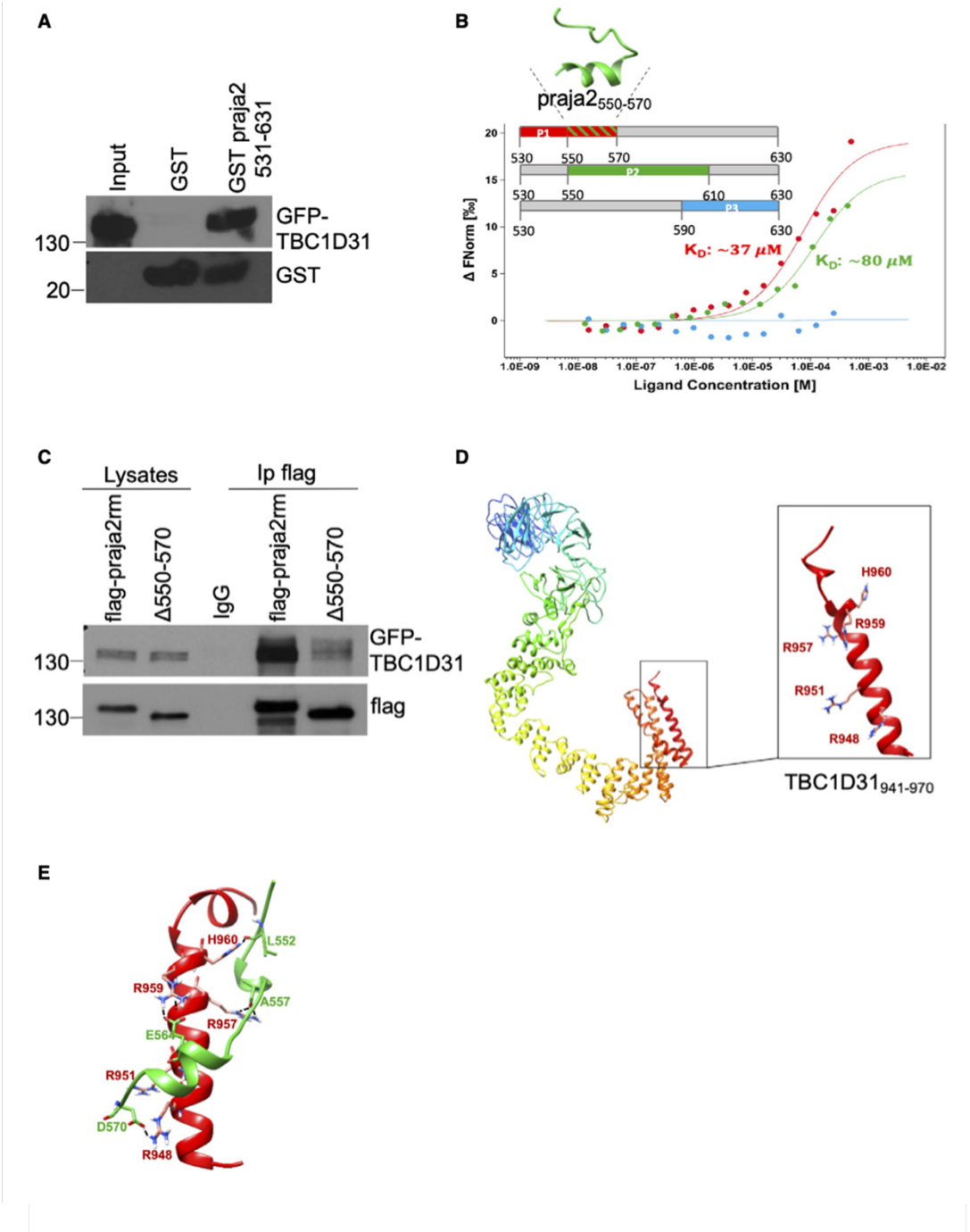


Figure 46. ***Modelling TBC1D31/praja2 binding in vitro and in silico.*** (A) Lysates from HEK293 cells expressing GFP-TBC1D31 were subjected to pull down assay with GST and GST-praja2$_{531–631}$ polypeptides. (B) MST signal (normalized fluorescence) of P1 (*red* curve), P2 (*green* curve) and P3 (*cyan* curve) plotted against TBC1D31, at increasing concentrations of peptides. The threading modelled structure of the overlapping binding segment of praja2 (praja2$_{550–570}$) is shown. (C) Co-immunoprecipitation of GFP-TBC1D31 and flag-praja2 ring mutant (flag-praja2rm) or praja2$_{\Delta550–570}$. (D) Threading modelled structure of TBC1D31, with a zoom of its C-terminus. Mutated residues are highlighted in stick coloured by atom type. (E) MD

derived binding mode of praja2$_{530–570}$ (*green* cartoon) to the C-terminal region of TBC1D31 (*red* cartoon). Credits for the figure go to Ref.[204]

A two-step docking procedure was then performed. First, docking calculation was made using AutodockVina software (Figure 47A).[205] The resulting binding pose was then subjected to a refinement using FlexPepDock webserver,[206] (Figures 46E and 47A) which also gave the plot of the energy landscape (kcal/mol) sampled by Rosetta FlexPepDock (Figure 47B). The funnel-like shape of the docking scores suggests the global minimum as a near-native complex.

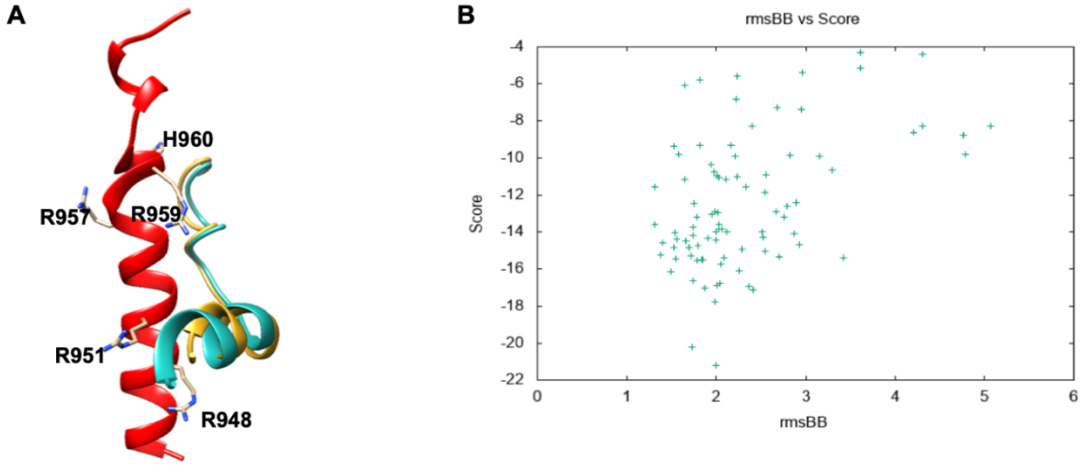


Figure 47. ***Binding mode of praja2 to TBC1D31.*** (A) Superimposition of praja2$_{550-570}$ binding mode to TBC1D31$_{941–970}$ (*red* cartoon) obtained with AutodockVina (*gold* cartoon) and the resulting refinement of FlexPepDock (*green* cartoon). (B) Plot of the energy landscape (kcal/mol) sampled by Rosetta FlexPepDock. Credits for the figure go to Ref.[204]

The docking procedure was followed by classic 2 μs MD simulations, in order to further validate the binding mode. The analysis of the MD trajectory reported a binding mode represented by three main clusters (Figure 48A and B). The binding is mainly driven by the cation-π and ionic interactions between the arginine-rich stretch Arg957-Arg961 (Arg-Ala-Arg-His-Arg) of TBC1D31 and the praja2 stretch Phe553-Asp558

and Glu564. Moreover, the residues Arg948 and Arg951 of TBC1D31 made discontinuous interactions with the praja2 Asp570 residue. As shown by the RMSD average plot computed on the praja2 protein backbone (Figure 48C), the binding mode reaches stability after 400 ns to then be stable until the end of the simulation.

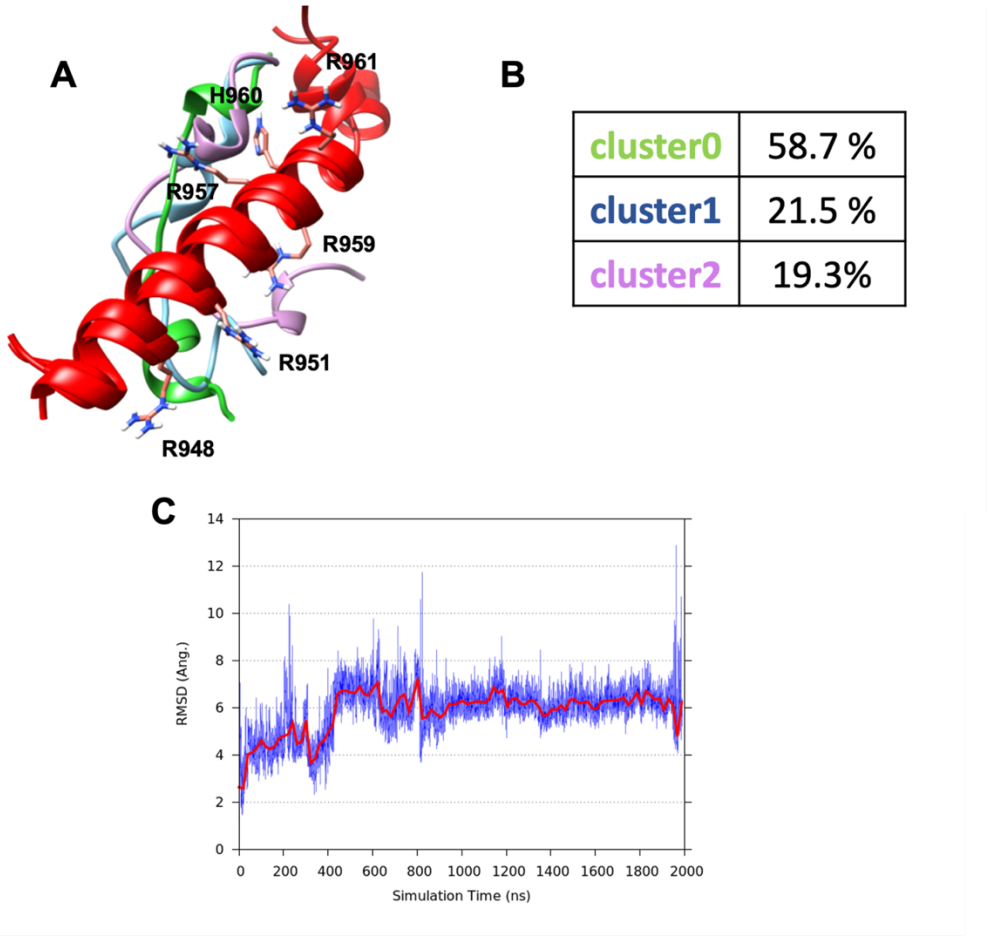


Figure 48. ***Molecular dynamics simulation of the complex praja2/TBC1D31.*** (A) Superposition among the three most populated praja2$^{550\text{-}570}$ clusters (*green*, *cyan* and *pink* cartoon, respectively) with TBC1D31^{wt} (*red* cartoon). (B) % of the population in the three main clusters. (C) Plot of praja2$^{550\text{-}570}$. RMSD average computed on the protein backbone.

In order to validate the proposed binding mode, two different mutants of the C-terminal TBC1D31 peptide were designed: a TBC1D31ADA triple mutant (R957A, R959D and H960A) peptide and a TBC1D31AA double-mutant (R948A and R951A)

peptide. 550 ns of MD simulations were performed and, in both cases, the binding is not stable, as shown by the RMSD plots, in which 14 Å in the RMSD average calculation were reached (Figures 49C and D).

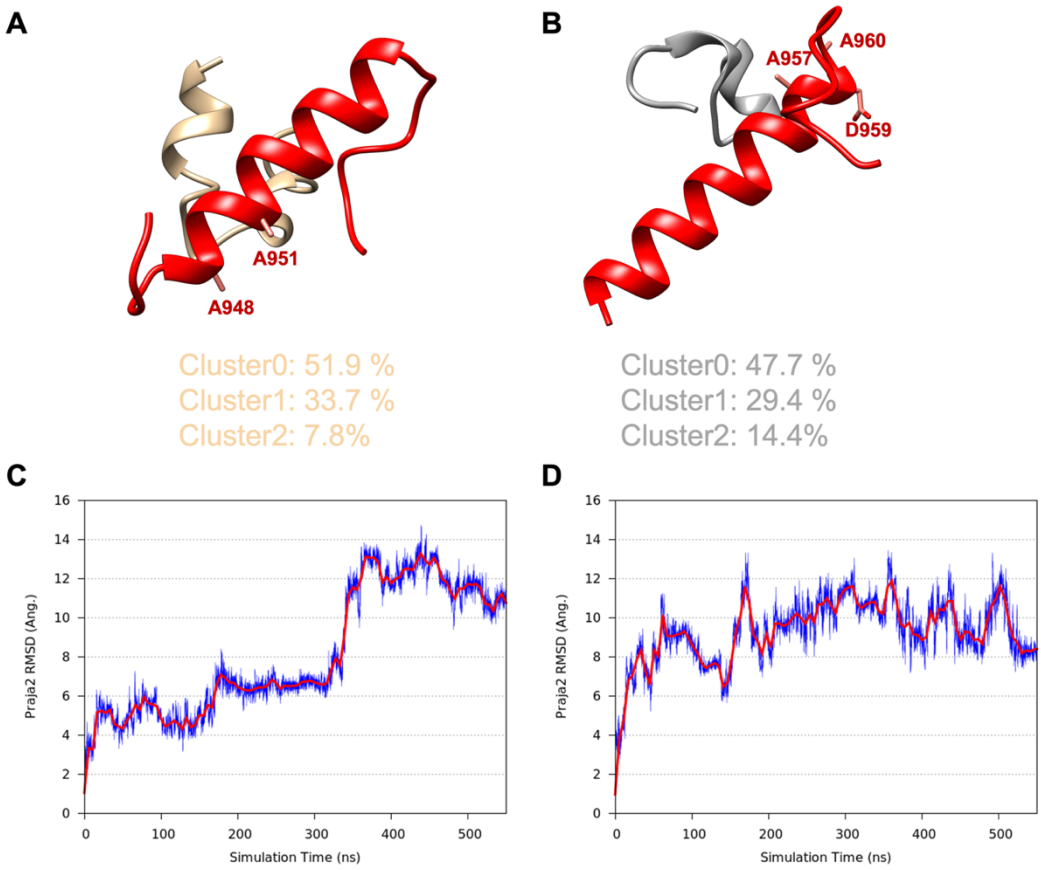


Figure 49. ***Centroids retrieved from molecular dynamics simulation of the complex praja2/TBC1D31.*** (A) Binding mode of praja2$_{550\text{-}570}$ (*tan* cartoon) to TBC1D31$_{AA}$ double-mutant (R948A and R951A) peptide (*red* cartoon)**.** (B) Binding mode of praja2$_{550\text{-}570}$ (*grey* cartoon) to TBC1D31$_{ADA}$ triple mutant (R957A, R959D and H960A) peptide (*red* cartoon)**.** (C and D) RMSD average plot computed on the praja2$_{550\text{-}570}$ protein backbone.

The interaction between praja2 and the two mutants was also studied through pharmacological assays. In particular, microscale thermophoresis experiments showed that the interaction between praja2$_{530–570}$ and TBC1D31 double mutant was preserved, even if the K_D resulted decreased a lot, whereas the interaction was almost abolished

with TBC1D31^{ADA} (Figure 50A). In addition, CD spectra showed a partial helical structure for both mutant peptides, without any appreciable difference with wild-type (Figure 50B), supporting the MD-derived hypothesis of a specific role of residues Arg957, Arg959 and His960 of TBC1D31 in praja2 binding activity.

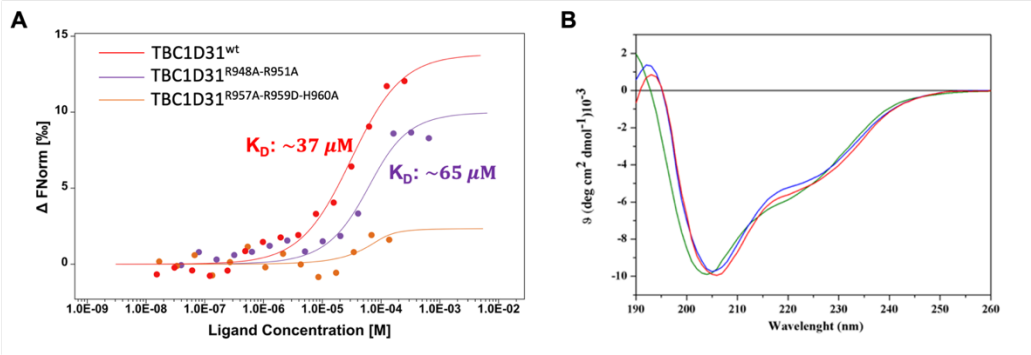


Figure 50. ***Modelling TBC1D31/praja2 binding in vitro.*** (A) MST signal of P1 plotted against increasing concentrations of TBC1D31 peptides: wild-type (*red* curve), R948A-R951A (*violet* curve) and R957-R959D-H960A (*orange* curve). (B) Far-UV CD spectra of: TBC1D31$_{wt}$ (*green* line), TBC1D31$_{AA}$ (*blue* line) and TBC1D31$_{ADA}$ (*red* line). Credits for the figure go to Ref.[204]

The study was also enriched by *in vivo* experiments.

To further demonstrate the role of TBC1D3 in ciliogenesis, an *in vivo* analysis in the Medaka fish model system was conducted using gene knock-down, gene overexpression and rescue experiments. In particular, morpholino, an oligomeric molecule used in molecular biology to modify gene expression, was designed against the medaka orthologist TBC1D31. It has been seen that from stage (St.) 24 onwards, depletion of Ol-TBC1D31 caused a delay in embryonic development and evident embryonic morphological abnormalities, causing microcephaly, microphthalmia, pigmentation defects and pericardial edema. Instead, it was found that injecting human TBC1D31 mRNA into morphants rescued the entire phenotype, thus demonstrating that the function of TBC1D31 is critical for the development of the embryo (Figure 51A).

To determine whether the Ol-TBC1D31 knock-down phenotype was indeed related to abnormal ciliogenesis, a whole-mount immunostaining assay with anti-acetylated α-tubulin was performed to target the biogenesis of cilia on the apical surface of the cells of the neural tube at St.24-26 (2 days after fertilization) of developing Medaka embryos. The result was a significant reduction in cilium length in a large percentage of embryos with Mo-TBC1D31 morphant (Figures 51B and C). Human TBC1D31 was then co-injected, which is not recognized by morpholino Ol-TBC1D31, resulting in an increase in cilium length.

Subsequently, the possible correlation of the Ol-TBC1D31 morphant phenotype with the abnormal phosphorylation/ubiquitylation of OFD1 was investigated, since, if TBC1D31 directly controls OFD1 ubiquitylation *in vivo*, the overexpression of the $OFD1_{S735A}$ mutant should induce a phenotype similar to due to alterations in ciliogenesis. Hence, it was found that $OFD1_{S735A}$ injection caused defects in ciliogenesis. In this case, co-injection of wild-type OFD1 with Ol-TBC1D31 morpholino was seen not to save cilium length and Medaka embryogenesis (Figures 51A-C).

To further support TBC1D31-mediated regulation of OFD1, $OFD1_{S735A}$ was also co-injected with Mo-Ol-TBC1D31, resulting in similar defects in embryo development and ciliogenesis, in which cilia were barely sketched (Figures 51A-C).

To further support TBC1D31-mediated regulation of OFD1, OFD1S735A was also co-injected with Mo-Ol-TBC1D31, resulting in similar defects in embryo development and ciliogenesis, in which cilia were barely sketched (Figures 51A-C). Therefore, after finding that most of the changes in ciliogenesis caused by Ol-TBC1D31 KD are related to the altered phosphorylation and proteolysis of OFD1 by praja2 activity, we co-injected the $OFD1_{S735D}$ mutant, which mimics the phosphorylated form of OFD1, combined with the dominant negative variant of human

praja2 (hpraja2rm), in order to re-establish ciliogenesis in Mo-Ol-TBC1D31 morphants and the phenotype of the rescue larva. The result showed that OFD1$_{S735D}$/hpraja2rm injection was sufficient to save the normal development of the larva (~20%) and partially recovered ciliogenesis defects in a substantial percentage of Mo-Ol-TBC1D31 morphants.

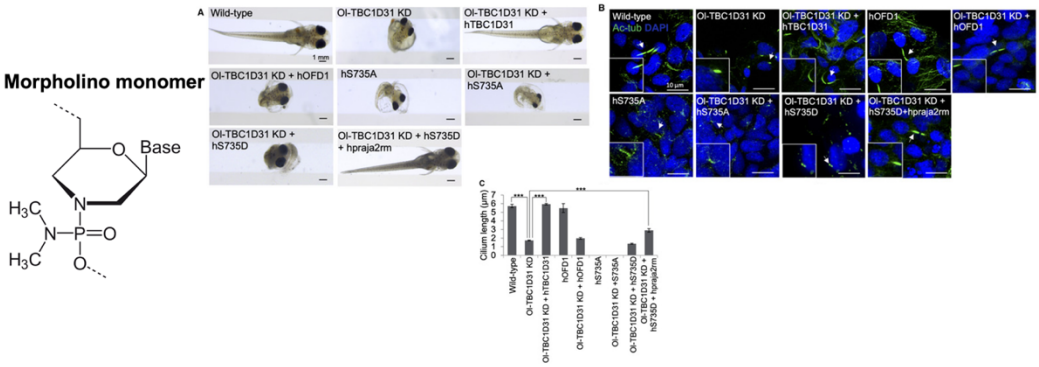


Figure 51. ***TBC1D1 and PKA/OFD1 pathway controls Medaka fish development.*** TBC1D1 and PKA/OFD1 pathway controls Medaka fish development. (A) Stereo-microscopic images of wild-type, Ol-TBC1D31 KD, Ol-TBC1D31 KD + hTBC1D31, Ol-TBC1D31 KD + wild-type hOFD1, hOFD1$_{S735A}$, Ol-TBC1D31 KD + hOFD1$_{S735A}$, Ol-TBC1D31 KD + hOFD1$_{S735D}$ and Ol-TBC1D31 KD + hOFD1$_{S735D}$ + hpraja2rm injected Medaka larvae, at stage 40. (B) Confocal images of cilia of the neural tube cells in the wild-type, Ol-TBC1D31 KD, Ol-TBC1D31 KD + hTBC1D31, wild-type hOFD1, Ol-TBC1D31 KD + wild-type hOFD1, hOFD1$_{S735A}$, Ol-TBC1D31 KD + hOFD1$_{S735A}$, Ol-TBC1D31 KD + hOFD1$_{S735D}$ and Ol-TBC1D31 KD + hOFD1$_{S735D}$ + hpraja2rm stained with anti-acetylated α-tubulin antibody (green) and DAPI (blue). (C) In the graph is reported the cilia length in wild-type, Ol-TBC1D31 KD, Ol-TBC1D31 KD + hTBC1D31, wild-type hOFD1, Ol-TBC1D31 KD + wild-type hOFD1, hOFD1$_{S735A}$, Ol-TBC1D31 KD + hOFD1$_{S735A}$, Ol-TBC1D31 KD + hOFD1$_{S735D}$ and Ol-TBC1D31 KD + hOFD1$_{S735D}$ + hpraja2rm. The data are expressed as mean value ± SE of twelve independent experiments. Student's *t* test, ***$P \leq 0.001$. Credits for the figure go to Ref.[204]

Thus, considering that OFD1 is a pathology related to primary cilia dysfunction linked to the lack of degradation of the OFD1 protein, the role of the praja2/PKA/OFD1/TBC1D31 complex has been studied. First, pharmacological assays demonstrated that TBC1D31 binds to and targets praja2 at the centrosome and

that the minimal core domain on praja2 that binds TBC1D31$_{941\text{-}970}$ was praja2$_{530–630}$. The binding mode was studied through molecular docking calculations and validated through 2 μs of MD simulations. Subsequently, two TBC1D31 mutants (R948A and R951A/R957A, R959D and H960A) were constructed and analyzed through pharmacological assays and *in silico* studies to further validate the binding mode, finding that the interaction between TBC1D31 and praja2 is lost in the two mutant mutants.

Finally, this study validated the role of TBC1D31 as a molecular scaffold located at the centrosome and its essential role in the praja2/PKA/OFD1 molecular network necessary for the correct ciliogenesis and development of Medaka fish.

This study has been published in The EMBO Journal (ref: Senatore, E.; Chiuso, F.; Rinaldi, L.; Intartaglia, D.; Delle Donne, R.; Pedone, E.; Catalanotti, B.; Pirone, L.; Fiorillo, B.; Moraca, F.; Giamundo, G.; Scala, G.; Raffeiner, A.; Torres-Quesada, O.; Stefan, E.; Kwiatkowski, M.; van Pijkeren, A.; Morleo, M.; Franco, B.; Garbi, C.; Conte, I.; Feliciello, A. The TBC1D31/praja2 complex controls primary ciliogenesis through PKA-directed OFD1 ubiquitylation. *EMBO J.* **2021**, *40*, e106503.)

3.2 Discovery of pelargonidin as a potential inhibitor of the SARS-CoV-2 interaction and angiotensin converting enzyme 2 (ACE2)

3.2.1 Introduction

Another project in which I was involved in collaboration with Professor Stefano Fiorucci of the University of Perugia, once again saw as its object the search for molecules capable of inhibiting the SARS-CoV-2/ACE2 interaction. Considering that the expression of ACE2 is induced in response to inflammation and that, therefore, its expression in the colon is upregulated in patients with inflammatory bowel disease (IBD), it has been thought that the entry of SARS-CoV-2 in the human body could have caused intestinal inflammation.[61] Unfortunately, however, the mechanisms underlying ACE2 expression in the intestine are still poorly understood.

Therefore, targeting the interaction of Spike RBD with ACE2 could have the potential for treating COVID-19 and the need, linked to the current pandemic situation due to COVID-19, to find anti-SARS-CoV-2 therapies, has led us to study the action of several agents in targeting the SARS-CoV-2/ACE2 interaction, including several monoclonal antibodies.

We also investigated whether modulation of ACE2 expression could reduce SARS-CoV-2 entry into target cells.[57,58,59,61,66,207]

ACE2 is differently expressed among human tissues, like in the lung and gastrointestinal system - particularly in the duodenum, jejunum, ileum and colon.[208,209,210,211,212,213] In particular, in ulcerative colitis patients with active inflammation, an elevated expression of ACE2 has been noted which normalizes after therapy with anticytokines. The effect of anticytokine therapy in patients with

inflammatory bowel disease affected by COVID-19 was then studied, noting lower morbidity than in the general population.[214]

Thus, we undertook the study of the aryl hydrocarbon receptor (AhR), a transcription factor activated by the ligand and belonging to the basic superfamily of helix-ring-helix (bHLH)/PerARNT-Sim (PAS). AhR plays the role of xenobiotic sensor and metabolites of the food/intestinal microbiota are known to represent physiological ligands.[215] In the absence of a ligand, AhR resides in the cytoplasm as a component of a chaperone complex.[216] AhR is expressed by innate/adaptive immunity cells[217] and intestinal epithelial cells where it plays the role of maintaining the integrity of the intestinal barrier, thus also regulating the inflammatory state of the gastrointestinal tract.[218,219]

In the previous work by my research group, the activity of pelargonidin, a water-soluble anthocyanidin widely spread in nature as glycosylated derivatives beneficial for human health,[220] was demonstrated,[221,222,223] but with poor systemic bioavailability and unclear mechanisms of action, as an AhR ligand *in vitro* and attenuating intestinal inflammation in an AhR-dependent manner.[224]

In this work, we investigated the activity of natural pelargonidin in regulating ACE2 expression in the colon in models of intestinal inflammation caused by the exposure of wild type and AhR -/- mice to high caloric intake and intestinal irritants.

3.2.2 Results and Discussion

Pharmacological assays were performed in the laboratory of Professor Stefano Fiorucci at the Department of Medicine and Surgery of the University of Perugia.

First, the agonistic activity of pelargonidin on AhR was assessed using a Luciferase reporter assay using HepG2 cells, transiently transfected with a AhR reporter gene cloned upstream to the *LUCIFERASE*. HepG2 was incubated with 5 nM TCDD, an

AhR agonist as control, or vehicle (0.1% v/v DMSO) in the presence of pelargonidin (10–50 μM) for 18 hours. Pelargonidin was found to be effective in transactivating AhR with an EC_{50} of 12 μM (Figure 52).

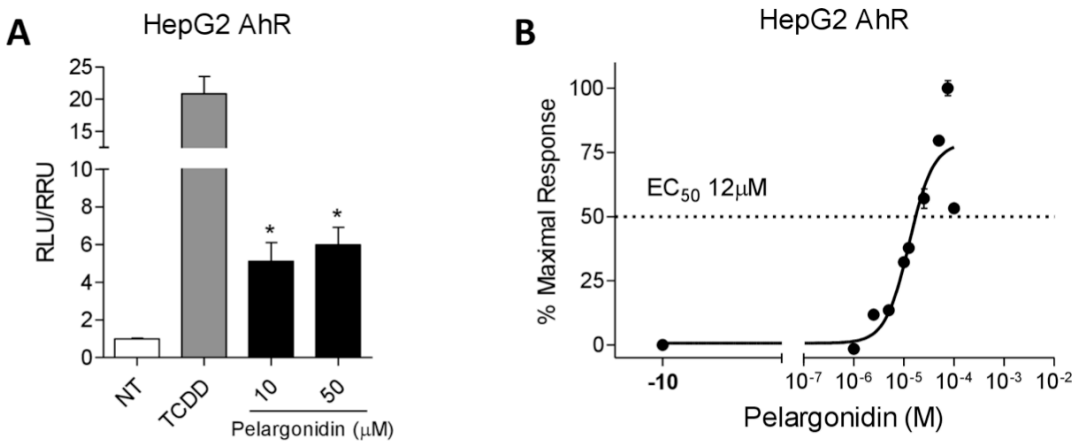


Figure 52. ***The activity of pelargonidin toward AhR.*** (A) Fold of induction of luciferase activity in cells transfected with *AhR* reporter gene and incubated with TCDD (5 nM) or pelargonidin (10–50 μM). (B) Dose-response curve of pelargonidin to evaluate AhR activation; cells were stimulated with increasing concentrations of pelargonidin 1 μM to 100 μM. Results are expressed as mean ± standard error. ∗ $p < 0.05$ versus not treated cells (NT). Credits for the figure go to Ref.[225]

Subsequently, the intestinal anti-inflammatory activity of pelargonidin was evaluated in a mouse model of colitis, first determining the effective dose by administering a dose of 1, 5 or 10 mg/kg in a mouse model of TNBS-induced colitis (Figure 53). Clinical data and analysis of the macroscopic and microscopic characteristics of the colon have shown for doses of 1 and 5 mg/kg of pelargonidin a dose-dependent effect with the lowest dose exerting only mild beneficial effects. Instead, at the administration of 10 mg/kg, intermediate effects between the dose of 1 and 5 mg/kg were observed. Hence, the effective dose to decrease intestinal inflammation is 5 mg/kg. Then, using this assay, we investigated the mechanism of action of pelargonidin on intestinal immunity, inducing inflammation of the colon by

administering TNBS to Ahr +/+ and Ahr -/- mice. Unfortunately, this study could not be performed in the Ahr knock-out group as a mortality rate of 80% was observed, which often occurs in Ahr -/- mice.[224]

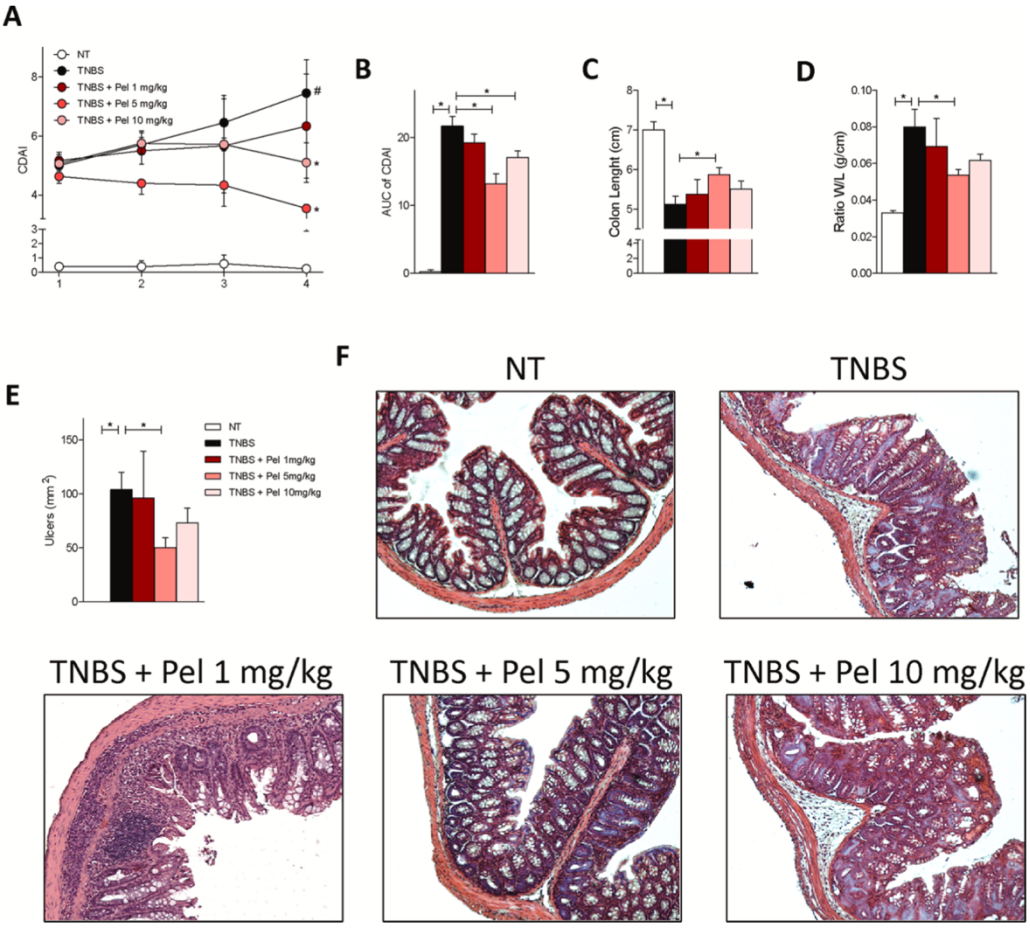


Figure 53. ***Pelargonidin reduces the severity of TNBS colitis in a dose-dependent manner.*** Colitis was induced by TNBS. After induction of colitis the mice were treated daily with pelargonidin (1, 5 or 10 mg/Kg) or vehicle. The disease was monitored by daily evaluation of (A) changes in colitis disease activity index (CDAI) and by evaluation of the (B) Area Under the Curve (AUC). At the end of the experiment, we evaluated (C) colon length (cm) and (D) ratio of colon weight/colon length (g/cm). (E) Area of ulcers and (F) H&E staining of colon sections (10× magnification) and Histological Score from each experimental group. Results are expressed as mean ± SEM (n = 5–7); In graph A # TNBS Vs NT; * TNBS + Pel Vs TNBS; in all graphs # and * $p < 0.05$. Credits for the figure go to Ref.[225]

On the other hand, administration of pelargonidin in wild-type mice caused a TNBS-induced reversal of intestinal inflammation, as measured by lower body weight

loss, lower CDAI, and evaluation of the macroscopic and microscopic features of the colon (Figures 54A-D).

Therefore, in order to study the expression of ACE2 protein and mRNA in the colon, we performed immunohistochemistry and qPCR assays which confirmed that the induction of colitis by TNBS administration increased ACE2 and mRNA expression as found in patients with IBD and that the administration of pelargonidin was able to contrast the increase in ACE2 expression (Figures 54E and F).[212]

We further confirmed the agonist activity of pelargonidin towards the AhR receptor by evaluating *in vivo* the expression of one of the main AhR target genes, *Cyp1a1*, obtaining that the administration of the flavonoid regulates the expression of *Cyp1a1* in the colon (Figure 54G). We then analyzed the colonic expression of Mas and cytokines (Figures 54F-J), finding that exposure to TNBS increased the regulation of pro-inflammatory cytokines (Figures 54I-J) and, in particular, Il-1β was upregulated approximately 20 times in the colon of TNBS-treated mice compared to naïve mice (Figure 54I), while the Tnf-α mRNA is increased by 5 times (Figure 54J). Conversely, Mas mRNA expression is reduced (Figure 54H). The administration of pelargonidin to TNBS mice showed an induction of Tgf-β expression (Figures 54I-K) and a strong reduction of ACE2 mRNA expression (Figure 54F). The correlation analysis between Il1β/ACE2 and Tnf-α/ACE2 confirmed a statistically significant correlation between the expression of the two pro-inflammatory cytokines and ACE2 (P-value of Il-1β/ACE2 = 0.0244; P-value of Tnf-α/AcE2<0.0001) (Figures 54L and M).

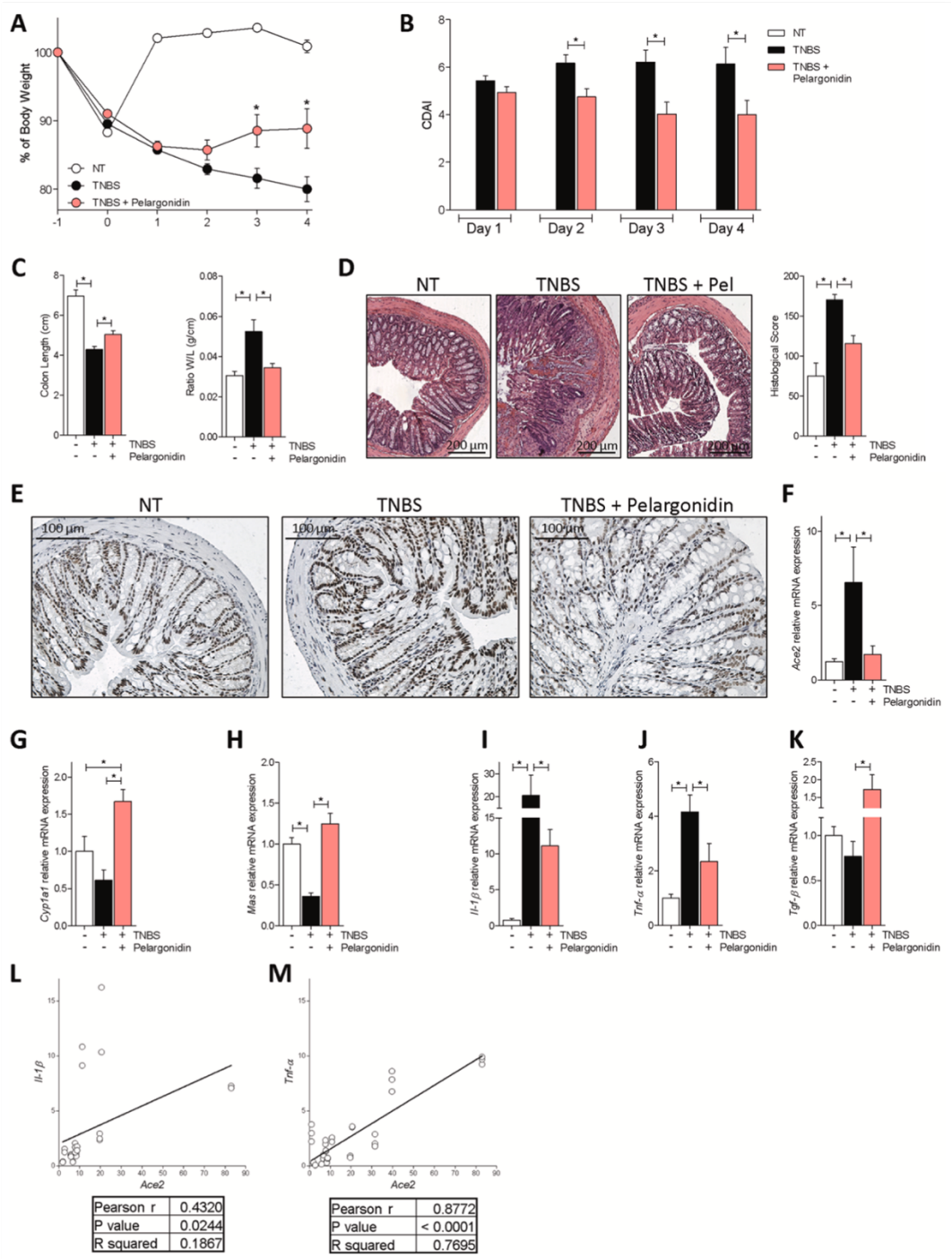


Figure 54. ***Pelargonidin effect on acute colitis.*** Colitis was induced by TNBS. After induction of colitis the mice were treated daily with pelargonidin (5 mg/Kg) or vehicle. The disease was monitored by daily evaluation of (A) changes in body weight (%), (B) colitis disease activity index (CDAI), (C) colon length (cm) and ratio of colon weight/colon length (g/cm). (D) H&E staining of colon sections (10× magnification) and Histological Score from each experimental group. (E) The figure shows immunohistochemistry representative images of the colon of one mouse for each experimental group stained with anti-ACE2 Ab (20x magnification). RNA extracted from the colon was used to evaluate, by quantitative real-time PCR, the relative mRNA expression of (F) *ACE2,* (G) *Cyp1a1*, (H) *Mas*, (I) *Il-1β*, (J) *Tnf-α* and (K) *Tgf-β*.

Values are normalized relative to *Gapdh* mRNA. The values are expressed relative to those of the control group (NT) which are arbitrarily set to one. Correlation graph of *ACE2* mRNA expression and (L) *Il-1β* (M) *Tnf-α.* Results are expressed as mean ± SEM (n = 7–12); In graph A * TNBS + pelargonidin Vs TNBS; in all graphs * $p < 0.05$. Credits for the figure go to Ref.[225]

Due to the inability to analyze the effects of pelargonidin treatment in the colitis model in Ahr -/- mice due to the high mortality, we tested the immunomodulatory activity of pelargonidin towards AhR on murine macrophages purified from the spleen of Ahr mice +/+ and Ahr -/-. Then, after exposing the macrophages to pro-inflammatory stimuli (LPS + IFN-γ), pelargonidin was administered. Stimulation with LPS + IFN-γ induced an increase in pro-inflammatory cytokine production (Figures 55A and B) and a reduction in Tgf-β production (Figure 55C) in macrophages. The up-regulation of Il-6 was much higher in Ahr -/- macrophages than in the wild type (Figure 55B). Treatment with pelargonidin induced an anti-inflammatory effect only in wild-type macrophages, confirming, once again, the agonist activity of pelargonidin towards the AhR receptor.

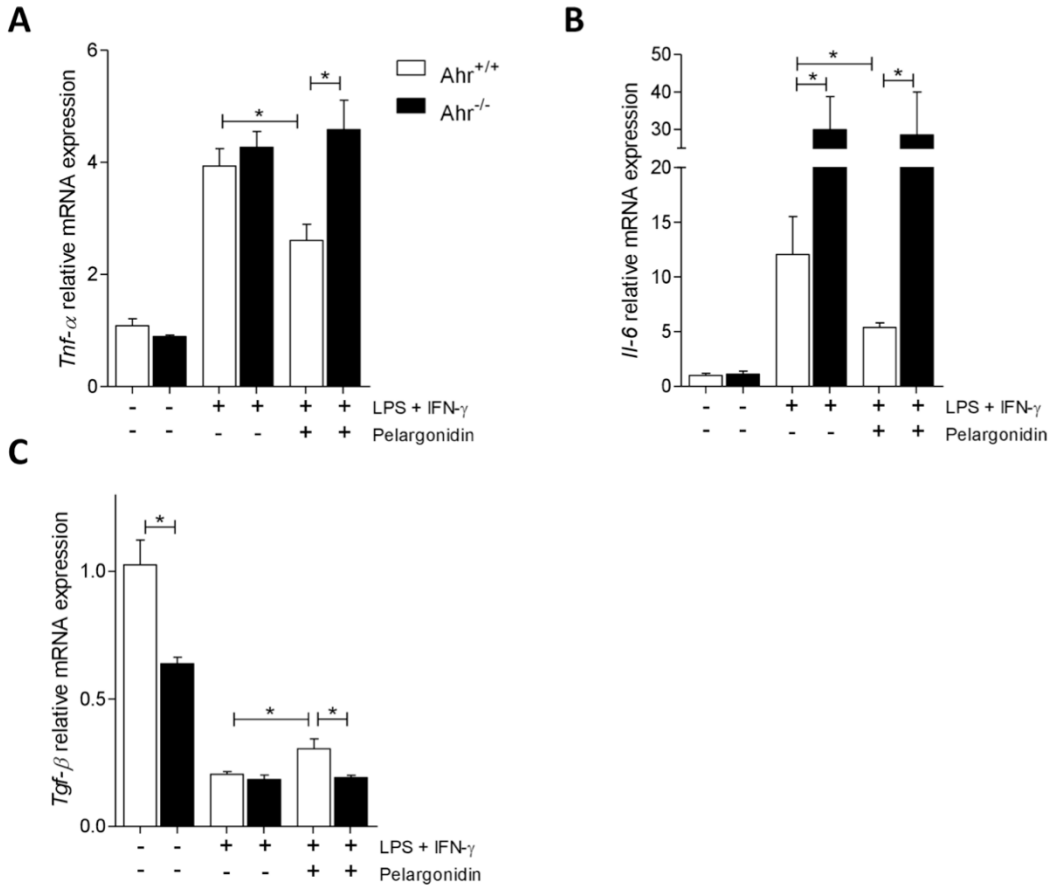


Figure 55. ***Pelargonidin exerts immunomodulatory effects through AhR***. Spleen macrophages purified from *AhR*$^{+/+}$ and *AhR*$^{-/-}$ mice were activated *in vitro* with LPS (5 ng/mL) in combination with IFN-γ (20 ng/mL) alone or plus pelargonidin (20 μM) for 16 h. At the end of stimulation, the relative mRNA expression of pro-inflammatory cytokines (A) *Tnf-α* and (B) *Il-6*, and anti-inflammatory cytokines (C) *Tgf-β*, was evaluated by Real-Time PCR. Values are normalized relative to Gapdh mRNA and the values are expressed relative to those of the control group which are arbitrarily set to one. Results are the mean ± SEM (n = 6); * $p < 0.05$. Credits for the figure go to Ref.[225]

Considering that obesity is a well-defined risk factor for the development of severe COVID-19, we also tested the effects of pelargonidin in a mouse model of a high-fat diet.[226,227]

In this experiment, the Ahr +/+ and Ahr -/- mice were treated with an exposed chronic caloric diet,[224] through a diet enriched in cholesterol and fructose (HFD-F) for 8 weeks. Starting from day 8, an experimental group of mice for each genotype was treated daily with pelargonidin. mice fed a high-fat diet gained body weight compared

to mice fed a normal diet and had a significantly higher BMI at the end of the study with no significant differences between the two genotypes (Figures 56A and C). Administration of 5 mg/kg/day of pelargonidin for 7 weeks protected against body weight gain and resulted in a lower BMI than mice feeding only one HFD-F. No beneficial effects of pelargonidin were observed in Ahr - - mice (Figures 56A and C). We, therefore, focused our attention on intestinal inflammation. Histological analysis of the colon showed no morphological abnormalities, except an increase in intestinal permeability in mice exposed to HFD-F (Figures 56D and E), restored following treatment with pelargonidin in wild-type mice, but in Ahr-/- mice which, on the other hand, showed worsening in permeability compared to Ahr +/+ mice (Figures 56D and E).

Since the correlation between ACE2 and colon inflammation in IBD patients is known,[212] we studied the expression of ACE2 and MAS in the colon and compared them with the colonic profiles of various cytokines. HFD-F increased ACE2 expression and reduced Mas expression in Ahr +/+ and Ahr -/- mice, while it increased colonic expression of Il-1β and Tnf-α and downregulated mRNA of Tgf-β. Pelargonidin showed its inflammatory properties by decreasing the expression of both Il-1β and Tnf-α, but only in Ahr +/+ mice (Figures 56F-J). On the other hand, a reduced expression of ACE2 was also observed in wild type mice (Figures 56F-J).

The correlation between ACE2 expression and pro-inflammatory cytokines was also studied in this mouse model with mild inflammation in the colon, obtaining positive results in both Ahr +/+ and Ahr -/- mice (Figures 56K and L).

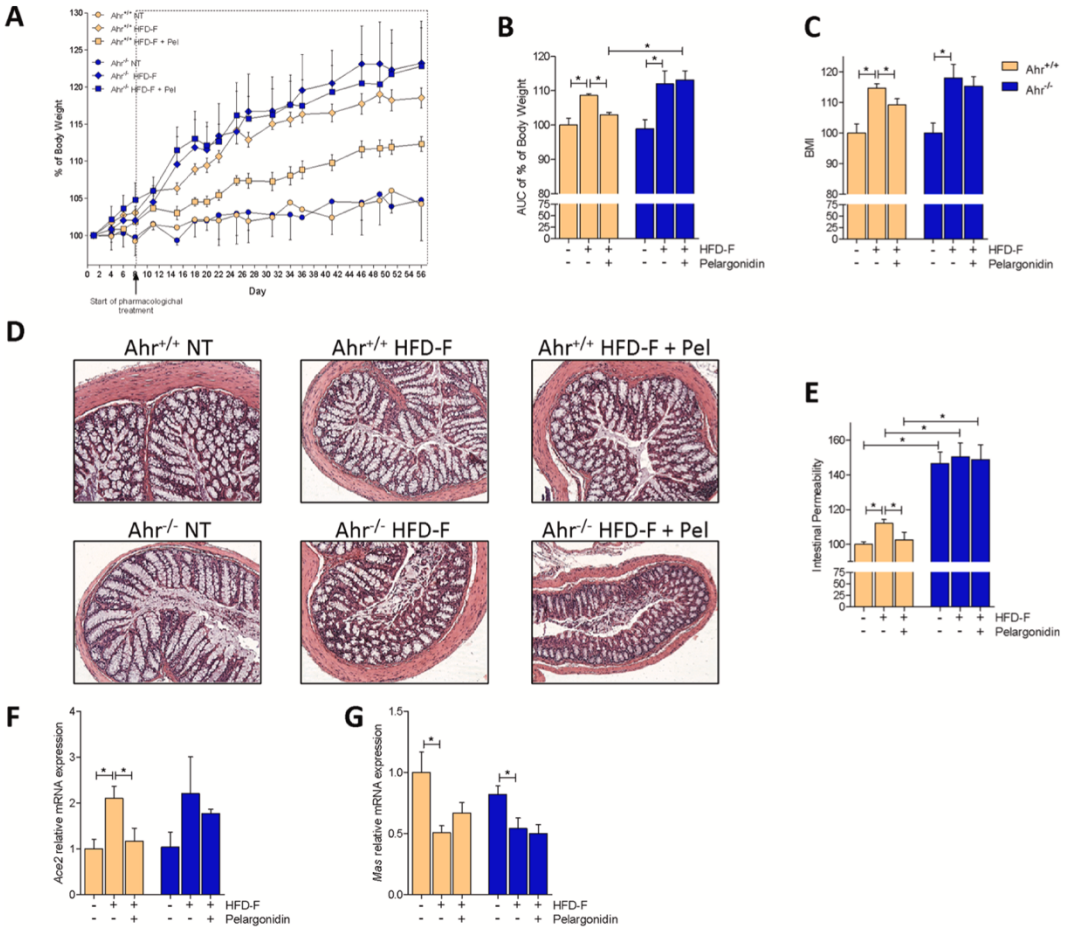


Figure 56. ***The benefit of pelargonidin administration in a mouse model of NASH is lost in Ahr$^{-/-}$ strains.*** C57BL/6 male mice, (*Ahr*$^{+/+}$) and their congenic littermates *Ahr* knock out (*Ahr*$^{-/-}$) were fed a normal chow diet (NT) or a high fat diet with fructose in water (HFD-F). (A) Changes in body weight (%) assessed for 56 days. (B) Areas under curve (AUC) of body weight expressed in arbitrary units. (C) Body Mass Index (BMI) is calculated at the end of the study as the ratio between body weight (g) and body length^2(cm^2). (D) Histological sections, performed with H&E staining on the colon (10x magnification) of *Ahr*$^{+/+}$ and *Ahr* $^{-/-}$ mice for each experimental group. (E) Intestinal permeability was measured after 4 weeks of diet with FITC-dextran administration. At the end of the experiment the total RNA extracted from the colon was used to evaluate, by quantitative real-time PCR, the relative mRNA expression of (F) *Ace2*, (G) *Mas*, (H) *Il-1β*, (I) *Tnf-α*, (J) *Tgf-β*. Values are normalized relative to *Gapdh* mRNA. The values are expressed relative to those of the control group (NT) which are arbitrarily set to one. Correlation graph of Ace2 mRNA expression and (K) *Il-1β* (L) *Tnf-α*. Results are expressed as mean ± SEM (n = 6–10); * $p < 0.05$. Credits for the figure go to Ref.[225]

Since ACE2 is expressed in intestinal epithelial cells, we then examined whether exposure of Caco-2 cells, a human intestinal epithelial cell line, to TNF-α modulates ACE2 expression. Exposure to TNF-α caused increased expression of the pro-inflammatory genes Il-8, Il-6 and Il-1B by Caco-2 cells (Figures 57B-D) and by ACE2 (≈2,5 times) confirming the close correlation between TNF-α and ACE2 (Figure 57A). Also in this experiment, pelargonidin reversed expressions in a concentration-dependent manner (Figures 57A-D).

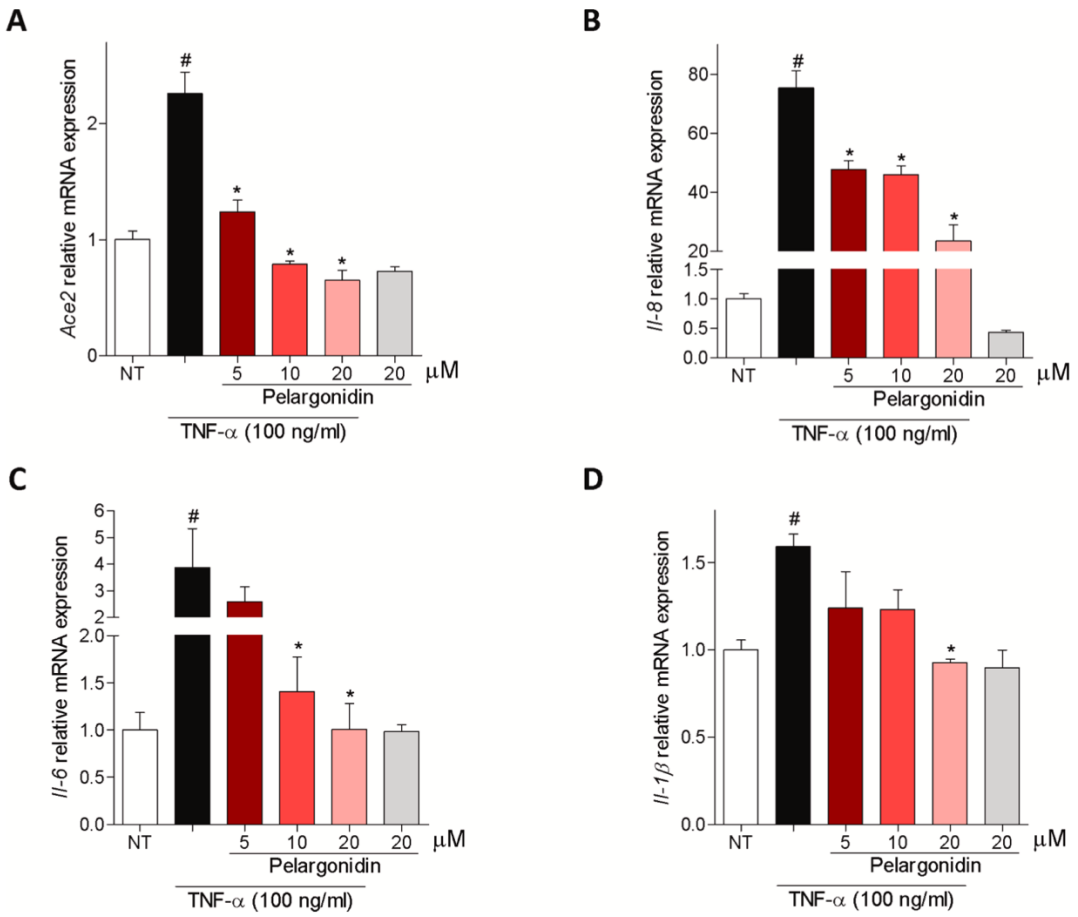


Figure 57. ***Pelargonidin counteracst TNF-α-inflammatory activation on Caco2-cells.*** Caco-2 cells, a human intestinal epithelial cell line, activated with TNF-α 100 ng/ml for 24 h alone or in combination with pelargonidin (5, 10, and 20 μM). At the end of stimulation, the relative mRNA expression of (A) *ACE2*, (B) *Il-8*, (C) *Il-6* and (D) *Il-1B*, was evaluated by Real-Time PCR. Values are normalized relative to Gapdh mRNA and the values are expressed relative to those of the control group (NT) which are arbitrarily set to one. Results are the mean ± SEM (n = 5); # NT Vs TNF-α; * TNF-α Vs TNF-α + pelargonidin; # and * $p < 0.05$. Credits for the figure go to Ref.[225]

To study the molecular mechanism underlying the correlation between TNF-α and ACE2 and the beneficial effect due to the administration of pelargonidin, we purified intestinal epithelial cells (IEC) from the colon of Ahr +/+ and Ahr -/- mice and we stimulated with TNF-α, demonstrating that also in these primary murine intestinal epithelial cells TNF-α induces inflammation by upregulating both genes (Figures 58A and B) and that the absence of the AhR receptor implied greater inflammation induced by TNF-α. After administration of pelargonidin, inflammatory states were restored in purified cells from wild-type mice (Figures 58A and B). In fact, the anti-inflammatory effect of pelargonidin did not manifest itself in Ahr -/- mice lacking the AhR receptor.

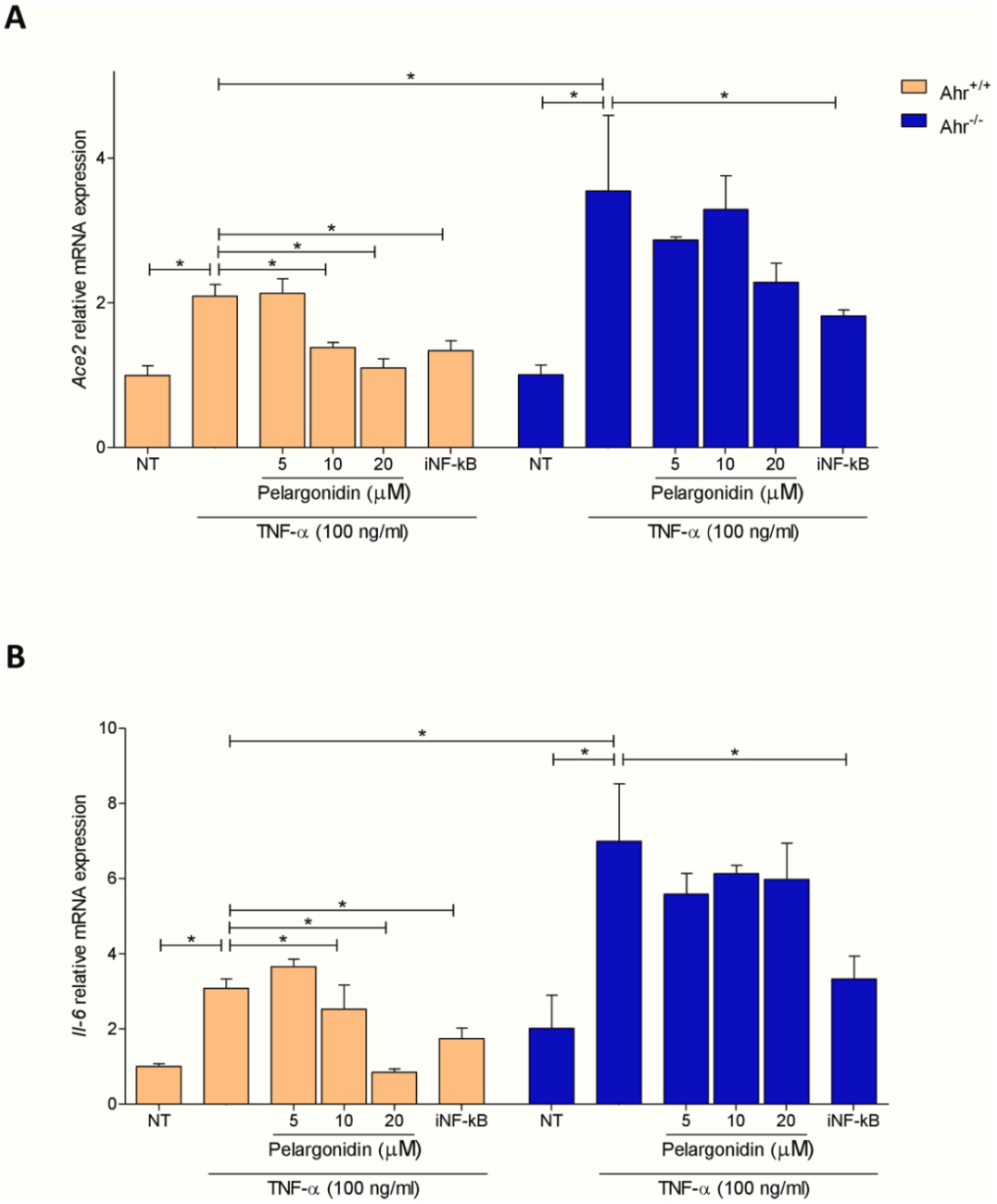


Figure 58. ***TNF-α up-regulates ACE2 expression by NF-kB on intestinal epithelial cells and pelargonidin inhibits this pathway by activating AhR.*** Intestinal epithelial cells were purified from the colon of *Ahr*$^{+/+}$and *Ahr*$^{-/-}$ mice. Intestinal epithelial cells were cultured for 24 h with TNF-α 100 ng/ml and treated with pelargonidin (5, 10, and 20 μM) or with the NF-κB inhibitor (iNF-kB 100 nM). At the end of stimulation, the relative mRNA expression of (A) *ACE2* and (B) *Il-6* was evaluated by Real-Time PCR. Values are normalized relative to *Gapdh* mRNA and the values are expressed relative to those of the control group (*Ahr*$^{+/+}$ NT) which are arbitrarily set to one. are expressed as mean ± SEM (n = 5); * $p < 0.05$. Credits for the figure go to Ref.[225]

Since SARS-CoV-2 enters the host cell thanks to the ACE2 target and the reduction of ACE2 expression could represent an interesting defense mechanism of the host against virus invasion, we also investigated whether pelargonidin was able to reduce the binding of the viral Spike protein on ACE2 through molecular docking studies and *in vitro* experiments (Figure 59).

The ability of pelargonidin to bind several pockets suggested in previous studies was investigated: i) the hydrophobic pockets on the β-sheet core of the RBD,[153] ii) the fatty acid (FA) pocket,[56,228] and, finally, iii) the binding pocket of flavonoids.[229] The Glide software package was used.[117,118] The scores of the best poses showed a marked preference for the results obtained in the FA pocket (best docking score −7.7 kcal/mol), compared to the results obtained in the other pockets, thus suggesting that pelargonidin interacts with the RBD of the Spike protein by binding to the FA pocket.[228] Analysis of the docking pose in the FA pocket (Figures 59A-C) showed that the polyphenolic ring of pelargonidin is in contact with Leu368, Leu387, Phe388, Phe342 and Ile434. Furthermore, Phe377 makes a p-cation interaction with the oxygen of the C ring, while Tyr365 and Tyr369 are engaged in a π–π stacking with the biphenyl ring and the B ring respectively. Finally, the binding mode is further stabilized by the H bond formed between the hydroxyl group of the B ring and the backbone of Ala372.

On the basis of docking calculations results, using a Spike/ACE2 Inhibitor Screening Assay Kit, the ability of pelargonidin to inhibit the Spike/ACE2 interaction was evaluated, obtaining a concentration-dependent inhibition. In fact, at 50 μM, we measured a reduction of the Spike bond on ACE2 of about 40% (Figure 59D).

In the SARS-CoV-2 virus infection test on the Vero E6 cell line, monkey-extracted renal epithelial cells, pelargonidin showed an ability to reduce virus penetration by approximately 70% at the highest concentration of 100 μM (Figure 59E).

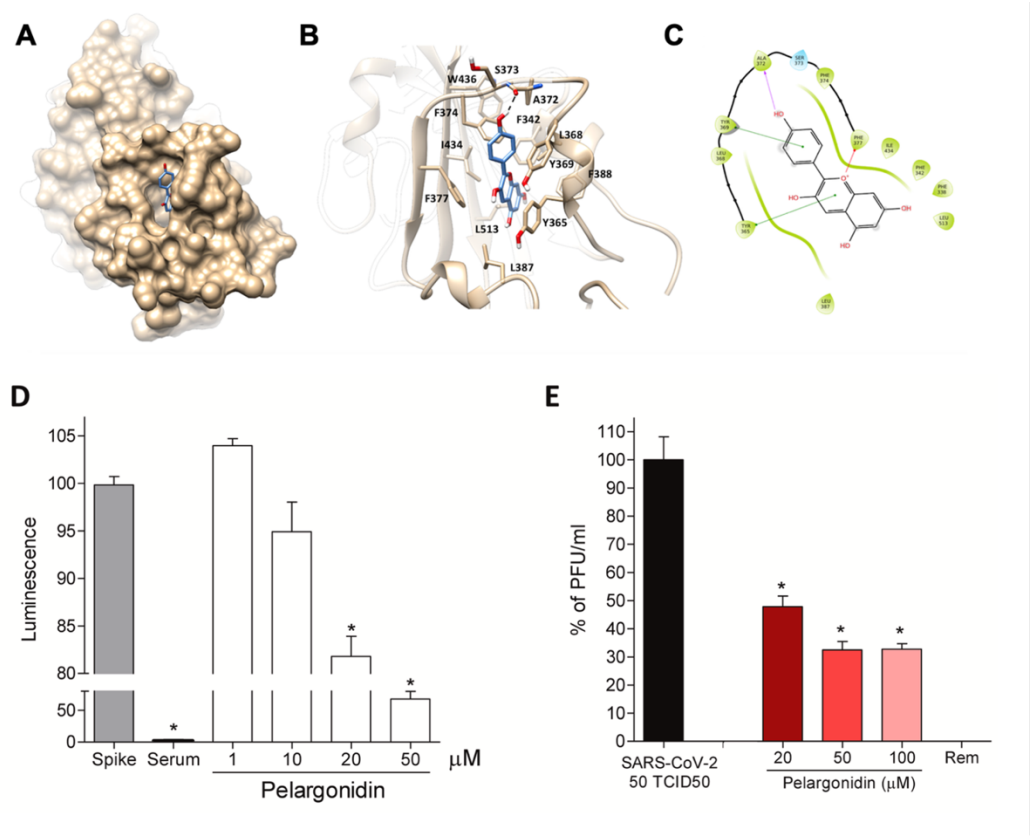


Figure 59. ***Pelargonidin inhibits the binding of the SARS-CoV-2 virus on the host cells.*** (A) Hydrophobic FA binding pocket in a surface representation; (B) Cartoon representation of the binding mode of pelargonidin to SARS-CoV-2 receptor. The ligand is represented as blue sticks, whereas the interacting residues of the receptor are shown in tan and labelled. Oxygen atoms are depicted in red and nitrogens in blue. The receptors are represented as tan ribbons. Hydrogens are omitted for the sake of clarity; (C) Diagram of pelargonidin interaction. (D) SARS-CoV-2 Spike binding to immobilized ACE2; Luminescence was measured using a Fluo-Star Omega fluorescent microplate reader. Pelargonidin was tested at different concentration (1, 10, 20 and 50 µM), to evaluate their ability to inhibit the binding of Spike protein (5 nM) to immobilized ACE2, by using the ACE2:SARS-CoV-2 Spike Inhibitor Screening assay Kit. Results are expressed as mean ± SEM (n = 5); * $p < 0.05$. To confirm the validity of the assay used in this study, we tested plasma samples of post COVID-19 patients as a control. (E) Virus growth in Vero 6E cells analyzed by plaque assay. Pelargonidin was tested at concentration of 20, 50 and 100 µM. To confirm the validity of the assay used in this study, we tested Remdesivir. Credits for the figure go to Ref.[225]

The set of data obtained from the *in silico* study and the *in vitro* tests associated with the down-regulation of ACE2 expression exerted by pelargonidin and anti-inflammatory activity make this flavonoid very interesting in the prevention and treatment of SARS-CoV-2 infection.

This study has been published in Biochemical Pharmacology (ref: Biagioli, M.; Marchianò, S.; Roselli, R.; Di Giorgio, C.; Bellini, R.; Bordoni, M.; Gidari, A.; Sabbatini, S.; Francisci, D.; Fiorillo, B.; Catalanotti, B.; Distrutti, E.; Carino, A.; Zampella, A.; Costantino, G.; Fiorucci, S. Discovery of a AHR pelargonidin agonist that counter-regulates ACE2 expression and attenuates ACE2-SARS-CoV-2 interaction. *Biochem Pharmacol*. **2021**, *188*, 114564.)

CONCLUSIONS

This thesis reports the initial identification of two bile acid derivatives, compounds **3** and **7**, with multitarget activity towards the nuclear receptor, ROR-γ, and the membrane receptor, GPBAR1, useful for the treatment of metabolic disorders. On the basis of the recent discovery about the dual activity of the first non-steroidal ligand, REV5901, as agonist against the membrane bile acid receptor, GPBAR1, and antagonist against $CysLT_1R$ as a promising drug for the treatment of colitis, we have undertaken a lead optimization study of the quinolinic scaffold in order to obtain derivatives with improved activity and pharmacokinetic profile. We designed two ligands, compounds 5 and 6, with a simplified chemical structure and better synthetic accessibility that maintain potency against the two GPCRs with therapeutic potential in the treatment of colitis, syndromes and other diseases related to GPBAR1/$CysLT_1R$. Furthermore, we rationalized the first structure-based rationalization of ligand binding to $CysLT_1R$, achieved through the combined application of experimental and *in silico* techniques, which will ultimately help guide future drug discovery studies on $CysLT_1R$ and GPBAR1.

The discovery of dual modulators is crucial not only to provide new opportunities for the treatment of lipid and glucose disorders, where these receptors play a pivotal role, but also to unravel the physiological actions and pathological implications of signaling pathways under their control. These extensive ligand/receptor binding studies will allow to clarify the structural requirements for the recognition of the aforementioned targets, in order to obtain new chemical entities with a double modulatory profile. This multi-target drug design represents a promising approach in identifying novel drug protocols for hepatic and metabolic disorders, in which different pathways and processes are involved. The ultimate goal of this project is to test the activity of these ligands against NR, M-BAR and other bile acid receptors to amplify the investigation of a multi-target profile.

Another project discussed involved the identification, through a virtual screening campaign, of natural and clinically available compounds capable of interacting with Spike RBD that could be useful in the prevention or treatment of SARS-CoV-2 infection. We firstly identified druggable pockets on the surface of the central core of the β sheet of RBD protein S and, subsequently, found several triterpenoids, such as glycyrrhetinic and oleanolic acids, and natural bile acids and their semisynthetic derivatives capable of binding Spike RBD and reducing interaction with ACE2 *in vitro*.

Subsequently, the discovery of UDCA derivatives, BAR107 and BAR708, reached once again through pharmacological assays and *in silico* studies, as potent activators of ACE2, is reported, supported by the identification of their mechanism of action by means of atomistic MD simulations of the PD apo ACE2 domain and *in silico* analysis of the ligand/ACE2 interaction (e.g., PCA, RMSF and conformational clusterization analysis). Furthermore, considering our recent results on the ability of UDCA and other BA derivatives to influence interactions between the SARS-CoV-2 RBD spike protein and ACE2 *in vitro*, we investigated the ability of a set of UDCA derivatives. to inhibit the SARS-CoV-2/ACE2 interaction, identifying BARn501 and BAR501–6α as promising inhibitors of this interaction and mild ACE2 activators. Instead, the best ACE2 activators in this series show very weak *in vitro* inhibition of RBD interaction with ACE2. Thus, ACE2 activation was not related to the inhibitory capacity of BA derivatives, thus stating that there is no relationship between ACE2 activation and inhibition of interaction with RBD.

Finally, two projects not related to the topic of bile acids were discussed.

In the first project, the mechanism underlying OFD1 and the signaling system involved in this pathology were identified. Underlying it is centrosomal transduction which controls the timing and spread of cAMP signaling to ciliary targets. In response

to GPCR activation, the signaling system, which includes the ubiquitin ligase E3 praja2, PKA and OFD1 and driven by TBC1D31, dynamically couples the PKA phosphorylation of OFD1 to its ubiquitylation and proteolysis, an essential process for the correct biogenesis of the ciliary body. After identifying the essential praja2 and TBC1D31 sequences through *in vitro* studies, 3D structures were generated. The molecular basis of the praja2 and TBC1D31 interaction were investigated by docking and MD studies and the binding modes were supported and validated by *in silico* mutagenesis studies. Finally, the alteration of the TBC1D31/praja2/OFD1 axis in ciliogenesis, leading to developmental defects, was investigated *in vivo*.

The second and final study looked at the ability of natural flavonoids to regulate Ace2 expression in intestinal models of inflammation. The result of *in vitro* analyses demonstrated that pelargonidin is able to activate AhR and reverse intestinal inflammation caused by chronic exposure to a high-fat diet or intestinal barrier agent TNBS in an AhR-dependent manner and associated to the upregulation of ACE2 expression, which they rate correlated to high levels of Tnf-α. Molecular docking studies have identified that pelargonidin binds to a fatty acid binding pocket on the receptor binding domain of the SARS-CoV-2 Spike protein, while *in vitro* studies have shown that the ability of this flavonoid to significantly reduce the binding of the Spike RBD protein to ACE2 and reduces SARS-CoV-2 replication in a concentration-dependent manner, ultimately yielding a potential candidate for the treatment of intestinal inflammation.

COMPUTATIONAL METHODS

4.1 Design of dual modulators against the nuclear retinoic acid-related orphan receptor γ (ROR-γ) and the membrane receptor G-protein bile acid receptor 1 (GPBAR1)

Receptors and ligands preparation

ROR-γ. The crystal structures of the human retinoic acid-related orphan receptors γ in the active and inactive conformation (PDB ID 3l0j and 5ntk, respectively)[119,230] were downloaded from the Protein Data Bank website. The nuclear receptor coactivator 2 (Src-2) in the ROR-γ active conformation, the co-crystallized ligands and water molecules were removed. Aminoacids residues protonation states were assigned in accordance with the most populated ones predicted by the H++ webserver[231] at pH 7.4.

GPBAR1. GPBAR1 homology model reported in D'Amore *et al.*[120] was employed for docking calculations. The receptor was prepared as reported in the work of Biagioli *et al.*[43]

Both the receptors were treated with the Protein Preparation Wizard[232] tool implemented in Maestro ver. 11.8.[233]

Ligands. The 3D structure of compounds **3** and **7** was built using the Graphical User Graphical User Interface (GUI) of Maestro ver. 11.8.[233] The protonation state of such compounds at pH 7.4 in water has been calculated using the Epik[234] module. Finally, **3** and **7** were then minimized with the OPLS 2005 force field[235] using the Polak-Ribiere Conjugate Gradient (PRCG) algorithm[236] and 2500 iteration steps.

Docking calculations

ROR-γ. Preliminary docking calculations were performed using Glide, Autodock 4.2 and FLAP docking programs[117, 118, 156, 237,238] to reproduce the binding pose of the ligand co-crystallized in the ROR-γ receptor (PDB ID 3l0j and 5ntk).[119, 228, 230] This redocking step allowed to assess the binding mode of steroidal scaffold-based on ROR-γ ligands and identify the most suitable parameters and scoring function for docking of compounds **3** and **7**. All docking calculations were performed using two different conformations of ROR-γ receptor: the closed form (i.e., H12 in the active conformation, PDB ID 3l0j)[119] and the opened form (i.e., H12 in the inactive conformation, PDB ID 5ntk).[230] The accuracy of each docking program was assessed by comparing the binding modes predicted by the calculation with those of the co-crystallized ligands retrieved from the PDB.

The validation step showed that the Glide with Standard Precision (SP)[117,118] algorithm was the best software to reproduce the ligand binding mode of agonists and inverse agonists to ROR-γ.

The docking procedure was carried out with the Glide software package,[117,118] using the Standard Precision (SP) algorithm of the GlideScore function and the OPLS 2005 force field.[117,118,235] A grid box of 25 × 25 × 25 Å for ROR-γ receptor and one of 25 × 25 × 25 Å for GPBAR1 centered on the ligand binding cavity were created to compute the interaction grids. Default parameters were applied. A total amount of 100 poses was generated and the conformational sampling of the ligand was enhanced by two times, with respect to the default setting of Glide. Docking conformations of compounds **3** and **7** were then clustered based on their atomic RMSD with a threshold of 2Å. Globally, seven clusters were obtained and, among them, only the

conformation included in the most populated cluster owing both the Glide Emodel and GlideScore lowest-energy value was considered.

Molecular dynamics simulations

ROR-γ. MDs were performed with NAMD ver. 2.12,[239] using the Amber *ff14SB* and the General Amber Force Field (GAFF) parameters[240,241] for the protein and the ligands, respectively. Each complex was solvated in a 10.0 Å layered cubic water box using the TIP3P water-model parameters[111] (about 90000 atoms each). 89 Na^+ and 89 Cl^- ions were added to reproduce the experimental buffer conditions of 200 mM. The obtained system was thus subjected to three minimization steps using the conjugate gradient algorithm in the following conditions: *i)* energy minimization of water molecules and ions, keeping the solute restrained (50000 steps); *ii)* energy minimization of the system, keeping the protein backbone and ligand's atoms restrained (50000 steps); iii) energy minimization of the entire system without any restriction (50000 steps).

Thus, each system was gradually heated from 50 to 300 K using a stepwise approach in which the molecules were first simulated for 250 ps in the NVT ensemble, followed by 250 ps of simulation in the NPT ensemble at 1 atm, before increasing the temperature by 50 K. This cycle was repeated until reaching 300 K and at each step the restraints were reduced by 2 Kcal/mol. Afterward, the proteins were allowed to relax without constraints for 5 ns at 300 K in the NPT ensemble before launching the production runs. Finally, a production run of 1 μs was carried out in the NPT ensemble at 1 atm and 300 K. A 2 fs integration time step was employed in each step.

The same MDs protocol has been applied to the simulation of the apo-form of ROR-γ receptor (PDB ID 3l0j).[119]

GPBAR1. Each receptor-ligand complex has been embedded in a 94 Å × 94 Å (in *x* and *y* axes) lipid bilayer composed by 1-Palmitoyl-2-oleoylphosphatidylcholine (POPC) and cholesterol with a ratio POPC:cholesterol equal to 7:3., using the membrane-builder tool of CHARMM-GUI.org,[242,243] and then surrounded by explicit waters. Each membrane-receptor complex was solvated using the TIP3P water model[111] and neutralized with the addition of 1 Cl^- ion. The ionic strength was kept at 0.15 M by NaCl. The Amber *ff14SB,* GAFF and *lipid14* Amber force fields[240,241,244] were used to parametrize the protein, the ligand, and the lipids, respectively. All simulations were performed with NAMD2.12 code.[239] The SHAKE algorithm was applied to constrain bonds involving hydrogen atoms, and thus integration of 1 fs time step interval until the thermalization at 200 K, then increased to 2 fs time step, was used.

System thermalization was carried out using the same protocol designed for ROR-γ. Finally, a production run of 1 μs was carried out in the NPT ensemble at 1 atm and 300 K.

4.2 Dual modulators against Cysteinyl Leukotriene Receptor 1 ($CysLT_1R$) and G-protein coupled bile acid receptor 1 (GPBAR1)

Receptors and ligands preparation

$CysLT_1R$. The X-ray structure of the homo sapiens Cysteinyl leukotriene receptor 1 (PDB ID 6rz4)[134] was downloaded from the Protein Data Bank website. The soluble cytochrome b562 fragment, the co-crystallized ligand (pranlukast) and water molecules were removed and the residue Gln274 was reconstructed. The missing 2 residues of ECL3 and the missing transmembrane helix 8 (TM8) were modeled using the Modeller 9.2 software package.[245,246] For TM8, the crystallographic structure of $CysLT_2R$ (PDB ID 6rz6)[247] was employed as template and its secondary structure was confirmed using the prediction tools PSIpred and Spider3.[248,249] Residues protonation states were assigned using the prediction H++ webserver[231] at pH 7.4. The final model was validated via 1 μs long molecular dynamics simulation. The protein was put in a box of size 10x10x12 nm and embedded in a lipid bilayer composed of and 1-parlmitoyl-2-oleoylphosphocholine (POPC) using the CHARMM-GUI webserver.[242,243] For solvation, TIP3P water molecules[111] were employed and a 0.150 mM concentration of NaCl was added to reach neutrality. The simulation was performed using the Amber *ff14SB* and *lipid17* force fields (for protein and lipid, respectively) with the GROMACS 2020.4 software package.[240,241,250]

GPBAR1. GPBAR1 homology model reported in D'Amore *et al.*[120] was employed for docking calculations. The receptor was prepared as in Biagioli *et al.*[43]

Both the receptors were treated with the Protein Preparation Wizard[232] tool implemented in Maestro ver. 11.8.[233]

Ligands. 3D structures of compounds **1-15** were built using the Graphical User Interface (GUI) of Maestro ver. 11.8.[233] The protonation state of **1-15** at pH 7.4 has been calculated using the Epik module.[234] Finally, **1-15** were then minimized using the OPLS 2005[235] force field through 2500 iteration steps of the PRCG algorithm.[236]

Docking calculations

Preliminary docking calculations were performed using Glide[251] and Autodock 4.2[156] to reproduce the binding pose of the co-crystalized ligand, pranlukast, with $CysLT_1R$ (PDB ID 6rz4).[134] This redocking step allowed to identify the most suitable parameters and scoring function for docking of compounds **1-15**. Glide SP software was employed for the docking calculations, thanks to the capability to reproduce the pranlukast crystallographic binding pose. The results were clustered and successively ranked according to the Glide Emodel and the Glide Score.

Docking calculations of **1-15** on GPBAR1 were performed using the same approach described in *Biagioli et al.*[43]

Additionally, in order to consider the ligand induced fit effect on the receptors' binding sites, we performed docking calculations on **5** in both centroid of the most populated GPBAR1 and $CysLT_1R$ protein conformation retrieved from the MD simulations in complex with compound **5**.

In detail, the docking procedure was carried out with the Glide software package,[251] using the Standard Precision (SP) algorithm of the GlideScore function[118] and the OPLS 2005 force field.[235] A grid box of 2.5 × 1.6 × 1.7 nm for GPBAR1 receptor and one of 1.6 × 2.0 × 1.8 nm for $CysLT_1R$ centered on the ligand binding cavity were created. A total amount of 100 poses was generated and the conformational sampling of the ligand was enhanced by two times, as reported by the default setting of Glide. Docking conformations of **1-15** were then clustered based on their atomic RMSD.

Globally, seven clusters were obtained and, among them, only the conformation included in the most populated cluster with both the Glide Emodel and GlideScore lowest-energy value was considered.

Molecular dynamics simulations

MDs were performed with GROMACS suite ver. 2020.4,[252] using the Amber *ff14SB*, *lipid17* and GAFF parameters[240,241] for the proteins, lipids and ligands, respectively. Protein/ligand complexes were prepared as previously reported for $CysLT_1R$ and GPBAR1 and embedded in a phosphatidylcholine (POPC) lipid bilayer of sizes 10x10nm. The resulting membrane was then solvated with TIP3P water[111] and a 0.150 mM concentration of NaCl into a 10x10x12 nm box. The whole procedure was carried on using the CHARMM-GUI webserver.[242,243] The systems were minimized using the steepest descent algorithm in a two steps procedure. First, the protein and ligand heavy atoms were restrained, whereas water molecules and ions were left free and only the movement on the Z axis of hydroxyl group of CHL and the phosphate group of POPC was restrained. Afterward, the restraints were removed, and a second round of minimization was performed. The systems were then gradually heated from 50 to 300 K using a stepwise approach of NVT/NPT simulations at fixed temperature, before increasing it by 50 K. Each NVT/NPT step lasted 1 ns. An initial restraint of 1000 kJ/mol at 50 K was applied on proteins, ligands and lipids as described for the minimization procedure. After each NVT/NPT cycle, the restraints were lowered by 160 kJ/mol. The Langevin dynamics integrator and the Berendsen barostat with semi-isotropic coupling at 1 atm were employed. After reaching 300 K, a preliminary production run of 10ns without restraints was performed using the Langevin dynamics integrator and the Parrinello-Rahman barostat with semi-isotropic coupling at 1 atm. The same parameters were employed for the following production runs of 1 μs. In all

these simulations, a time step integration of 2 fs. For the calculation of electrostatic and Van der Waals interactions, the Particle-Mesh Ewald (PME) and the cutoff algorithms were used, respectively, with a threshold of 1.2 nm. The cluster analysis trajectory was carried out using the GROMACS gmx cluster tools with the GROMOS method[253] and a 0.2 nm cutoff.

Free-energy calculations

Well-tempered MetaD simulations were performed using the same protocol described for MD calculations. However, the GROMACS suite ver. 2020.4[254] was patched with the Plumed software package ver. 2.6.2[255] and the C_α atoms of the protein structured parts (i.e., alpha helices, beta strands) were restrained around the initial conformation using a RMSD-based harmonic potential with constant 10000 kJ/mol and threshold 0.1 nm. The distance between the heavy atoms of the quinoline moiety and the C_β of $CysLT_1R$ $Arg79^{2.60}$ was chosen was chosen as the collective variable, allowing to explore the values from 0 to 3.0 nm and limit the sampling of the free energy landscape within the binding pocket. To do so, an upper wall with constant 10000 kJ/mol was placed at 3.0 nm of the distance CV to prevent the ligand from exiting the $CysLT_1R$ cavity. A bias of 1 kJ/mol was deposited every 5 ps with a sigma of 0.05 nm and a bias factor of 15. The MetaD simulations were performed using 10 multiple walkers lasting 150 ns each, for a total of 1.5 μs of calculation.

4.3 Natural and semi-synthetic steroidal agent acting on Spike protein Receptor Binding Domain (RBD) of SARS-CoV-2

Receptor and ligands preparation

Ligands preparation. The library of FDA approved drugs has been obtained both from DrugBank[256] (2106 compounds) and from the Selleckchem website (FDA-approved Drug Library, 2020)[154] (tot. 2638). Each database was converted to 3D and prepared with the LigPrep tool (Schrödinger, 2019)[233] considering a protonation state at a physiological pH of 7.4. Subsequently, the two libraries were merged and deduplicated with Open Babel,[257] giving a total amount of 2,906 drugs. The bile acids (BA) focused library was prepared with the same protocol described above.

Receptor preparation. The electron microscopy (EM) model of SARS-CoV-2 Spike glycoprotein was downloaded from the Protein Data Bank (PDB ID 6vsb).[150] Missing loops were added from the Swiss-Model web-site.[258] The obtained model was submitted to the Protein Preparation Wizard tool implemented into Maestro ver. 11.8[233] to assign bond orders, adding all hydrogen atoms and adjusting disulfide bonds.

Virtual Screening

The pocket search was performed by using the Fpocket website.[259]

The AutoDock4.2.6 suite[156] and the Raccoon2 graphical interface[157] were employed to carry out the virtual screening approach using the Lamarckian genetic algorithm (LGA). This hybrid algorithm combines two conformational research methods, the genetic algorithm and the local research. For the first low-accuracy screening, for each of the 2906 drugs, 3 poses were generated using 250,000 steps of genetic algorithm and 300 steps of local search, while in the second high-accuracy screening protocol,

20 poses for each ligand were generated, increasing the number of genetic algorithm steps to 25,000,000. The MGLTools were used to convert both ligands and each pocket into appropriate pdbqt files. Virtual screening was performed on a hybrid CPU/GPU HPC cluster equipped with 2 NVIDIA® Tesla® V100 GPUs and 560 Intel® Xeon® Gold and 64 AMD® EPYC® processors.

Each of the six selected RBD pockets was submitted to the AutoGrid4 tool, which calculates, for each bonding pocket, maps (or grids) of interaction, considering the different ligands and receptor-atom types through the definition of a cubic box. Subsequently, for each grid AutoDock4 calculates interaction energies (ADscore) that express the affinity of a given ligand for the receptor.

Molecular dynamics simulations

MD simulations were performed using the CUDA version of the AMBER18 suite[260] on NVIDIA Titan Xp and K20 GPUs, using the Amber *ff14SB* force field[240] to treat the protein. The RBD was then immersed in a pre-equilibrated octahedral box of TIP3P water[111] and the system was neutralized by adding Na^+ and Cl^- ions. The system was then minimized using energy gradient convergence criterion set to 0.01 kcal/mol $Å^2$ in four steps involving: (i) an initial 5,000 minimization steps (2,500 with the steepest descent and 2,500 with the conjugate gradient) of only hydrogen atoms, (ii) 20,000 minimization steps (10,000 with the steepest descent and 10,000 with the conjugate gradient) of water and hydrogen atoms, keeping the solute restrained, (iii) 50,000 minimization steps (25,000 with the steepest descent and 25,000 with the conjugate gradient) of protein side chains, water and hydrogen atoms, (iv) 100,000 (50,000 with the steepest descent and 50,000 with the conjugate gradient) of complete minimization. Successively, the water, ions and protein side chains were thermally equilibrated in three steps: (i) 5 ns of NVT equilibration with the Langevin thermostat by gradually

heating from 0K to 300K, while gradually rescaling solute restraints from a force constant of 10 to 1 kcal/mol Å^2, (ii) 5 ns of NPT equilibration at 1 atm with the Berendsen thermostat, gradually rescaling restraints from 1.0 to 0.1 kcal/mol Å^2, (ii) 5 ns of NPT equilibration with no restraints. Finally, a production run of 500 ns was performed using a timestep of 2 fs. The SHAKE algorithm was used for those bonds containing hydrogen atoms in conjunction with periodic boundary conditions at constant pressure and temperature, particle mesh Ewald for the treatment of long range electrostatic interactions, and a cutoff of 10 Å for nonbonded interactions.

Dynamical Network Analysis

The Dynamical Network Analysis was performed on 500 ns long MD trajectories of the RBD domain using the plugin Carma ver. 0.8[261] implemented in VMD 1.9.2.[262] The optimal community distribution is calculated by using the Girvan–Newman algorithm.[263] Edges between each node (here defined as C_α atoms) were drawn between those nodes whose residues were within a default cut-off distance (4.5 Å) for at least 75% of our MD trajectories. Communities map analysis and representation were obtained using the NetworkView tool, implemented in VMD 1.9.2.[262]

4.4 Bile acids derivatives as potent angiotensin converting enzyme 2 ACE2 activators

Virtual Screening

The crystal structure of the open apo form of homo sapiens ACE2 (PDB ID 1r42)[166] was downloaded from the Protein Data Bank website. The disordered segment of the collectrin homology domain and water molecules were removed. The receptor was treated with the Protein Preparation Wizard tool[232] implemented in Maestro ver. 11.8[233] to assign bond orders, to add hydrogen atoms, adjust disulfide bonds, and assign residues protonation state at pH 7.4.

Virtual screening (VS) was performed on an *in house* library of 67 bile acids (BAs), 10 natural and 57 semisynthetic derivatives, enriched with previously identified ACE2 activators, hydroxyzine, minithixen, and DIZE.[92] Chemical/physical properties of all of the 67 compounds were calculated with QikProp tool ver. 5.8.[264] Since even minor structural changes of steroids can produce potential biological activities, we build our *in house* BAs library to include compounds sharing a 17-carbon-atom skeleton composed of four fused rings, which form the typical steroidal scaffold. They vary from one another in the position and name of the substituent groups. The steroidal carbons hydrogens that have been replaced in our *in house* library are: (i) those in positions 3 and 7, which have been replaced with a hydroxyl group in both different configurations (α and β); (ii) the hydrogen at C7, which has been replaced with an ethyl group in both configurations (α and β); (iii) finally, the C24 has been substituted with different polar and apolar groups. Docking calculations were performed in a box including the hinge-bending region of ACE2 on the Protein Data Bank deposited structure of ACE2 in the open conformation (PDB ID 1r42),[166] according to previous

reports on the discovery of ACE2 activators.[90] The VS procedure was carried out with the AutoDock4.2.6 suite[156] and the Raccoon2 graphical interface[157] using LGA algorithm. The VS protocol adopted was the same described in our previous work.[153] To further assess our docking protocol, re-docking calculations were performed on the potent ACE2 inhibitor MLN-4760 in the ACE2 binding site (PDB ID 1r4l).[166] Given the presence of a Zn2+ ion coordinating the ligand, the improved AutoDock4(Zn) force field was used for the calculation.[265]

The receptor was submitted to the AutoGrid4 tool, which calculated interaction grids, considering the two ligands and receptor-atom types through the definition of a cubic box of 46 × 46 × 46 Å. Subsequently, for each grid, AutoDock4 calculated interaction energies (ADscore) that express the affinity of a given ligand for the receptor.

Molecular dynamics simulations

MD simulations of apo ACE2 and ACE2 in complex with BAR708 and BAR107 were performed with the CUDA version of the AMBER18 suite[266,260] using the Amber *ff14SB*[240] to treat the protein, while ligands charges were computed using the restrained electrostatic potential (RESP) fitting procedure.[267] First, the ligand ESP was calculated through the Gaussian16 package[268] using the 6-31G* basis set at Hartree–Fock level of theory. Then, RESP charges and the ligand force field parameters were obtained from the two-stage fitting procedure using Antechamber[269] and GAFF2 parameters.[270] The system was then immersed in a preequilibrated octahedral box of TIP3P water molecules[111] and the system was neutralized. The system was then minimized and successively equilibrated in a multistep procedure as previously described.[153] Specifically, each system was minimized in four steps using the energy gradient convergence criterion set to 0.01 kcal/mol Å^2 involving: (i) 5000 minimization steps

(2500 with the steepest descent and 2500 with the conjugate gradient) of only hydrogen atoms; (ii) 20 000 minimization steps (10 000 with the steepest descent and 10 000 with the conjugate gradient) of water and hydrogen atoms, keeping the solute restrained; (iii) 50 000 minimization steps (25 000 with the steepest descent and 25 000 with the conjugate gradient) of only the side chains of the protein, water, and hydrogen atoms; (iv) 100 000 (50 000 with the steepest descent and 50 000 with the conjugate gradient) of complete minimization. Successively, water molecules, ions, and protein side chains were thermally equilibrated in three steps: (i) 5 ns of NVT equilibration with the Langevin thermostat by gradually heating from 0 to 300 K, while gradually rescaling solute restraints from a force constant of 10 to 1 kcal/mol $Å^2$; (ii) 5 ns of NPT equilibration at 1 atm with the Berendsen thermostat by gradually rescaling restraints from 1.0 to 0.1 kcal/mol $Å^2$; and (iii) 5 ns of NPT equilibration with no restraints. Finally, three independent MD production runs of 500 ns each were performed for each system using a timestep of 2 fs. The SHAKE algorithm was used for those bonds containing hydrogen atoms in conjunction with periodic boundary conditions at constant pressure and temperature, particle mesh Ewald (PME)[271] for the treatment of long-range electrostatic interactions and a cutoff of 10 Å for nonbonded interactions.

Principal Component Analysis (PCA)

The principal component analysis (PCA)[272] of apo ACE2 and ACE2 complexed with BAR708 and BAR107 was carried out using the CPPTRAJ module[273] of the AMBER18 Suite.[260] First, the overall 1.5 μs of MD trajectories of each system was stripped of solvent and ions. Then, to take into account the internal dynamics of ACE2, global rotational/translational motions of the protein were removed by fitting the stripped trajectories to the protein heavy atoms of the first MD frame. This allowed us to generate the average structure of the protein of each system, which was used as the

reference structure for the PCA analysis. Finally, we have generated the coordinate covariance matrix and diagonalized it, thus obtaining the first four principal components (PCs) as eigenvectors and eigenvalues. The pseudotrajectory of the protein motion was then imported and visualized into the Normal Mode Wizard GUI (NMWiz)[274] of VMD, to generate the porcupine plot of each motion, with the arrows representing the magnitude and direction of the eigenvectors.

4.5 The TBC1D31/praja2 complex controls primary ciliogenesis through PKA-directed OFD1 ubiquitylation

TBC1D31 and praja2 homology model

The wild-type FASTA sequence of human TBC1D31 was obtained from Uniprot website (Q96DN5) and submitted to I-TASSER server[275] in order to build the 3D homology model. Similarly, the segment from W550 to D570 (praja2$_{W550\text{-}D570}$) of E3 ubiquitin-protein ligase praja2 (PJA2) human sequence was searched on Uniprot (O43164) and submitted to I-TASSER server. I-TASSER gave five homology models of both TBC1 and praja2$_{W550\text{-}D570}$. The quality of the predicted models was ranked based on their C-score, which is calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulation. A higher value of the C-score signifies a model with high confidence. In our case, among the predicted models we selected the first model for both systems, having a C-score of −0.99 and −0.57 for TBC1D31 and praja2$_{W550\text{-}D570}$, respectively. Specifically, as concerns the TBC1D31 model, according to experimental data, the C-terminal alpha-helix region from Q941 to A970 was used for docking and MD calculations.

Docking calculations

TBC1D31/praja2 protein–protein docking was performed in two steps. The first step was aimed to obtain a starting TBC1D31/praja2 complex conformation, and it was performed using AutoDock Vina ver. 1.0.2 software.[205] Specifically, the grid box (size: 102 x 100 x 100) was built in order to include the whole alpha helix of TBC1D31 C-terminal segment Gln941-Ala970, an exhaustiveness of 32 was used and 20 docking

poses of Trp550-Asp570 praja2 were generated. In the second step, an AutoDock Vina docking pose refinement was carried out with FlexPepDock, a high-resolution peptide-protein docking protocol for the modelling of peptide-protein complexes implemented in the Rosetta framework.[206] In particular, the best pose found by AutoDock Vina was submitted to FlexPepDock webserver, which confirmed the goodness of the docking sampling.

Molecular dynamics simulations

The TBC1D31/praja2 FlexPepDock refined pose was submitted to 2μs-long MD simulation using the AMBER16 suite. The complex was firstly parameterized with the *LEaP* module of AmberTools16 suite, using the *ff14SB* force field.[240] The system was immersed in a pre-equilibrated octahedral box of TIP3P water molecules[111] and neutralized by adding Na^+ and Cl^- counterions. The final system, of about 22,000 atoms, was minimized in three steps using an energy gradient convergence criterion set to 0.01 kcal/mol $Å^2$ involving: (i) only the hydrogen atoms of the system (2,000 steps of steepest descent); (ii) hydrogen atoms, water molecules and counterions (4,000 steps of steepest descent); (iii) minimization of the whole system (10,000 steps of steepest descent). Subsequently, water, ions and protein side chains were thermalized in two equilibration steps: (i) 200 ps heating water and ions from 0 to 298 K with constant volume, restraining protein atoms; (ii) 800 ps of thermalization step with pressure control at 1 atm (NPT ensemble) of the whole system, without any restraint; (iii) additional 400 ps were performed in order to further equilibrate the system density in NPT ensemble. Finally, the 2 μs-long of production run was performed in NPT using a time step of 2 fs. MD's trajectory was analysed with VMD 1.9.4.[262] A cluster analysis of MD trajectory was conducted considering a praja2 peptide RMSD cut-off of 5.0 Å.

Protein Network Analysis

The protein–protein interaction sub-network has been obtained by using GeneMANIA.[276,277] In particular, the network has been built using ODF1, PJA2 and TBC1D31 as input seeds and the GeneMANIA "Physical Interactions" catalogue as protein–protein interaction database.[278] This catalogue is composed by all interactions reported in different experiments and collected in different databases such as BioGRID and PathwayCommons. Starting from the input proteins and the chosen set of protein–protein interaction networks, the GeneMANIA algorithm extracts a single association network, centred on the input proteins and summarizing the information from all the different networks. After the physical interaction network has been obtained, the subnetwork induced by all the first-order interactors of ODF1, PJA2 and TBC1D31 was computed by filtering out from the network all the nodes not having any direct (physical) interaction with the input proteins. Finally, the set of genes associated with the proteins in the final network was functionally annotated on Gene Ontology terms using the g:GOst module of the gProfiler toolset, with G:SCS as multiple test correction method and 0.05 as *P*-value threshold.

4.6 Discovery of pelargonidin as a potential inhibitor of the SARS-CoV-2 interaction and angiotensin converting enzyme 2 (ACE2)

Receptor and ligand preparations

In order to study the binding mode of pelargonidin to SARS-CoV-2 RBD, docking calculations were performed in the central β-sheet core and in the flavonoids binding pocket using the cryo-electron microscopy structure of SARS-CoV-2 Spike protein in the "up" conformation (PDB ID 6vsb),[150] while for dockings in the fatty acids (FA) pocket the cryo-electron microscopy structure of SARS-CoV-2 Spike protein in the "down" conformation was employed (PDB ID 6zb5).[56]

Receptor. The receptor was prepared using the Protein Preparation tool[232] implemented in Maestro ver. 11.8[233] in order to assign bond orders, adding all hydrogen atoms and adjusting disulfide bonds.

Ligand. The 3D structure of the pelargonidin was built using the build panel implemented in Maestro ver. 11.8.[233] The protonation state of pelargonidin at pH 7.4 in water was calculated using the Epik module.[234] Finally, the compound was minimized using the OPLS 2005 force field[235] through 2500 iteration steps of the PRCG algorithm.[236]

Docking calculation

Docking calculation was performed with the Glide software package[251] using the Standard Precision (SP) algorithm of the GlideScore[117] function and the OPLS 2005 force field.[235] A 20 × 20 × 20 Å grid was created for the SARS-CoV-2 receptor centered on the presumed binding pocket. A total of 100 poses were generated and

clustered based on their atomic RMSD, obtaining five clusters. The conformation included in the most populated cluster with both the lowest Glide Emodel and GlideScore energy values was considered.

MD trajectories were visualized using VMD[262] software and all figures were rendered by UCSF Chimera.[279]

REFERENCES

(1) Fiorucci, S.; Distrutti, E. The Pharmacology of Bile Acids and Their Receptors. *Handb Exp Pharmacol* **2019**, *256*, 3-18. DOI: 10.1007/164_2019_238.

(2) Maruyama, T.; Miyamoto, Y.; Nakamura, T.; Tamai, Y.; Okada, H.; Sugiyama, E.; Itadani, H.; Tanaka, K. Identification of membrane-type receptor for bile acids (M-BAR). *Biochem Biophys Res Commun* **2002**, *298* (5), 714-719. DOI: 10.1016/s0006-291x(02)02550-0.

(3) Carino, A.; Biagioli, M.; Marchianò, S.; Fiorucci, C.; Zampella, A.; Monti, M. C.; Scarpelli, P.; Ricci, P.; Distrutti, E.; Fiorucci, S. Ursodeoxycholic acid is a GPBAR1 agonist and resets liver/intestinal FXR signaling in a model of diet-induced dysbiosis and NASH. *Biochim Biophys Acta Mol Cell Biol Lipids* **2019**, *1864* (10), 1422-1437. DOI: 10.1016/j.bbalip.2019.07.006.

(4) Bloch, K.; Berg, B. N.; Rittenberg, D. The biological conversion of cholesterol to cholic acid. *Journal of Biological Chemistry* **1943**, *149* (2), 511-517. DOI: 10.1016/S0021-9258(18)72197-2.

(5) Kullak-Ublick, G. A.; Stieger, B.; Meier, P. J. Enterohepatic bile salt transporters in normal physiology and liver disease. *Gastroenterology* **2004**, *126* (1), 322-342. DOI: 10.1053/j.gastro.2003.06.005.

(6) Stellaard, F.; Sackmann, M.; Sauerbruch, T.; Paumgartner, G. Simultaneous determination of cholic acid and chenodeoxycholic acid pool sizes and fractional

turnover rates in human serum using 13C-labeled bile acids. *J Lipid Res* **1984**, *25* (12), 1313-1319. DOI: 10.1016/S0022-2275(20)34446-1.

(7) Vlahcevic, Z. R.; Heuman, D. M.; Hylemon, P. B. Regulation of bile acid synthesis. *Hepatology* **1991**, *13* (3), 590-600. DOI: 10.1002/hep.1840130331.

(8) Hofmann, A.F.; Borgstroem, B. The intraluminal phase of fat digestion in man: the lipid content of the micellar and oil phase of intestinal content obtained during fat digestion and absorption. *J Clin Invest* **1964**, *43*, 247-257. DOI: 10.1172/JCI104909.

(9) Thomas, C.; Pellicciari, R.; Pruzanski, M.; Auwerx, J.; Schoonjans, K. Targeting bile-acid signalling for metabolic diseases. *Nat Rev Drug Discov* **2008**, *7* (8), 678-693. DOI: 10.1038/nrd2619.

(10) Solt, L. A.; Burris, T. P. Action of RORs and their ligands in (patho)physiology. *Trends Endocrinol Metab* **2012**, *23* (12), 619-627. DOI: 10.1016/j.tem.2012.05.012.

(11) Santori, F. R. Nuclear hormone receptors put immunity on sterols. *Eur J Immunol* **2015**, *45* (10), 2730-2741. DOI: 10.1002/eji.201545712.

(12) Zhang, Y.; Luo, X. Y.; Wu, D. H.; Xu, Y. ROR nuclear receptors: structures, related diseases, and drug discovery. *Acta Pharmacol Sin* **2015**, *36* (1), 71-87. DOI: 10.1038/aps.2014.120.

(13) Wang, H.; Chen, J.; Hollister, K.; Sowers, L. C.; Forman, B. M. Endogenous bile acids are ligands for the nuclear receptor FXR/BAR. *Mol Cell* **1999**, *3* (5), 543-553. DOI: 10.1016/s1097-2765(00)80348-2.

(14) Fiorucci, S.; Mencarelli, A.; Palladino, G.; Cipriani, S. Bile-acid-activated receptors: targeting TGR5 and farnesoid-X-receptor in lipid and glucose disorders. *Trends Pharmacol Sci* **2009**, *30* (11), 570-580. DOI: 10.1016/j.tips.2009.08.001.

(15) Chiang, J. Y. L.; Ferrell, J. M. Bile acid receptors FXR and TGR5 signaling in fatty liver diseases and therapy. *Am J Physiol Gastrointest Liver Physiol* **2020**, *318* (3), G554-G573. DOI: 10.1152/ajpgi.00223.2019.

(16) Evans, R. M. The steroid and thyroid hormone receptor superfamily. *Science* **1988**, *240* (4854), 889-895. DOI: 10.1126/science.3283939.

(17) Olefsky, J. M. Nuclear receptor minireview series. *J Biol Chem* **2001**, *276* (40), 36863-36864. DOI: 10.1074/jbc.R100047200.

(18) Amoutzias, G. D.; Pichler, E. E.; Mian, N.; De Graaf, D.; Imsiridou, A.; Robinson-Rechavi, M.; Bornberg-Bauer, E.; Robertson, D. L.; Oliver, S. G. A protein interaction atlas for the nuclear receptors: properties and quality of a hub-based dimerisation network. *BMC Syst Biol* **2007**, *1*, 34. DOI: 10.1186/1752-0509-1-34.

(19) Linja, M. J.; Porkka, K. P.; Kang, Z.; Savinainen, K. J.; Jänne, O. A.; Tammela, T. L.; Vessella, R. L.; Palvimo, J. J.; Visakorpi, T. Expression of androgen receptor coregulators in prostate cancer. *Clin Cancer Res* **2004**, *10* (3), 1032-1040. DOI: 10.1158/1078-0432.ccr-0990-3.

(20) Forman, B. M.; Goode, E.; Chen, J.; Oro, A. E.; Bradley, D. J.; Perlmann, T.; Noonan, D. J.; Burka, L. T.; McMorris, T.; Lamph, W. W.; et al. Identification of a

nuclear receptor that is activated by farnesol metabolites. *Cell* **1995**, *81* (5), 687-693. DOI: 10.1016/0092-8674(95)90530-8.

(21) Makishima, M.; Okamoto, A. Y.; Repa, J. J.; Tu, H.; Learned, R. M.; Luk, A.; Hull, M. V.; Lustig, K. D.; Mangelsdorf, D. J.; Shan, B. Identification of a nuclear receptor for bile acids. *Science* **1999**, *284* (5418), 1362-1365. DOI: 10.1126/science.284.5418.1362.

(22) Parks, D. J.; Blanchard, S. G.; Bledsoe, R. K.; Chandra, G.; Consler, T. G.; Kliewer, S. A.; Stimmel, J. B.; Willson, T. M.; Zavacki, A. M.; Moore, D. D.; et al. Bile acids: natural ligands for an orphan nuclear receptor. *Science* **1999**, *284* (5418), 1365-1368. DOI: 10.1126/science.284.5418.1365.

(23) Laffitte, B. A.; Kast, H. R.; Nguyen, C. M.; Zavacki, A. M.; Moore, D. D.; Edwards, P. A. Identification of the DNA binding specificity and potential target genes for the farnesoid X-activated receptor. *J Biol Chem* **2000**, *275* (14), 10638-10647. DOI: 10.1074/jbc.275.14.10638.

(24) Claudel, T.; Inoue, Y.; Barbier, O.; Duran-Sandoval, D.; Kosykh, V.; Fruchart, J.; Fruchart, J. C.; Gonzalez, F. J.; Staels, B. Farnesoid X receptor agonists suppress hepatic apolipoprotein CIII expression. *Gastroenterology* **2003**, *125* (2), 544-555. DOI: 10.1016/s0016-5085(03)00896-5.

(25) Lu, T. T.; Makishima, M.; Repa, J. J.; Schoonjans, K.; Kerr, T. A.; Auwerx, J.; Mangelsdorf, D. J. Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. *Mol Cell* **2000**, *6* (3), 507-515. DOI: 10.1016/s1097-2765(00)00050-2.

(26) Goodwin, B.; Jones, S. A.; Price, R. R.; Watson, M. A.; McKee, D. D.; Moore, L. B.; Galardi, C.; Wilson, J. G.; Lewis, M. C.; Roth, M. E.; et al. A regulatory cascade of the nuclear receptors FXR, SHP-1, and LRH-1 represses bile acid biosynthesis. *Mol Cell* **2000**, *6* (3), 517-526. DOI: 10.1016/s1097-2765(00)00051-4.

(27) Jiao, Y.; Lu, Y.; Li, X. Y. Farnesoid X receptor: a master regulator of hepatic triglyceride and glucose homeostasis. *Acta Pharmacol Sin* **2015**, *36* (1), 44-50. DOI: 10.1038/aps.2014.116.

(28) Jetten, A. M.; Kurebayashi, S.; Ueda, E. The ROR nuclear orphan receptor subfamily: critical regulators of multiple biological processes. *Prog Nucleic Acid Res Mol Biol* **2001**, *69*, 205-247. DOI: 10.1016/s0079-6603(01)69048-2.

(29) Sun, Z.; Unutmaz, D.; Zou, Y. R.; Sunshine, M. J.; Pierani, A.; Brenner-Morton, S.; Mebius, R. E.; Littman, D. R. Requirement for RORgamma in thymocyte survival and lymphoid organ development. *Science* **2000**, *288* (5475), 2369-2373. DOI: 10.1126/science.288.5475.2369.

(30) Jetten, A. M. Retinoid-related orphan receptors (RORs): critical roles in development, immunity, circadian rhythm, and cellular metabolism. *Nucl Recept Signal* **2009**, *7*, e003. DOI: 10.1621/nrs.07003.

(31) Jetten, A. M.; Takeda, Y.; Slominski, A.; Kang, H. S. Retinoic acid-related Orphan Receptor γ (RORγ): connecting sterol metabolism to regulation of the immune system and autoimmune disease. *Curr Opin Toxicol* **2018**, *8*, 66-80. DOI: 10.1016/j.cotox.2018.01.005.

(32) Bronner, S. M.; Zbieg, J. R.; Crawford, J. J. RORγ antagonists and inverse agonists: a patent review. *Expert Opin Ther Pat* **2017**, *27* (1), 101-112. DOI: 10.1080/13543776.2017.1236918.

(33) Burris, T. P.; Busby, S. A.; Griffin, P. R. Targeting orphan nuclear receptors for treatment of metabolic diseases and autoimmunity. *Chem Biol* **2012**, *19* (1), 51-59. DOI: 10.1016/j.chembiol.2011.12.011.

(34) Duboc, H.; Taché, Y.; Hofmann, A. F. The bile acid TGR5 membrane receptor: from basic research to clinical application. *Dig Liver Dis* **2014**, *46* (4), 302-312. DOI: 10.1016/j.dld.2013.10.021.

(35) Fiorucci, S.; Distrutti, E. Bile Acid-Activated Receptors, Intestinal Microbiota, and the Treatment of Metabolic Disorders. *Trends Mol Med* **2015**, *21* (11), 702-714. DOI: 10.1016/j.molmed.2015.09.001.

(36) Marino, S.; Finamore, C.; Biagioli, M.; Carino, A.; Marchianò, S.; Roselli, R.; Giorgio, C. D.; Bordoni, M.; Di Leva, F. S.; Novellino, E.; et al. GPBAR1 Activation by C6-Substituted Hyodeoxycholane Analogues Protect against Colitis. *ACS Med Chem Lett* **2020**, *11* (5), 818-824. DOI: 10.1021/acsmedchemlett.9b00636.

(37) Carino, A.; Cipriani, S.; Marchianò, S.; Biagioli, M.; Scarpelli, P.; Zampella, A.; Monti, M. C.; Fiorucci, S. Gpbar1 agonism promotes a Pgc-1α-dependent browning of white adipose tissue and energy expenditure and reverses diet-induced steatohepatitis in mice. *Sci Rep* **2017**, *7* (1), 13689. DOI: 10.1038/s41598-017-13102-y.

(38) Biagioli, M.; Carino, A.; Cipriani, S.; Francisci, D.; Marchianò, S.; Scarpelli, P.; Sorcini, D.; Zampella, A.; Fiorucci, S. The Bile Acid Receptor GPBAR1 Regulates the M1/M2 Phenotype of Intestinal Macrophages and Activation of GPBAR1 Rescues Mice from Murine Colitis. *J Immunol* **2017**, *199* (2), 718-733. DOI: 10.4049/jimmunol.1700183.

(39) Kida, T.; Omori, K.; Hori, M.; Ozaki, H.; Murata, T. Stimulation of G protein-coupled bile acid receptor enhances vascular endothelial barrier function via activation of protein kinase A and Rac1. *J Pharmacol Exp Ther* **2014**, *348* (1), 125-130. DOI: 10.1124/jpet.113.209288.

(40) Carino, A.; Marchianò, S.; Biagioli, M.; Bucci, M.; Vellecco, V.; Brancaleone, V.; Fiorucci, C.; Zampella, A.; Monti, M. C.; Distrutti, E.; et al. Agonism for the bile acid receptor GPBAR1 reverses liver and vascular damage in a mouse model of steatohepatitis. *FASEB J* **2019**, *33* (2), 2809-2822. DOI: 10.1096/fj.201801373RR.

(41) Fiorucci, S.; Cipriani, S.; Baldelli, F.; Mencarelli, A. Bile acid-activated receptors in the treatment of dyslipidemia and related disorders. *Prog Lipid Res* **2010**, *49* (2), 171-185. DOI: 10.1016/j.plipres.2009.11.001.

(42) Tiwari, A.; Maiti, P. TGR5: an emerging bile acid G-protein-coupled receptor target for the potential treatment of metabolic disorders. *Drug Discov Today* **2009**, *14* (9-10), 523-530. DOI: 10.1016/j.drudis.2009.02.005.

(43) Biagioli, M.; Carino, A.; Marchianò, S.; Roselli, R.; Di Giorgio, C.; Bordoni, M.; Fiorucci, C.; Sepe, V.; Conflitti, P.; Limongelli, V.; et al. Identification of cysteinyl-

leukotriene-receptor 1 antagonists as ligands for the bile acid receptor GPBAR1. *Biochem Pharmacol* **2020**, *177*, 113987. DOI: 10.1016/j.bcp.2020.113987.

(44) Nothacker, H. P.; Wang, Z.; Zhu, Y.; Reinscheid, R. K.; Lin, S. H.; Civelli, O. Molecular cloning and characterization of a second human cysteinyl leukotriene receptor: discovery of a subtype selective agonist. *Mol Pharmacol* **2000**, *58* (6), 1601-1608. DOI: 10.1124/mol.58.6.1601.

(45) Fiorillo, B.; Sepe, V.; Conflitti, P.; Roselli, R.; Biagioli, M.; Marchianò, S.; De Luca, P.; Baronissi, G.; Rapacciuolo, P.; Cassiano, C.; et al. Structural Basis for Developing Multitarget Compounds Acting on Cysteinyl Leukotriene Receptor 1 and G-Protein-Coupled Bile Acid Receptor 1. *J Med Chem* **2021**, *64* (22), 16512-16529. DOI: 10.1021/acs.jmedchem.1c01078.

(46) Yokomizo, T.; Nakamura, M.; Shimizu, T. Leukotriene receptors as potential therapeutic targets. *J Clin Invest* **2018**, *128* (7), 2691-2701. DOI: 10.1172/JCI97946.

(47) Colazzo, F.; Gelosa, P.; Tremoli, E.; Sironi, L.; Castiglioni, L. Role of the Cysteinyl Leukotrienes in the Pathogenesis and Progression of Cardiovascular Diseases. *Mediators Inflamm* **2017**, *2017*, 2432958. DOI: 10.1155/2017/2432958.

(48) Singh, R. K.; Gupta, S.; Dastidar, S.; Ray, A. Cysteinyl leukotrienes and their receptors: molecular and functional characteristics. *Pharmacology* **2010**, *85* (6), 336-349. DOI: 10.1159/000312669.

(49) Claesson, H. E.; Dahlén, S. E. Asthma and leukotrienes: antileukotrienes as novel anti-asthmatic drugs. *J Intern Med* **1999**, *245* (3), 205-227. DOI: 10.1046/j.1365-2796.1999.00418.x.

(50) Drazen, J. M.; Israel, E.; O'Byrne, P. M. Treatment of asthma with drugs modifying the leukotriene pathway. *N Engl J Med* **1999**, *340* (3), 197-206. DOI: 10.1056/NEJM199901213400306.

(51) Haberal, I.; Corey, J. P. The role of leukotrienes in nasal allergy. *Otolaryngol Head Neck Surg* **2003**, *129* (3), 274-279. DOI: 10.1016/S0194-5998(03)00601-6.

(52) Zhang, J.; Migita, O.; Koga, M.; Shibasaki, M.; Arinami, T.; Noguchi, E. Determination of structure and transcriptional regulation of CYSLTR1 and an association study with asthma and rhinitis. *Pediatr Allergy Immunol* **2006**, *17* (4), 242-249. DOI: 10.1111/j.1399-3038.2005.00347.x.

(53) Singh, R. K.; Tandon, R.; Dastidar, S. G.; Ray, A. A review on leukotrienes and their receptors with reference to asthma. *J Asthma* **2013**, *50* (9), 922-931. DOI: 10.3109/02770903.2013.823447.

(54) Sarau, H. M.; Ames, R. S.; Chambers, J.; Ellis, C.; Elshourbagy, N.; Foley, J. J.; Schmidt, D. B.; Muccitelli, R. M.; Jenkins, O.; Murdock, P. R.; et al. Identification, molecular cloning, expression, and characterization of a cysteinyl leukotriene receptor. *Mol Pharmacol* **1999**, *56* (3), 657-663. DOI: 10.1124/mol.56.3.657.

(55) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **2020**, *382* (8), 727-733. DOI: 10.1056/NEJMoa2001017.

(56) Toelzer, C.; Gupta, K.; Yadav, S. K. N.; Borucu, U.; Davidson, A. D.; Kavanagh Williamson, M.; Shoemark, D. K.; Garzoni, F.; Staufer, O.; Milligan, R.; et al. Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. *Science* **2020**, *370* (6517), 725-730. DOI: 10.1126/science.abd3255.

(57) Gui, M.; Song, W.; Zhou, H.; Xu, J.; Chen, S.; Xiang, Y.; Wang, X. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res* **2017**, *27* (1), 119-129. DOI: 10.1038/cr.2016.152.

(58) Yuan, Y.; Cao, D.; Zhang, Y.; Ma, J.; Qi, J.; Wang, Q.; Lu, G.; Wu, Y.; Yan, J.; Shi, Y.; et al. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat Commun* **2017**, *8*, 15092. DOI: 10.1038/ncomms15092.

(59) Walls, A. C.; Xiong, X.; Park, Y. J.; Tortorici, M. A.; Snijder, J.; Quispe, J.; Cameroni, E.; Gopal, R.; Dai, M.; Lanzavecchia, A.; et al. Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **2019**, *176* (5), 1026-1039.e1015. DOI: 10.1016/j.cell.2018.12.028.

(60) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *183* (6), 1735. DOI: 10.1016/j.cell.2020.11.032.

(61) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367* (6485), 1444-1448. DOI: 10.1126/science.abb2762.

(62) Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **2020**, *395* (10224), 565-574. DOI: 10.1016/S0140-6736(20)30251-8.

(63) Faheem; Kumar, B. K.; Sekhar, K. V. G. C.; Kunjiappan, S.; Jamalis, J.; Balaña-Fouce, R.; Tekwani, B. L.; Sankaranarayanan, M. Druggable targets of SARS-CoV-2 and treatment opportunities for COVID-19. *Bioorg Chem* **2020**, *104*, 104269. DOI: 10.1016/j.bioorg.2020.104269.

(64) Tang, T.; Bidon, M.; Jaimes, J. A.; Whittaker, G. R.; Daniel, S. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. *Antiviral Res* **2020**, *178*, 104792. DOI: 10.1016/j.antiviral.2020.104792.

(65) Huang, Y.; Yang, C.; Xu, X. F.; Xu, W.; Liu, S. W. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacol Sin* **2020**, *41* (9), 1141-1149. DOI: 10.1038/s41401-020-0485-4.

(66) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K. Y.; et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, *181* (4), 894-904.e899. DOI: 10.1016/j.cell.2020.03.045.

(67) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581* (7807), 215-220. DOI: 10.1038/s41586-020-2180-5.

(68) Henderson, R.; Edwards, R. J.; Mansouri, K.; Janowska, K.; Stalls, V.; Gobeil, S. M. C.; Kopp, M.; Li, D.; Parks, R.; Hsu, A. L.; et al. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat Struct Mol Biol* **2020**, *27* (10), 925-933. DOI: 10.1038/s41594-020-0479-4.

(69) Kirchdoerfer, R. N.; Wang, N.; Pallesen, J.; Wrapp, D.; Turner, H. L.; Cottrell, C. A.; Corbett, K. S.; Graham, B. S.; McLellan, J. S.; Ward, A. B. Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Sci Rep* **2018**, *8* (1), 15701. DOI: 10.1038/s41598-018-34171-7.

(70) Matsuyama, S.; Nao, N.; Shirato, K.; Kawase, M.; Saito, S.; Takayama, I.; Nagata, N.; Sekizuka, T.; Katoh, H.; Kato, F.; et al. Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proc Natl Acad Sci U S A* **2020**, *117* (13), 7001-7003. DOI: 10.1073/pnas.2002589117.

(71) Li, W.; Moore, M. J.; Vasilieva, N.; Sui, J.; Wong, S. K.; Berne, M. A.; Somasundaran, M.; Sullivan, J. L.; Luzuriaga, K.; Greenough, T. C.; et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **2003**, *426* (6965), 450-454. DOI: 10.1038/nature02145.

(72) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, *581* (7807), 221-224. DOI: 10.1038/s41586-020-2179-y.

(73) Harmer, D.; Gilbert, M.; Borman, R.; Clark, K. L. Quantitative mRNA expression profiling of ACE 2, a novel homologue of angiotensin converting enzyme. *FEBS Lett* **2002**, *532* (1-2), 107-110. DOI: 10.1016/s0014-5793(02)03640-2.

(74) Hamming, I.; Timens, W.; Bulthuis, M. L.; Lely, A. T.; Navis, G.; van Goor, H. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J Pathol* **2004**, *203* (2), 631-637. DOI: 10.1002/path.1570.

(75) Bertram, S.; Heurich, A.; Lavender, H.; Gierer, S.; Danisch, S.; Perin, P.; Lucas, J. M.; Nelson, P. S.; Pöhlmann, S.; Soilleux, E. J. Influenza and SARS-coronavirus activating proteases TMPRSS2 and HAT are expressed at multiple sites in human respiratory and gastrointestinal tracts. *PLoS One* **2012**, *7* (4), e35876. DOI: 10.1371/journal.pone.0035876.

(76) Benigni, A.; Cassis, P.; Remuzzi, G. Angiotensin II revisited: new roles in inflammation, immunology and aging. *EMBO Mol Med* **2010**, *2* (7), 247-257. DOI: 10.1002/emmm.201000080.

(77) Sparks, M. A.; Crowley, S. D.; Gurley, S. B.; Mirotsou, M.; Coffman, T. M. Classical Renin-Angiotensin system in kidney physiology. *Compr Physiol* **2014**, *4* (3), 1201-1228. DOI: 10.1002/cphy.c130040.

(78) Vaughan, D. E. Angiotensin and vascular fibrinolytic balance. *Am J Hypertens* **2002**, *15* (1 Pt 2), 3S-8S. DOI: 10.1016/s0895-7061(01)02273-7.

(79) Mehta, P. K.; Griendling, K. K. Angiotensin II cell signaling: physiological and pathological effects in the cardiovascular system. *Am J Physiol Cell Physiol* **2007**, *292* (1), C82-97. DOI: 10.1152/ajpcell.00287.2006.

(80) Simões e Silva, A. C.; Silveira, K. D.; Ferreira, A. J.; Teixeira, M. M. ACE2, angiotensin-(1-7) and Mas receptor axis in inflammation and fibrosis. *Br J Pharmacol* **2013**, *169* (3), 477-492. DOI: 10.1111/bph.12159.

(81) Burrell, L. M.; Johnston, C. I.; Tikellis, C.; Cooper, M. E. ACE2, a new regulator of the renin-angiotensin system. *Trends Endocrinol Metab* **2004**, *15* (4), 166-169. DOI: 10.1016/j.tem.2004.03.001.

(82) Cohen, J. B.; Hanff, T. C.; Bress, A. P.; South, A. M. Relationship Between ACE2 and Other Components of the Renin-Angiotensin System. *Curr Hypertens Rep* **2020**, *22* (7), 44. DOI: 10.1007/s11906-020-01048-y.

(83) Thomas, M. C.; Pickering, R. J.; Tsorotes, D.; Koitka, A.; Sheehy, K.; Bernardi, S.; Toffoli, B.; Nguyen-Huu, T. P.; Head, G. A.; Fu, Y.; et al. Genetic Ace2 deficiency accentuates vascular inflammation and atherosclerosis in the ApoE knockout mouse. *Circ Res* **2010**, *107* (7), 888-897. DOI: 10.1161/CIRCRESAHA.110.219279.

(84) Sahara, M.; Ikutomi, M.; Morita, T.; Minami, Y.; Nakajima, T.; Hirata, Y.; Nagai, R.; Sata, M. Deletion of angiotensin-converting enzyme 2 promotes the development

of atherosclerosis and arterial neointima formation. *Cardiovasc Res* **2014**, *101* (2), 236-246. DOI: 10.1093/cvr/cvt245.

(85) Souza, L. L.; Costa-Neto, C. M. Angiotensin-(1-7) decreases LPS-induced inflammatory response in macrophages. *J Cell Physiol* **2012**, *227* (5), 2117-2122. DOI: 10.1002/jcp.22940.

(86) Xu, J.; Fan, J.; Wu, F.; Huang, Q.; Guo, M.; Lv, Z.; Han, J.; Duan, L.; Hu, G.; Chen, L.; et al. The ACE2/Angiotensin-(1-7)/Mas Receptor Axis: Pleiotropic Roles in Cancer. *Front Physiol* **2017**, *8*, 276. DOI: 10.3389/fphys.2017.00276.

(87) Verdecchia, P.; Cavallini, C.; Spanevello, A.; Angeli, F. The pivotal link between ACE2 deficiency and SARS-CoV-2 infection. *Eur J Intern Med* **2020**, *76*, 14-20. DOI: 10.1016/j.ejim.2020.04.037.

(88) Banu, N.; Panikar, S. S.; Leal, L. R.; Leal, A. R. Protective role of ACE2 and its downregulation in SARS-CoV-2 infection leading to Macrophage Activation Syndrome: Therapeutic implications. *Life Sci* **2020**, *256*, 117905. DOI: 10.1016/j.lfs.2020.117905.

(89) Gheblawi, M.; Wang, K.; Viveiros, A.; Nguyen, Q.; Zhong, J. C.; Turner, A. J.; Raizada, M. K.; Grant, M. B.; Oudit, G. Y. Angiotensin-Converting Enzyme 2: SARS-CoV-2 Receptor and Regulator of the Renin-Angiotensin System: Celebrating the 20th Anniversary of the Discovery of ACE2. *Circ Res* **2020**, *126* (10), 1456-1474. DOI: 10.1161/CIRCRESAHA.120.317015.

(90) Hernández Prada, J. A.; Ferreira, A. J.; Katovich, M. J.; Shenoy, V.; Qi, Y.; Santos, R. A.; Castellano, R. K.; Lampkins, A. J.; Gubala, V.; Ostrov, D. A.; et al. Structure-based identification of small-molecule angiotensin-converting enzyme 2 activators as novel antihypertensive agents. *Hypertension* **2008**, *51* (5), 1312-1317. DOI: 10.1161/HYPERTENSIONAHA.107.108944.

(91) Kulemina, L. V.; Ostrov, D. A. Prediction of off-target effects on angiotensin-converting enzyme 2. *J Biomol Screen* **2011**, *16* (8), 878-885. DOI: 10.1177/1087057111413919.

(92) Qaradakhi, T.; Gadanec, L. K.; McSweeney, K. R.; Tacey, A.; Apostolopoulos, V.; Levinger, I.; Rimarova, K.; Egom, E. E.; Rodrigo, L.; Kruzliak, P.; et al. The potential actions of angiotensin-converting enzyme II (ACE2) activator diminazene aceturate (DIZE) in various diseases. *Clin Exp Pharmacol Physiol* **2020**, *47* (5), 751-758. DOI: 10.1111/1440-1681.13251.

(93) Shenoy, V.; Ferreira, A. J.; Katovich, M.; Raizada, M. K. Angiotensin-converting enzyme 2/Angiotensin-(1-7)/Mas receptor axis: emerging pharmacological target for pulmonary diseases. *The Protective Arm of the Renin Angiotensin System (RAS)* **2015**, 269.

(94) Hasan, H. F.; Elgazzar, E. M.; Mostafa, D. M. Diminazene aceturate extenuate the renal deleterious consequences of angiotensin-II induced by γ-irradiation through boosting ACE2 signaling cascade. *Life Sci* **2020**, *253*, 117749. DOI: 10.1016/j.lfs.2020.117749.

(95) Rajapaksha, I. G.; Mak, K. Y.; Huang, P.; Burrell, L. M.; Angus, P. W.; Herath, C. B. The small molecule drug diminazene aceturate inhibits liver injury and biliary fibrosis in mice. *Sci Rep* **2018**, *8* (1), 10175. DOI: 10.1038/s41598-018-28490-y.

(96) Qi, Y.; Zhang, J.; Cole-Jeffrey, C. T.; Shenoy, V.; Espejo, A.; Hanna, M.; Song, C.; Pepine, C. J.; Katovich, M. J.; Raizada, M. K. Diminazene aceturate enhances angiotensin-converting enzyme 2 activity and attenuates ischemia-induced cardiac pathophysiology. *Hypertension* **2013**, *62* (4), 746-752. DOI: 10.1161/HYPERTENSIONAHA.113.01337.

(97) Fang, Y.; Gao, F.; Liu, Z. Angiotensin-converting enzyme 2 attenuates inflammatory response and oxidative stress in hyperoxic lung injury by regulating NF-κB and Nrf2 pathways. *QJM* **2019**, *112* (12), 914-924. DOI: 10.1093/qjmed/hcz206.

(98) Nicolau, L. A. D.; Nolêto, I. R. S. G.; Medeiros, J. V. R. Could a specific ACE2 activator drug improve the clinical outcome of SARS-CoV-2? A potential pharmacological insight. *Expert Rev Clin Pharmacol* **2020**, *13* (8), 807-811. DOI: 10.1080/17512433.2020.1798760.

(99) Baldissera, M. D.; Sagrillo, M. R.; Grando, T. H.; Dalla Rosa, L.; de Sá, M. F.; da Luz, S. C.; Silveira, S. O.; Nascimento, K.; Peres, D. S.; Copetti, P. M. Cytotoxic and genotoxic effects of the trypanocidal drug diminazene aceturate. *Comparative Clinical Pathology* **2017**, *26* (1), 219-227. DOI: 10.1007/s00580-016-2373-3.

(100) Lengauer, T.; Rarey, M. Computational methods for biomolecular docking. *Current opinion in structural biology* **1996**, *6* (3), 402-406. DOI: 10.1016/S0959-440X(96)80061-3.

(101) Seifert, M.; Kraus, J.; Kramer, B. Virtual high-throughput screening of molecular databases. *Current opinion in drug discovery & development* **2007**, *10* (3), 298-307.

(102) Rollinger, J. M.; Stuppner, H.; Langer, T. Virtual screening for the discovery of bioactive natural products. *Prog Drug Res* **2008**, *65*, 211, 213-249. DOI: 10.1007/978-3-7643-8117-2_6.

(103) Rester, U. From virtuality to reality - Virtual screening in lead discovery and lead optimization: a medicinal chemistry perspective. *Curr Opin Drug Discov Devel* **2008**, *11* (4), 559-568.

(104) McInnes, C. Virtual screening strategies in drug discovery. *Curr Opin Chem Biol* **2007**, *11* (5), 494-502. DOI: 10.1016/j.cbpa.2007.08.033.

(105) Santana, K.; do Nascimento, L. D.; Lima E Lima, A.; Damasceno, V.; Nahum, C.; Braga, R. C.; Lameira, J. Applications of Virtual Screening in Bioprospecting: Facts, Shifts, and Perspectives to Explore the Chemo-Structural Diversity of Natural Products. *Front Chem* **2021**, *9*, 662688. DOI: 10.3389/fchem.2021.662688.

(106) Sun, H. Pharmacophore-based virtual screening. *Curr Med Chem* **2008**, *15* (10), 1018-1024. DOI: 10.2174/092986708784049630.

(107) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic

molecules. *Journal of the American Chemical Society* **1996**, *117* (19), 5179-5197. DOI: 10.1021/ja955032e.

(108) Scott, W. R.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. The GROMOS biomolecular simulation program package. *The Journal of Physical Chemistry A* **1999**, *103* (19), 3596-3607. DOI: 10.1021/jp984217f.

(109) MacKerell Jr, A. D.; Brooks, B.; Brooks III, C. L.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. CHARMM: the energy function and its parameterization. *Encyclopedia of computational chemistry* **2002**, *1*. DOI: 10.1002/0470845015.cfa007.

(110) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved Peptide and Protein Torsional Energetics with the OPLSAA Force Field. *J Chem Theory Comput* **2015**, *11* (7), 3499-3509. DOI: 10.1021/acs.jctc.5b00356.

(111) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics* **1983**, *79* (2), 926-935. DOI: 10.1063/1.445869.

(112) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc Natl Acad Sci U S A* **2002**, *99* (20), 12562-12566. DOI: 10.1073/pnas.202427399.

(113) Tiwary, P.; Parrinello, M. A time-independent free energy estimator for metadynamics. *J Phys Chem B* **2015**, *119* (3), 736-742. DOI: 10.1021/jp504920s.

(114) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys Rev Lett* **2008**, *100* (2), 020603. DOI: 10.1103/PhysRevLett.100.020603.

(115) Bussi, G.; Branduardi, D. Free-energy calculations with metadynamics: Theory and practice. *Rev. Comput. Chem* **2015**, *28*, 1-49.

(116) Piana, S.; Laio, A. A bias-exchange approach to protein folding. *J Phys Chem B* **2007**, *111* (17), 4553-4559. DOI: 10.1021/jp067873l.

(117) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* **2004**, *47* (7), 1750-1759. DOI: 10.1021/jm030644s.

(118) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* **2004**, *47* (7), 1739-1749. DOI: 10.1021/jm0306430.

(119) Jin, L.; Martynowski, D.; Zheng, S.; Wada, T.; Xie, W.; Li, Y. Structural basis for hydroxycholesterols as natural ligands of orphan nuclear receptor RORgamma. *Mol Endocrinol* **2010**, *24* (5), 923-929. DOI: 10.1210/me.2009-0507.

(120) D'Amore, C.; Di Leva, F. S.; Sepe, V.; Renga, B.; Del Gaudio, C.; D'Auria, M. V.; Zampella, A.; Fiorucci, S.; Limongelli, V. Design, synthesis, and biological

evaluation of potent dual agonists of nuclear and membrane bile acid receptors. *J Med Chem* **2014**, *57* (3), 937-954. DOI: 10.1021/jm401873d.

(121) Sepe, V.; Renga, B.; Festa, C.; D'Amore, C.; Masullo, D.; Cipriani, S.; Di Leva, F. S.; Monti, M. C.; Novellino, E.; Limongelli, V.; et al. Modification on ursodeoxycholic acid (UDCA) scaffold. discovery of bile acid derivatives as selective agonists of cell-surface G-protein coupled bile acid receptor 1 (GP-BAR1). *J Med Chem* **2014**, *57* (18), 7687-7701. DOI: 10.1021/jm500889f.

(122) Di Leva, F. S.; Festa, C.; Carino, A.; De Marino, S.; Marchianò, S.; Di Marino, D.; Finamore, C.; Monti, M. C.; Zampella, A.; Fiorucci, S.; et al. Discovery of ((1,2,4-oxadiazol-5-yl)pyrrolidin-3-yl)ureidyl derivatives as selective non-steroidal agonists of the G-protein coupled bile acid receptor-1. *Sci Rep* **2019**, *9* (1), 2504. DOI: 10.1038/s41598-019-38840-z.

(123) De Marino, S.; Carino, A.; Masullo, D.; Finamore, C.; Marchianò, S.; Cipriani, S.; Di Leva, F. S.; Catalanotti, B.; Novellino, E.; Limongelli, V.; et al. Hyodeoxycholic acid derivatives as liver X receptor α and G-protein-coupled bile acid receptor agonists. *Sci Rep* **2017**, *7*, 43290. DOI: 10.1038/srep43290.

(124) Di Leva, F. S.; Festa, C.; Renga, B.; Sepe, V.; Novellino, E.; Fiorucci, S.; Zampella, A.; Limongelli, V. Structure-based drug design targeting the cell membrane receptor GPBAR1: exploiting the bile acid scaffold towards selective agonism. *Sci Rep* **2015**, *5*, 16605. DOI: 10.1038/srep16605.

(125) Macchiarulo, A.; Gioiello, A.; Thomas, C.; Pols, T. W.; Nuti, R.; Ferrari, C.; Giacchè, N.; De Franco, F.; Pruzanski, M.; Auwerx, J.; et al. Probing the Binding Site

of Bile Acids in TGR5. *ACS Med Chem Lett* **2013**, *4* (12), 1158-1162. DOI: 10.1021/ml400247k.

(126) Gertzen, C. G.; Spomer, L.; Smits, S. H.; Häussinger, D.; Keitel, V.; Gohlke, H. Mutational mapping of the transmembrane binding site of the G-protein coupled receptor TGR5 and binding mode prediction of TGR5 agonists. *Eur J Med Chem* **2015**, *104*, 57-72. DOI: 10.1016/j.ejmech.2015.09.024.

(127) Renga, B.; Cipriani, S.; Carino, A.; Simonetti, M.; Zampella, A.; Fiorucci, S. Reversal of Endothelial Dysfunction by GPBAR1 Agonism in Portal Hypertension Involves a AKT/FOXOA1 Dependent Regulation of H2S Generation and Endothelin-1. *PLoS One* **2015**, *10* (11), e0141082. DOI: 10.1371/journal.pone.0141082.

(128) Renga, B.; Bucci, M.; Cipriani, S.; Carino, A.; Monti, M. C.; Zampella, A.; Gargiulo, A.; d'Emmanuele di Villa Bianca, R.; Distrutti, E.; Fiorucci, S. Cystathionine γ-lyase, a H2S-generating enzyme, is a GPBAR1-regulated gene and contributes to vasodilation caused by secondary bile acids. *Am J Physiol Heart Circ Physiol* **2015**, *309* (1), H114-126. DOI: 10.1152/ajpheart.00087.2015.

(129) Biagioli, M.; Carino, A.; Fiorucci, C.; Marchianò, S.; Di Giorgio, C.; Bordoni, M.; Roselli, R.; Baldoni, M.; Distrutti, E.; Zampella, A.; et al. The Bile Acid Receptor GPBAR1 Modulates CCL2/CCR2 Signaling at the Liver Sinusoidal/Macrophage Interface and Reverses Acetaminophen-Induced Liver Toxicity. *J Immunol* **2020**, *204* (9), 2535-2551. DOI: 10.4049/jimmunol.1901427.

(130) Nagata, M.; Saito, K.; Tsuchiya, K.; Sakamoto, Y. Leukotriene D4 upregulates eosinophil adhesion via the cysteinyl leukotriene 1 receptor. *J Allergy Clin Immunol* **2002**, *109* (4), 676-680. DOI: 10.1067/mai.2002.122841.

(131) Boehmler, A. M.; Drost, A.; Jaggy, L.; Seitz, G.; Wiesner, T.; Denzlinger, C.; Kanz, L.; Möhle, R. The CysLT1 ligand leukotriene D4 supports alpha4beta1- and alpha5beta1-mediated adhesion and proliferation of CD34+ hematopoietic progenitor cells. *J Immunol* **2009**, *182* (11), 6789-6798. DOI: 10.4049/jimmunol.0801525.

(132) Meliton, A. Y.; Muñoz, N. M.; Osan, C. M.; Meliton, L. N.; Leff, A. R. Leukotriene D4 activates {beta}2-integrin adhesion in human polymorphonuclear leukocytes. *Eur Respir J* **2010**, *35* (2), 402-409. DOI: 10.1183/09031936.00009309.

(133) Zhou, X.; Cai, J.; Liu, W.; Wu, X.; Gao, C. Cysteinyl leukotriene receptor type 1 (CysLT1R) antagonist zafirlukast protects against TNF-α-induced endothelial inflammation. *Biomed Pharmacother* **2019**, *111*, 452-459. DOI: 10.1016/j.biopha.2018.12.064.

(134) Luginina, A.; Gusach, A.; Marin, E.; Mishin, A.; Brouillette, R.; Popov, P.; Shiriaeva, A.; Besserer-Offroy, É.; Longpré, J. M.; Lyapina, E.; et al. Structure-based mechanism of cysteinyl leukotriene receptor inhibition by antiasthmatic drugs. *Sci Adv* **2019**, *5* (10), eaax2518. DOI: 10.1126/sciadv.aax2518.

(135) Yang, F.; Mao, C.; Guo, L.; Lin, J.; Ming, Q.; Xiao, P.; Wu, X.; Shen, Q.; Guo, S.; Shen, D. D.; et al. Structural basis of GPBAR activation and bile acid recognition. *Nature* **2020**, *587* (7834), 499-504. DOI: 10.1038/s41586-020-2569-1.

(136) Chen, G.; Wang, X.; Ge, Y.; Ma, L.; Chen, Q.; Liu, H.; Du, Y.; Ye, R. D.; Hu, H.; Ren, R. Cryo-EM structure of activated bile acids receptor TGR5 in complex with stimulatory G protein. *Signal Transduct Target Ther* **2020**, *5* (1), 142. DOI: 10.1038/s41392-020-00262-z.

(137) Ballesteros, J. A.; Weinstein, H. [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In *Methods in neurosciences*, Vol. 25; Elsevier, 1995; pp 366-428. DOI: 10.1016/S1043-9471(05)80049-7.

(138) Limongelli, V.; Marinelli, L.; Cosconati, S.; La Motta, C.; Sartini, S.; Mugnaini, L.; Da Settimo, F.; Novellino, E.; Parrinello, M. Sampling protein motion and solvent effect during ligand binding. *Proc Natl Acad Sci U S A* **2012**, *109* (5), 1467-1472. DOI: 10.1073/pnas.1112181108.

(139) Casasnovas, R.; Limongelli, V.; Tiwary, P.; Carloni, P.; Parrinello, M. Unbinding Kinetics of a p38 MAP Kinase Type II Inhibitor from Metadynamics Simulations. *J Am Chem Soc* **2017**, *139* (13), 4780-4788. DOI: 10.1021/jacs.6b12950.

(140) Moraca, F.; Amato, J.; Ortuso, F.; Artese, A.; Pagano, B.; Novellino, E.; Alcaro, S.; Parrinello, M.; Limongelli, V. Ligand binding to telomeric G-quadruplex DNA investigated by funnel-metadynamics simulations. *Proc Natl Acad Sci U S A* **2017**, *114* (11), E2136-E2145. DOI: 10.1073/pnas.1612627114.

(141) Limongelli, V.; Bonomi, M.; Parrinello, M. Funnel metadynamics as accurate binding free-energy method. *Proc Natl Acad Sci U S A* **2013**, *110* (16), 6358-6363. DOI: 10.1073/pnas.1303186110.

(142) Saleh, N.; Ibrahim, P.; Saladino, G.; Gervasio, F. L.; Clark, T. An Efficient Metadynamics-Based Protocol To Model the Binding Affinity and the Transition State Ensemble of G-Protein-Coupled Receptor Ligands. *J Chem Inf Model* **2017**, *57* (5), 1210-1217. DOI: 10.1021/acs.jcim.6b00772.

(143) Raniolo, S.; Limongelli, V. Ligand binding free-energy calculations with funnel metadynamics. *Nat Protoc* **2020**, *15* (9), 2837-2866. DOI: 10.1038/s41596-020-0342-4.

(144) Ghosh, A. K.; Brindisi, M.; Shahabi, D.; Chapman, M. E.; Mesecar, A. D. Drug Development and Medicinal Chemistry Efforts toward SARS-Coronavirus and Covid-19 Therapeutics. *ChemMedChem* **2020**, *15* (11), 907-932. DOI: 10.1002/cmdc.202000223.

(145) Smith, M.; Smith, J. C. Repurposing therapeutics for COVID-19: Supercomputer-based docking to the SARS-CoV-2 viral spike protein and viral spike protein-human ACE2 interface. **2020**. DOI: 10.26434/chemrxiv.11871402.v4.

(146) Senathilake, K.; Samarakoon, S.; Tennekoon, K. Virtual screening of inhibitors against spike glycoprotein of 2019 novel corona virus: A drug repurposing approach. **2020**. DOI: 10.20944/preprints202003.0042.v2.

(147) Utomo, R. Y.; Ikawati, M.; Meiyanto, E. Revealing the potency of citrus and galangal constituents to halt SARS-CoV-2 infection. **2020**. DOI: 10.20944/preprints202003.0214.v1.

(148) Wang, C.; Wang, S.; Li, D.; Wei, D. Q.; Zhao, J.; Wang, J. Human Intestinal Defensin 5 Inhibits SARS-CoV-2 Invasion by Cloaking ACE2. *Gastroenterology* **2020**, *159* (3), 1145-1147.e1144. DOI: 10.1053/j.gastro.2020.05.015.

(149) Zhou, Y.; Hou, Y.; Shen, J.; Huang, Y.; Martin, W.; Cheng, F. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov* **2020**, *6*, 14. DOI: 10.1038/s41421-020-0153-3.

(150) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C. L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367* (6483), 1260-1263. DOI: 10.1126/science.abb2507.

(151) Bertoni, M.; Kiefer, F.; Biasini, M.; Bordoli, L.; Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep* **2017**, *7* (1), 10480. DOI: 10.1038/s41598-017-09654-8.

(152) Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics* **2009**, *10*, 168. DOI: 10.1186/1471-2105-10-168.

(153) Carino, A.; Moraca, F.; Fiorillo, B.; Marchianò, S.; Sepe, V.; Biagioli, M.; Finamore, C.; Bozza, S.; Francisci, D.; Distrutti, E.; et al. Hijacking SARS-CoV-2/ACE2 Receptor Interaction by Natural and Semi-synthetic Steroidal Agents Acting on Functional Pockets on the Receptor Binding Domain. *Front Chem* **2020**, *8*, 572885. DOI: 10.3389/fchem.2020.572885.

(154) FDA- approved Drug Library, (2020). Available online at: https://www.selleckchem.com/screening/fda-approved-drug-library.html (accessed March 4, 2020).

(155) Drugbank (2020). Available online at: https://www.drugbank.ca/releases/latest#structures (accessed March 5, 2020).

(156) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* **2009**, *30* (16), 2785-2791. DOI: 10.1002/jcc.21256.

(157) Forli, S.; Huey, R.; Pique, M. E.; Sanner, M. F.; Goodsell, D. S.; Olson, A. J. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat Protoc* **2016**, *11* (5), 905-919. DOI: 10.1038/nprot.2016.051.

(158) Sepe, V.; Distrutti, E.; Limongelli, V.; Fiorucci, S.; Zampella, A. Steroidal scaffolds as FXR and GPBAR1 ligands: from chemistry to therapeutical application. *Future Med Chem* **2015**, *7* (9), 1109-1135. DOI: 10.4155/fmc.15.54.

(159) De Marino, S.; Festa, C.; Sepe, V.; Zampella, A. Chemistry and Pharmacology of GPBAR1 and FXR Selective Agonists, Dual Agonists, and Antagonists. *Handb Exp Pharmacol* **2019**, *256*, 137-165. DOI: 10.1007/164_2019_237.

(160) Distrutti, E.; Santucci, L.; Cipriani, S.; Renga, B.; Schiaroli, E.; Ricci, P.; Donini, A.; Fiorucci, S. Bile acid activated receptors are targets for regulation of integrity of

gastrointestinal mucosa. *J Gastroenterol* **2015**, *50* (7), 707-719. DOI: 10.1007/s00535-015-1041-8.

(161) Wang, L. Y.; Cheng, K. C.; Li, Y.; Niu, C. S.; Cheng, J. T.; Niu, H. S. Glycyrrhizic acid increases glucagon like peptide-1 secretion via TGR5 activation in type 1-like diabetic rats. *Biomed Pharmacother* **2017**, *95*, 599-604. DOI: 10.1016/j.biopha.2017.08.087.

(162) Rezanka, T.; Siristova, L.; Sigler, K. Sterols and triterpenoids with antiviral activity. *Anti-Infective Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Infective Agents)* **2009**, *8* (3), 193-210.

(163) Yang, Y.; He, H. J.; Chang, H.; Yu, Y.; Yang, M. B.; He, Y.; Fan, Z. C.; Iyer, S. S.; Yu, P. Multivalent oleanolic acid human serum albumin conjugate as nonglycosylated neomucin for influenza virus capture and entry inhibition. *Eur J Med Chem* **2018**, *143*, 1723-1731. DOI: 10.1016/j.ejmech.2017.10.070.

(164) Hong, E. H.; Song, J. H.; Kang, K. B.; Sung, S. H.; Ko, H. J.; Yang, H. Anti-Influenza Activity of Betulinic Acid from Zizyphus jujuba on Influenza A/PR/8 Virus. *Biomol Ther (Seoul)* **2015**, *23* (4), 345-349. DOI: 10.4062/biomolther.2015.019.

(165) Fiorucci, S.; Biagioli, M.; Zampella, A.; Distrutti, E. Bile Acids Activated Receptors Regulate Innate Immunity. *Front Immunol* **2018**, *9*, 1853. DOI: 10.3389/fimmu.2018.01853.

(166) Towler, P.; Staker, B.; Prasad, S. G.; Menon, S.; Tang, J.; Parsons, T.; Ryan, D.; Fisher, M.; Williams, D.; Dales, N. A.; et al. ACE2 X-ray structures reveal a large

hinge-bending motion important for inhibitor binding and catalysis. *J Biol Chem* **2004**, *279* (17), 17996-18007. DOI: 10.1074/jbc.M311191200.

(167) Limongelli, V. Ligand binding free energy and kinetics calculation in 2020. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2020**, *10* (4), e1455. DOI: 10.1002/wcms.1455.

(168) Fiorillo, B.; Marchianò, S.; Moraca, F.; Sepe, V.; Carino, A.; Rapacciuolo, P.; Biagioli, M.; Limongelli, V.; Zampella, A.; Catalanotti, B.; et al. Discovery of Bile Acid Derivatives as Potent ACE2 Activators by Virtual Screening and Essential Dynamics. *J Chem Inf Model* **2022**, *62* (1), 196-209. DOI: 10.1021/acs.jcim.1c01126.

(169) Hildebrandt, F.; Benzing, T.; Katsanis, N. Ciliopathies. *N Engl J Med* **2011**, *364* (16), 1533-1543. DOI: 10.1056/NEJMra1010172.

(170) Ishikawa, H.; Marshall, W. F. Ciliogenesis: building the cell's antenna. *Nat Rev Mol Cell Biol* **2011**, *12* (4), 222-234. DOI: 10.1038/nrm3085.

(171) Oh, E. C.; Vasanth, S.; Katsanis, N. Metabolic regulation and energy homeostasis through the primary Cilium. *Cell Metab* **2015**, *21* (1), 21-31. DOI: 10.1016/j.cmet.2014.11.019.

(172) Hilgendorf, K. I.; Johnson, C. T.; Jackson, P. K. The primary cilium as a cellular receiver: organizing ciliary GPCR signaling. *Curr Opin Cell Biol* **2016**, *39*, 84-92. DOI: 10.1016/j.ceb.2016.02.008.

(173) Nachury, M. V.; Mick, D. U. Establishing and regulating the composition of cilia for signal transduction. *Nat Rev Mol Cell Biol* **2019**, *20* (7), 389-405. DOI: 10.1038/s41580-019-0116-4.

(174) Valente, E. M.; Rosti, R. O.; Gibbs, E.; Gleeson, J. G. Primary cilia in neurodevelopmental disorders. *Nat Rev Neurol* **2014**, *10* (1), 27-36. DOI: 10.1038/nrneurol.2013.247.

(175) Reiter, J. F.; Leroux, M. R. Genes and molecular pathways underpinning ciliopathies. *Nat Rev Mol Cell Biol* **2017**, *18* (9), 533-547. DOI: 10.1038/nrm.2017.60.

(176) Macca, M.; Franco, B. The molecular basis of oral-facial-digital syndrome, type 1. *Am J Med Genet C Semin Med Genet* **2009**, *151C* (4), 318-325. DOI: 10.1002/ajmg.c.30224.

(177) Bruel, A. L.; Franco, B.; Duffourd, Y.; Thevenon, J.; Jego, L.; Lopez, E.; Deleuze, J. F.; Doummar, D.; Giles, R. H.; Johnson, C. A.; et al. Fifteen years of research on oral-facial-digital syndromes: from 1 to 16 causal genes. *J Med Genet* **2017**, *54* (6), 371-380. DOI: 10.1136/jmedgenet-2016-104436.

(178) Tang, Z.; Lin, M. G.; Stowe, T. R.; Chen, S.; Zhu, M.; Stearns, T.; Franco, B.; Zhong, Q. Autophagy promotes primary ciliogenesis by removing OFD1 from centriolar satellites. *Nature* **2013**, *502* (7470), 254-257. DOI: 10.1038/nature12606.

(179) Anvarian, Z.; Mykytyn, K.; Mukhopadhyay, S.; Pedersen, L. B.; Christensen, S. T. Cellular signalling by primary cilia in development, organ function and disease. *Nat Rev Nephrol* **2019**, *15* (4), 199-219. DOI: 10.1038/s41581-019-0116-9.

(180) Jeng, K. S.; Chang, C. F.; Lin, S. S. Sonic Hedgehog Signaling in Organogenesis, Tumors, and Tumor Microenvironments. *Int J Mol Sci* **2020**, *21* (3). DOI: 10.3390/ijms21030758.

(181) Taylor, S. S.; Zhang, P.; Steichen, J. M.; Keshwani, M. M.; Kornev, A. P. PKA: lessons learned after twenty years. *Biochim Biophys Acta* **2013**, *1834* (7), 1271-1278. DOI: 10.1016/j.bbapap.2013.03.007.

(182) Newton, A. C.; Bootman, M. D.; Scott, J. D. Second Messengers. *Cold Spring Harb Perspect Biol* **2016**, *8* (8). DOI: 10.1101/cshperspect.a005926.

(183) Rinaldi, L.; Delle Donne, R.; Catalanotti, B.; Torres-Quesada, O.; Enzler, F.; Moraca, F.; Nisticò, R.; Chiuso, F.; Piccinin, S.; Bachmann, V.; et al. Feedback inhibition of cAMP effector signaling by a chaperone-assisted ubiquitin system. *Nat Commun* **2019**, *10* (1), 2572. DOI: 10.1038/s41467-019-10037-y.

(184) Yang, L.; McKnight, G. S. Hypothalamic PKA regulates leptin sensitivity and adiposity. *Nat Commun* **2015**, *6*, 8237. DOI: 10.1038/ncomms9237.

(185) Jones, B. W.; Deem, J.; Younts, T. J.; Weisenhaus, M.; Sanford, C. A.; Slack, M. C.; Chin, J.; Nachmanson, D.; McKennon, A.; Castillo, P. E.; et al. Targeted deletion of AKAP7 in dentate granule cells impairs spatial discrimination. *Elife* **2016**, *5*. DOI: 10.7554/eLife.20695.

(186) Reggi, E.; Diviani, D. The role of A-kinase anchoring proteins in cancer development. *Cell Signal* **2017**, *40*, 143-155. DOI: 10.1016/j.cellsig.2017.09.011.

(187) Rinaldi, L.; Sepe, M.; Delle Donne, R.; Conte, K.; Arcella, A.; Borzacchiello, D.; Amente, S.; De Vita, F.; Porpora, M.; Garbi, C.; et al. Mitochondrial AKAP1 supports mTOR pathway and tumor growth. *Cell Death Dis* **2017**, *8* (6), e2842. DOI: 10.1038/cddis.2017.241.

(188) Torres-Quesada, O.; Mayrhofer, J. E.; Stefan, E. The many faces of compartmentalized PKA signalosomes. *Cell Signal* **2017**, *37*, 1-11. DOI: 10.1016/j.cellsig.2017.05.012.

(189) Rinaldi, L.; Delle Donne, R.; Borzacchiello, D.; Insabato, L.; Feliciello, A. The role of compartmentalized signaling pathways in the control of mitochondrial activities in cancer cells. *Biochim Biophys Acta Rev Cancer* **2018**, *1869* (2), 293-302. DOI: 10.1016/j.bbcan.2018.04.004.

(190) Bucko, P. J.; Lombard, C. K.; Rathbun, L.; Garcia, I.; Bhat, A.; Wordeman, L.; Smith, F. D.; Maly, D. J.; Hehnly, H.; Scott, J. D. Subcellular drug targeting illuminates local kinase action. *Elife* **2019**, *8*. DOI: 10.7554/eLife.52220.

(191) Lignitto, L.; Carlucci, A.; Sepe, M.; Stefan, E.; Cuomo, O.; Nisticò, R.; Scorziello, A.; Savoia, C.; Garbi, C.; Annunziato, L.; et al. Control of PKA stability and signalling by the RING ligase praja2. *Nat Cell Biol* **2011**, *13* (4), 412-422. DOI: 10.1038/ncb2209.

(192) Lignitto, L.; Sepe, M.; Carlucci, A.; Feliciello, A. An intimate connection between ubiquitination and compartmentalized cAMP signaling. *Cell Cycle* **2011**, *10* (13), 2051-2052. DOI: 10.4161/cc.10.13.15839.

(193) Lignitto, L.; Arcella, A.; Sepe, M.; Rinaldi, L.; Delle Donne, R.; Gallo, A.; Stefan, E.; Bachmann, V. A.; Oliva, M. A.; Tiziana Storlazzi, C.; et al. Proteolysis of MOB1 by the ubiquitin ligase praja2 attenuates Hippo signalling and supports glioblastoma growth. *Nat Commun* **2013**, *4*, 1822. DOI: 10.1038/ncomms2791.

(194) Sepe, M.; Lignitto, L.; Porpora, M.; Delle Donne, R.; Rinaldi, L.; Belgianni, G.; Colucci, G.; Cuomo, O.; Viggiano, D.; Scorziello, A.; et al. Proteolytic control of neurite outgrowth inhibitor NOGO-A by the cAMP/PKA pathway. *Proc Natl Acad Sci U S A* **2014**, *111* (44), 15729-15734. DOI: 10.1073/pnas.1410274111.

(195) Zhang, P.; Fu, W. Y.; Fu, A. K.; Ip, N. Y. S-nitrosylation-dependent proteasomal degradation restrains Cdk5 activity to regulate hippocampal synaptic strength. *Nat Commun* **2015**, *6*, 8665. DOI: 10.1038/ncomms9665.

(196) Rinaldi, L.; Delle Donne, R.; Sepe, M.; Porpora, M.; Garbi, C.; Chiuso, F.; Gallo, A.; Parisi, S.; Russo, L.; Bachmann, V.; et al. praja2 regulates KSR1 stability and mitogenic signaling. *Cell Death Dis* **2016**, *7*, e2230. DOI: 10.1038/cddis.2016.109.

(197) Song, J.; Wang, T.; Chi, X.; Wei, X.; Xu, S.; Yu, M.; He, H.; Ma, J.; Li, X.; Du, J.; et al. Kindlin-2 Inhibits the Hippo Signaling Pathway by Promoting Degradation of MOB1. *Cell Rep* **2019**, *29* (11), 3664-3677.e3665. DOI: 10.1016/j.celrep.2019.11.035.

(198) Mukherjee, S.; Jansen, V.; Jikeli, J. F.; Hamzeh, H.; Alvarez, L.; Dombrowski, M.; Balbach, M.; Strünker, T.; Seifert, R.; Kaupp, U. B.; et al. A novel biosensor to study cAMP dynamics in cilia and flagella. *Elife* **2016**, *5*. DOI: 10.7554/eLife.14052.

(199) Siljee, J. E.; Wang, Y.; Bernard, A. A.; Ersoy, B. A.; Zhang, S.; Marley, A.; Von Zastrow, M.; Reiter, J. F.; Vaisse, C. Subcellular localization of MC4R with ADCY3 at neuronal primary cilia underlies a common pathway for genetic predisposition to obesity. *Nat Genet* **2018**, *50* (2), 180-185. DOI: 10.1038/s41588-017-0020-9.

(200) Sherpa, R. T.; Mohieldin, A. M.; Pala, R.; Wachten, D.; Ostrom, R. S.; Nauli, S. M. Sensory primary cilium is a responsive cAMP microdomain in renal epithelia. *Sci Rep* **2019**, *9* (1), 6523. DOI: 10.1038/s41598-019-43002-2.

(201) Chen, Y.; Yue, S.; Xie, L.; Pu, X. H.; Jin, T.; Cheng, S. Y. Dual Phosphorylation of suppressor of fused (Sufu) by PKA and GSK3beta regulates its stability and localization in the primary cilium. *J Biol Chem* **2011**, *286* (15), 13502-13511. DOI: 10.1074/jbc.M110.217604.

(202) Vuolo, L.; Herrera, A.; Torroba, B.; Menendez, A.; Pons, S. Ciliary adenylyl cyclases control the Hedgehog pathway. *J Cell Sci* **2015**, *128* (15), 2928-2937. DOI: 10.1242/jcs.172635.

(203) Porpora, M.; Sauchella, S.; Rinaldi, L.; Delle Donne, R.; Sepe, M.; Torres-Quesada, O.; Intartaglia, D.; Garbi, C.; Insabato, L.; Santoriello, M.; et al. Counterregulation of cAMP-directed kinase activities controls ciliogenesis. *Nat Commun* **2018**, *9* (1), 1224. DOI: 10.1038/s41467-018-03643-9.

(204) Senatore, E.; Chiuso, F.; Rinaldi, L.; Intartaglia, D.; Delle Donne, R.; Pedone, E.; Catalanotti, B.; Pirone, L.; Fiorillo, B.; Moraca, F.; et al. The TBC1D31/praja2 complex controls primary ciliogenesis through PKA-directed OFD1 ubiquitylation. *EMBO J* **2021**, *40* (10), e106503. DOI: 10.15252/embj.2020106503.

(205) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **2010**, *31* (2), 455-461. DOI: 10.1002/jcc.21334.

(206) London, N.; Raveh, B.; Cohen, E.; Fathi, G.; Schueler-Furman, O. Rosetta FlexPepDock web server--high resolution modeling of peptide-protein interactions. *Nucleic Acids Res* **2011**, *39* (Web Server issue), W249-253. DOI: 10.1093/nar/gkr431.

(207) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181* (2), 281-292.e286. DOI: 10.1016/j.cell.2020.02.058.

(208) Li, G.; He, X.; Zhang, L.; Ran, Q.; Wang, J.; Xiong, A.; Wu, D.; Chen, F.; Sun, J.; Chang, C. Assessing ACE2 expression patterns in lung tissues in the pathogenesis of COVID-19. *J Autoimmun* **2020**, *112*, 102463. DOI: 10.1016/j.jaut.2020.102463.

(209) Valyaeva, A. A.; Zharikova, A. A.; Kasianov, A. S.; Vassetzky, Y. S.; Sheval, E. V. Expression of SARS-CoV-2 entry factors in lung epithelial stem cells and its potential implications for COVID-19. *Sci Rep* **2020**, *10* (1), 17772. DOI: 10.1038/s41598-020-74598-5.

(210) Ziegler, C. G. K.; Allon, S. J.; Nyquist, S. K.; Mbano, I. M.; Miao, V. N.; Tzouanas, C. N.; Cao, Y.; Yousif, A. S.; Bals, J.; Hauser, B. M.; et al. SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. *Cell* **2020**, *181* (5), 1016-1035.e1019. DOI: 10.1016/j.cell.2020.04.035.

(211) Xu, J.; Chu, M.; Zhong, F.; Tan, X.; Tang, G.; Mai, J.; Lai, N.; Guan, C.; Liang, Y.; Liao, G. Digestive symptoms of COVID-19 and expression of ACE2 in digestive tract organs. *Cell Death Discov* **2020**, *6*, 76. DOI: 10.1038/s41420-020-00307-w.

(212) Potdar, A. A.; Dube, S.; Naito, T.; Li, K.; Botwin, G.; Haritunians, T.; Li, D.; Casero, D.; Yang, S.; Bilsborough, J.; et al. Altered Intestinal ACE2 Levels Are Associated With Inflammation, Severe Disease, and Response to Anti-Cytokine Therapy in Inflammatory Bowel Disease. *Gastroenterology* **2021**, *160* (3), 809-822.e807. DOI: 10.1053/j.gastro.2020.10.041.

(213) Verstockt, B.; Verstockt, S.; Abdu Rahiman, S.; Ke, B. J.; Arnauts, K.; Cleynen, I.; Sabino, J.; Ferrante, M.; Matteoli, G.; Vermeire, S. Intestinal Receptor of SARS-CoV-2 in Inflamed IBD Tissue Seems Downregulated by HNF4A in Ileum and Upregulated by Interferon Regulating Factors in Colon. *J Crohns Colitis* **2021**, *15* (3), 485-498. DOI: 10.1093/ecco-jcc/jjaa185.

(214) Higgins, P. D. R.; Ng, S.; Danese, S.; Rao, K. The Risk of SARS-CoV-2 in Immunosuppressed IBD Patients. *Crohns Colitis 360* **2020**, *2* (2), otaa026. DOI: 10.1093/crocol/otaa026.

(215) Lamas, B.; Natividad, J. M.; Sokol, H. Aryl hydrocarbon receptor and intestinal immunity. *Mucosal Immunol* **2018**, *11* (4), 1024-1038. DOI: 10.1038/s41385-018-0019-2.

(216) Denison, M. S.; Soshilov, A. A.; He, G.; DeGroot, D. E.; Zhao, B. Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action

of the aryl hydrocarbon (dioxin) receptor. *Toxicol Sci* **2011**, *124* (1), 1-22. DOI: 10.1093/toxsci/kfr218.

(217) Zhou, L. AHR Function in Lymphocytes: Emerging Concepts. *Trends Immunol* **2016**, *37* (1), 17-31. DOI: 10.1016/j.it.2015.11.007.

(218) Garg, M.; Royce, S. G.; Tikellis, C.; Shallue, C.; Batu, D.; Velkoska, E.; Burrell, L. M.; Patel, S. K.; Beswick, L.; Jackson, A.; et al. Imbalance of the renin-angiotensin system may contribute to inflammation and fibrosis in IBD: a novel therapeutic target? *Gut* **2020**, *69* (5), 841-851. DOI: 10.1136/gutjnl-2019-318512.

(219) Hashimoto, T.; Perlot, T.; Rehman, A.; Trichereau, J.; Ishiguro, H.; Paolino, M.; Sigl, V.; Hanada, T.; Hanada, R.; Lipinski, S.; et al. ACE2 links amino acid malnutrition to microbial ecology and intestinal inflammation. *Nature* **2012**, *487* (7408), 477-481. DOI: 10.1038/nature11228.

(220) Mazza, G. Compositional and functional properties of saskatoon berry and blueberry. *International journal of fruit science* **2005**, *5* (3), 101-120. DOI: 10.1300/J492v05n03_10.

(221) Jennings, A.; Welch, A. A.; Fairweather-Tait, S. J.; Kay, C.; Minihane, A. M.; Chowienczyk, P.; Jiang, B.; Cecelja, M.; Spector, T.; Macgregor, A.; et al. Higher anthocyanin intake is associated with lower arterial stiffness and central blood pressure in women. *Am J Clin Nutr* **2012**, *96* (4), 781-788. DOI: 10.3945/ajcn.112.042036.

(222) Pervin, M.; Hasnat, M. A.; Lim, J. H.; Lee, Y. M.; Kim, E. O.; Um, B. H.; Lim, B. O. Preventive and therapeutic effects of blueberry (Vaccinium corymbosum) extract

against DSS-induced ulcerative colitis by regulation of antioxidant and inflammatory mediators. *J Nutr Biochem* **2016**, *28*, 103-113. DOI: 10.1016/j.jnutbio.2015.10.006.

(223) Roth, S.; Spalinger, M. R.; Gottier, C.; Biedermann, L.; Zeitz, J.; Lang, S.; Weber, A.; Rogler, G.; Scharl, M. Bilberry-Derived Anthocyanins Modulate Cytokine Expression in the Intestine of Patients with Ulcerative Colitis. *PLoS One* **2016**, *11* (5), e0154817. DOI: 10.1371/journal.pone.0154817.

(224) Biagioli, M.; Carino, A.; Fiorucci, C.; Annunziato, G.; Marchianò, S.; Bordoni, M.; Roselli, R.; Giorgio, C. D.; Castiglione, F.; Ricci, P.; et al. The Aryl Hydrocarbon Receptor (AhR) Mediates the Counter-Regulatory Effects of Pelargonidins in Models of Inflammation and Metabolic Dysfunctions. *Nutrients* **2019**, *11* (8). DOI: 10.3390/nu11081820.

(225) Biagioli, M.; Marchianò, S.; Roselli, R.; Di Giorgio, C.; Bellini, R.; Bordoni, M.; Gidari, A.; Sabbatini, S.; Francisci, D.; Fiorillo, B.; et al. Discovery of a AHR pelargonidin agonist that counter-regulates Ace2 expression and attenuates ACE2-SARS-CoV-2 interaction. *Biochem Pharmacol* **2021**, *188*, 114564. DOI: 10.1016/j.bcp.2021.114564.

(226) Chiappetta, S.; Sharma, A. M.; Bottino, V.; Stier, C. COVID-19 and the role of chronic inflammation in patients with obesity. *Int J Obes (Lond)* **2020**, *44* (8), 1790-1792. DOI: 10.1038/s41366-020-0597-4.

(227) McNeill, J. N.; Lau, E. S.; Paniagua, S. M.; Liu, E. E.; Wang, J. K.; Bassett, I. V.; Selvaggi, C. A.; Lubitz, S. A.; Foulkes, A. S.; Ho, J. E. The role of obesity in

inflammatory markers in COVID-19 patients. *Obes Res Clin Pract* **2021**, *15* (1), 96-99. DOI: 10.1016/j.orcp.2020.12.004.

(228) Shoemark, D. K.; Colenso, C. K.; Toelzer, C.; Gupta, K.; Sessions, R. B.; Davidson, A. D.; Berger, I.; Schaffitzel, C.; Spencer, J.; Mulholland, A. J. Molecular Simulations suggest Vitamins, Retinoids and Steroids as Ligands of the Free Fatty Acid Pocket of the SARS-CoV-2 Spike Protein. *Angew Chem Int Ed Engl* **2021**, *60* (13), 7098-7110. DOI: 10.1002/anie.202015639.

(229) Allam, A. E.; Assaf, H. K.; Hassan, H. A.; Shimizu, K.; Elshaier, Y. A. An in silico perception for newly isolated flavonoids from peach fruit as privileged avenue for a countermeasure outbreak of COVID-19. *RSC advances* **2020**, *10* (50), 29983-29998. DOI: 10.1039/D0RA05265E.

(230) Kallen, J.; Izaac, A.; Be, C.; Arista, L.; Orain, D.; Kaupmann, K.; Guntermann, C.; Hoegenauer, K.; Hintermann, S. Structural States of RORγt: X-ray Elucidation of Molecular Mechanisms and Binding Interactions for Natural and Synthetic Compounds. *ChemMedChem* **2017**, *12* (13), 1014-1021. DOI: 10.1002/cmdc.201700278.

(231) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res* **2012**, *40* (Web Server issue), W537-541. DOI: 10.1093/nar/gks375.

(232) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening

enrichments. *J Comput Aided Mol Des* **2013**, *27* (3), 221-234. DOI: 10.1007/s10822-013-9644-8.

(233) Schrödinger Release 2019-1: Maestro, New York, NY: Schrödinger, LLC, 2**019.**

(234) Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des* **2007**, *21* (12), 681-691. DOI: 10.1007/s10822-007-9133-z.

(235) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J Comput Chem* **2005**, *26* (16), 1752-1780. DOI: 10.1002/jcc.20292.

(236) Grippo, L.; Lucidi, S. A globally convergent version of the Polak-Ribiere conjugate gradient method. *Mathematical Programming* **1997**, *78* (3), 375-391. DOI: 10.1007/BF02614362.

(237) Glide, version 7.1. New York, NY: Schrödinger, LLC, 2**019.**

(238) Baroni, M.; Cruciani, G.; Sciabola, S.; Perruccio, F.; Mason, J. S. A common reference framework for analyzing/comparing proteins and ligands. Fingerprints for Ligands and Proteins (FLAP): theory and application. *J Chem Inf Model* **2007**, *47* (2), 279-294. DOI: 10.1021/ci600253e.

(239) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J Comput Chem* **2005**, *26* (16), 1781-1802. DOI: 10.1002/jcc.20289.

(240) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **2015**, *11* (8), 3696-3713. DOI: 10.1021/acs.jctc.5b00255.

(241) 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. DOI: 10.1002/jcc.20035.

(242) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* **2008**, *29* (11), 1859-1865. DOI: 10.1002/jcc.20945.

(243) Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. *Biophys J* **2009**, *97* (1), 50-58. DOI: 10.1016/j.bpj.2009.04.013.

(244) Dickson, C. J.; Madej, B. D.; Skjevik, A. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: The Amber Lipid Force Field. *J Chem Theory Comput* **2014**, *10* (2), 865-879. DOI: 10.1021/ct4010307.

(245) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics* **2016**, *54*, 5.6.1-5.6.37. DOI: 10.1002/cpbi.3.

(246) Fiser, A.; Do, R. K.; Sali, A. Modeling of loops in protein structures. *Protein Sci* **2000**, *9* (9), 1753-1773. DOI: 10.1110/ps.9.9.1753.

(247) Gusach, A.; Luginina, A.; Marin, E.; Brouillette, R. L.; Besserer-Offroy, É.; Longpré, J. M.; Ishchenko, A.; Popov, P.; Patel, N.; Fujimoto, T.; et al. Structural basis of ligand selectivity and disease mutations in cysteinyl leukotriene receptors. *Nat Commun* **2019**, *10* (1), 5573. DOI: 10.1038/s41467-019-13348-2.

(248) Buchan, D. W. A.; Jones, D. T. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res* **2019**, *47* (W1), W402-W407. DOI: 10.1093/nar/gkz297.

(249) Heffernan, R.; Paliwal, K.; Lyons, J.; Singh, J.; Yang, Y.; Zhou, Y. Single-sequence-based prediction of protein secondary structures and solvent accessibility by deep whole-sequence learning. *J Comput Chem* **2018**, *39* (26), 2210-2216. DOI: 10.1002/jcc.25534.

(250) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19-25. DOI: 10.1016/j.softx.2015.06.001.

(251) Glide, version 7.1. New York, NY: Schrödinger, LLC, 2**019.**

(252) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29* (7), 845-854. DOI: 10.1093/bioinformatics/btt055.

(253) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; Van Gunsteren, W. F.; Mark, A. E. Peptide folding: when simulation meets experiment. *Angewandte Chemie International Edition* **1999**, *38* (1-2), 236-240. DOI: 10.1002/(SICI)1521-3773(19990115)38:1/2<236::AID-ANIE236>3.0.CO;2-M.

(254) Lindahl, E.; Abraham, M.; Hess, B.; van der Spoel, D. Gromacs 2020 Manual. Version: 2020.

(255) PLUMED consortium. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat Methods* **2019**, *16* (8), 670-673. DOI: 10.1038/s41592-019-0506-8.

(256) FDA-approved Drug Library, (2020). Available online at: (accessed March 4, 2020).

(257) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *J Cheminform* **2011**, *3*, 33. DOI: 10.1186/1758-2946-3-33.

(258) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic acids research* **2003**, *31* (13), 3381-3385. DOI: 10.1093/nar/gkg520.

(259) Schmidtke, P.; Le Guilloux, V.; Maupetit, J.; Tufféry, P. fpocket: online tools for protein ensemble pocket detection and tracking. *Nucleic Acids Res* **2010**, *38* (Web Server issue), W582-589. DOI: 10.1093/nar/gkq383.

(260) Lee, T. S.; Cerutti, D. S.; Mermelstein, D.; Lin, C.; LeGrand, S.; Giese, T. J.; Roitberg, A.; Case, D. A.; Walker, R. C.; York, D. M. GPU-Accelerated Molecular Dynamics and Free Energy Methods in Amber18: Performance Enhancements and New Features. *J Chem Inf Model* **2018**, *58* (10), 2043-2050. DOI: 10.1021/acs.jcim.8b00462.

(261) Glykos, N. M. Software news and updates. Carma: a molecular dynamics analysis program. *J Comput Chem* **2006**, *27* (14), 1765-1768. DOI: 10.1002/jcc.20482.

(262) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **1996**, *14* (1), 33-38, 27-38. DOI: 10.1016/0263-7855(96)00018-5.

(263) Girvan, M.; Newman, M. E. Community structure in social and biological networks. *Proc Natl Acad Sci U S A* **2002**, *99* (12), 7821-7826. DOI: 10.1073/pnas.122653799.

(264) **Schrödinger Release 2021-4** : QikProp, Schrödinger, LLC, New York, NY, 2021.

(265) Santos-Martins, D.; Forli, S.; Ramos, M. J.; Olson, A. J. AutoDock4(Zn): an improved AutoDock force field for small-molecule docking to zinc metalloproteins. *J Chem Inf Model* **2014**, *54* (8), 2371-2379. DOI: 10.1021/ci500209e.

(266) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E. I.; 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.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. AMBER 18. University of California: San Francisco (USA), 2**018.**

(267) 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. *The Journal of Physical Chemistry* **1993**, *97* (40), 10269-10280.

(268) Gaussian 16. Gaussian, Inc.: Wallingford CT, 2**016.**

(269) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* **2006**, *25* (2), 247-260. DOI: 10.1016/j.jmgm.2005.12.005.

(270) He, X.; Man, V. H.; Yang, W.; Lee, T. S.; Wang, J. A fast and high-quality charge model for the next generation general AMBER force field. *J Chem Phys* **2020**, *153* (11), 114502. DOI: 10.1063/5.0019056.

(271) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J Chem Theory Comput* **2013**, *9* (6), 2733-2748. DOI: 10.1021/ct400146w.

(272) David, C. C.; Jacobs, D. J. Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol Biol* **2014**, *1084*, 193-226. DOI: 10.1007/978-1-62703-658-0_11.

(273) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* **2013**, *9* (7), 3084-3095. DOI: 10.1021/ct400341p.

(274) Bakan, A.; Meireles, L. M.; Bahar, I. ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* **2011**, *27* (11), 1575-1577. DOI: 10.1093/bioinformatics/btr168.

(275) Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* **2010**, *5* (4), 725-738. DOI: 10.1038/nprot.2010.5.

(276) Mostafavi, S.; Ray, D.; Warde-Farley, D.; Grouios, C.; Morris, Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol* **2008**, *9 Suppl 1*, S4. DOI: 10.1186/gb-2008-9-s1-s4.

(277) Montojo, J.; Zuberi, K.; Rodriguez, H.; Kazi, F.; Wright, G.; Donaldson, S. L.; Morris, Q.; Bader, G. D. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* **2010**, *26* (22), 2927-2928. DOI: 10.1093/bioinformatics/btq562.

(278) Warde-Farley, D.; Donaldson, S. L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C. T.; et al. The GeneMANIA prediction

server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* **2010**, *38* (Web Server issue), W214-220. DOI: 10.1093/nar/gkq537.

(279) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **2004**, *25* (13), 1605-1612. DOI: 10.1002/jcc.20084.