Hotspot Identification and Drug Design of Protein-Protein Interaction Modulators using the Fragment Molecular Orbital Method

Stefania Monteleone[1], Dmitri G. Fedorov[2], Andrea Townsend-Nicholson[3], Michelle Southey[1], Michael Bodkin[1] and Alexander Heifetz[1]*

[1]Evotec UK Ltd., 114 Innovation Drive, Milton Park, Abingdon, Oxfordshire OX14 4RZ, United Kingdom

[2]Research Center for Computational Design of Advanced Functional Materials (CD-FMat), National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan

[3]Institute of Structural & Molecular Biology, Research Department of Structural & Molecular Biology, Division of Biosciences, University College London, London, WC1E 6BT, United Kingdom

ABSTRACT: Protein-protein interactions (PPIs) are essential for the function of many proteins. Aberrant PPIs have the potential to lead to disease, making PPIs promising targets for drug discovery. There are over 64,000 PPIs in the human interactome reference database, however, to date very few PPI modulators have been approved for clinical use. Further development of PPI-specific therapeutics is highly dependent on the availability of structural data and existence of reliable computational tools to explore the interface between two interacting proteins. The fragment molecular orbital (FMO) quantum mechanics method offers a comprehensive and computationally inexpensive mean of identifying the strength (in kcal/mol) and the chemical nature (electrostatic or hydrophobic) of the molecular interactions taking place at the protein-protein interface. We have integrated FMO and PPI exploration (FMO-PPI) to identify the residues that are critical for protein-protein binding (hotspots). To validate this approach, we have applied FMO-PPI to a dataset of protein-protein complexes representing several different protein subfamilies, and obtained FMO-PPI results that are in agreement with published mutagenesis data. We observed that critical PPIs can be divided into 3 major categories: interactions between residues of two proteins (intermolecular), interactions between residues within the same protein (intramolecular) and interactions between residues of two proteins that are mediated by water molecules (water bridges). We extended our findings by demonstrating how this information obtained by FMO-PPI can be utilized to support the structure-based drug design of PPI modulators (SBDD-PPI).

INTRODUCTION

Protein-protein interactions underpin all of cell biology from fundamental processes that take place in every cell such as DNA replication, transcription and translation, to the control of dynamic networks such as cell signaling and immune responses. Many proteins have functions that depend on their ability to recognize and bind to other molecules including proteins. Key cellular processes are often regulated through the formation of protein complexes. These protein complexes are typically controlled via protein-protein interactions (PPIs).[1]

PPIs generate a complex network, called the "interactome" [1, 2], which plays a crucial role in physiological processes, such as signal transduction, cell proliferation, growth, differentiation and apoptosis.[3] Deviations in PPIs affect the entire network of protein-protein signaling and can lead to human pathophysiological conditions, such as cancers, neurodegenerative disorders and infectious diseases. [4, 5] Estimates of the number of PPIs in the human interactome range from 130,000 to 650,000 [3] and over 64,000 [6] confirmed PPIs are listed in the human interactome reference database. In recent years, the increasing attention that PPIs have received has made them a promising target for drug discovery [3]. PPI modulators have been developed to assist in treatment of leukemia, lymphoma, carcinoma, melanoma, lung cancer, ulcerative colitis, liver cirrhosis, kidney transplantation and other diseases and disorders of human health.[3]

However, despite these promising developments, PPIs remain extremely challenging drug targets because the traditional small molecule drug discovery approaches focus primarily on protein targets that have a relatively well-defined ligand-binding site that small molecules can interact with. PPIs have a considerable number of interactions that form through a much larger contact area, which is difficult for a small molecule to compete with. [7] Fortunately, despite

the relatively large size of protein-protein interfaces and the significant number of interactions, the vast majority of PPIs are weak and the stability of the complex is dependent on a relatively small number of strong PPIs formed by a limited number of residues [3]. Mutation of even one of these residues can abolish the formation of the complex [8] or significantly weaken its stability. Therefore, although the interface of PPIs is relatively large, a small molecule only needs to act on a few of these key residues to intervene in the PPI.

These key residues are named "hotspots". Tryptophan, arginine, and tyrosine are more frequently found to appear as hotspots than other amino acids [3] and, as a result, are frequently targeted in the structure-based drug design (SBDD) of PPI modulators (SBDD-PPI) [3, 9, 10]. Engagement of these hotspot residues in the interactions with small molecules could potentially prevent the formation of the protein-protein complex. This strategy would be particularly efficient if the hotspot is located at or near the cavity on the protein-protein interface where the ligand can bind. For this reason, the identification of hotspot residues is vital for SBDD-PPI of PPI modulators.

Traditionally the identification of hotspots is performed by site-directed mutagenesis (SDM). Unfortunately, SDM experiments such as alanine-scanning [11] are time consuming, expensive and not always appropriate for every protein or protein complex. For these reasons SBDD-PPI campaigns are likely to be impractical if SDM is the only method used for hotspot identification. Several computational approaches have been proposed in the past for the identification of hotspots (e.g., SILCS [12], FTMap [13] and Allosteer [14, 15]). These methods are based on molecular mechanics instead of QM and use protein dynamics and/or fragment docking to identify potential binding pockets for SBDD-PPI. The definition of hotspots in these methods differs from ours, with the author's defined hotspots as fragment binding sites, whereas we define hotspots as the single amino acids crucial for protein-protein binding. In this work, we propose a new computational approach as a viable, if not preferable, alternative.

Further development of PPI-focused drugs is highly dependent on the availability of structural data for the target complex and on the existence of accurate computational tools with which to analyse the structural data. There is increasing evidence [16] that a number of underappreciated interactions, such as CH/π [17], halogen/π [18] cation/π [19], and non-classical hydrogen bonds [20], play important roles in protein-ligand binding that are not adequately parameterized in the most popular force fields [21, 22] implemented in molecular mechanics (MM) calculations. Furthermore, reliable MM-based predictive methods for quantifying hydrophobic interactions, which are vital for stabilizing protein-protein interactions [23], remain to be developed [16].

Quantum mechanical methods (QM) have always been considered to be a reliable approach for the exploration of molecular interactions [24, 25]. However, despite their many advantages, traditional QM approaches are generally not feasible for large biological systems like proteins, due to their high computational cost.

The fragment molecular orbital (FMO) is a well-established QM method [26] that provides a list of interactions formed between protein residues including their strength (in kcal/mol) and their chemical nature (electrostatic or hydrophobic). FMO [25] offers a considerable computational speed-up over traditional QM methods. It is achieved by dividing the system into smaller pieces called fragments (*Supporting Information* Figure S1). For example, in proteins, each residue can be represented by a fragment. Ligand can also be represented by one or multiple fragments. By using fragments, one can perform QM calculations in much shorter time.

An additional advantage of FMO is the pair interaction energy decomposition analysis (PIEDA), which quantitatively decomposes the interaction energy between two residues (fragments) into the single energy terms that define it: electrostatic, exchange-repulsion, charge transfer, dispersion (*Supporting Information* Figure S1), and solvation [27]. Polar interactions are generally given by electrostatic and charge transfer terms, whereas hydrophobic interactions by the dispersion term. In this way, the chemical nature of the interaction (electrostatic or hydrophobic) can be identified and quantified. [28]

Many research groups have used FMO to identify underappreciated interactions in protein-small molecule and protein-protein binding. For example, the FMO protocol has been extensively used for exploration of interactions between COVID-19 main protease and its inhibitors [29, 30], class A G protein-coupled receptors (GPCRs) and their ligands [31], in discovery of ITK (kinase) and of Hsp90 [32] inhibitors and in many other SBDD programs [26]. In recent years, the use of FMO has been expanded to structural analysis of proteins [26, 33], for example in the exploration of interactions between transmembrane helices of GPCRs [34, 35], SARS-Cov-2-related proteins [36, 37] and several other PPI targets [33, 36, 38-40].

In the current study, we have used FMO to identify PPI hotspots (FMO-PPI), applying FMO to 6 protein-protein complexes (Table 1) that represent different protein subfamilies. As a benchmark, we compared our computational results to published experimental SDM data. As backbone-backbone interactions are not impacted in SDM experiments, we focused this research on PPIs between sidechain-sidechain, sidechain-backbone and PPIs mediated by water molecules (water bridges). QM-based methods are known to be sensitive to even small structural changes. We have previously and extensively reported this phenomenon for GPCRs.[41] For this reason, we selected the crystal structures with the highest possible resolutions from the PDB for these tested systems.

We divided hotspot residues into 3 three categories: (a) residues involved in interactions between two proteins (intermolecular PPIs), (b) residues that form interactions with the residues within the same protein (intramolecular PPIs) and (c) residues involved in interactions with interface

water molecules (water bridges) and explored the importance of each hotspot category to protein-protein complex stability. Since the interaction interface in a PPI can cover a relatively large and mobile surface area [42], we assumed that intramolecular PPIs would exert an indirect effect on complex formation by stabilizing the 'bioactive' conformation (topology) of the interface and preparing it for complex formation. Interactions with interface water molecules were also included in the FMO-PPI calculations, since water molecules can mediate intermolecular interactions by forming water bridges between residues at the interface [43].

Finally, we demonstrate how our findings can be used to guide SBDD-PPI of novel and potent PPI modulators.

Table 1: A summary of crystal structures that have been analyzed with FMO-PPI *.

Class	Protein A	Protein B	Year	Res. (Å)	PDB-ID
β-lactamase	TEM1	BLIP	2001	1.73	1JTG [44]
Interferon	IFNa2	IFNAR2	2011	2.00	3S9D [45]
GPCR	PTH1R	ePTH	2018	2.50	6FJ3 [46]
Kinase	LIMK1	Cofilin-1	2016	3.50	5HVK [47]
GTPase	KRAS	SOS1	2019	2.55	6EPL [48] (Apo structure), 6EPM [48] (Compound **1**), 5OVE [48] (Compound **2**), 5OVG [48] (Compound **3**), 5OVI [48] (Compound **4**)
E3-ligase	CRBN	CK1α	2015	2.45	5FQD[49]

*The bold line in the middle of the table separates the 4 cases used for SDM comparison and the 2 cases used for drug discovery.

RESULTS

Detecting hotspots with FMO-PPI

In this first stage of our work, we wanted to demonstrate that hotspots detected by FMO-PPI agree with those detected in SDM experiments. We selected 4 protein-protein complexes extracted from the Protein Data Bank (PDB) that have available experimental SDM data (Table 1). These systems represent different protein classes. Based on previous FMO reports [50, 51], we considered any interaction with a differential pair interaction energy (ΔPIE, see *Methods section*) ≥ 3.0 kcal/mol to be significant and treated the residues involved in this interaction as hotspots. Results have been displayed as heat maps including intra- and inter-molecular interactions, and water bridges (see *Supporting Information* Figure S2) for both ΔPIE and *f* (ratio of electrostatics and its sum with the dispersion term) terms.

Class β-lactamase: TEM1 in complex with BLIP

TEM1 hydrolyzes the β-lactam bond in antibiotics, thus leading to resistance to antibiotics such as penicillin. β-lactamase inhibitors such as BLIP are usually used together with antibiotics to prevent this [52]. The TEM1–BLIP complex (Figure 1A) was subjected to FMO-PPI analysis.

FMO-PPI results (Table 2, heat map plot Figure 1B and PIEDA results *Supporting Information* Figure S3A) are consistent with the reported SDM data (*Supporting Information* Table S1). FMO-PPI detected 36 TEM1 and 33 BLIP hotspots on their interface (Table 2). These hotspots were mapped on the surfaces of the two proteins and a potential ligand-binding pocket was found by Site Finder (see Methods section) at the TEM1-BLIP interface (Figure 1C). These potential ligand-binding pockets together with the hotspot information can facilitate SBDD-PPI of new TEM1-BLIP PPI modulators.

SDM of $K74^{BLIP}$ into alanine showed the highest decrease in TEM1-BLIP binding. According to FMO-PPI, $K74^{BLIP}$ forms two intermolecular PPIs with $E104^{TEM1}$ (salt bridge) and $Y105^{TEM1}$ (hydrophobic), and two hydrophobic intramolecular PPIs with $Y143^{BLIP}$ and $G141^{BLIP}$ (Figure 1D). This evidence explains why $K74^{BLIP}$ is so critical for TEM1-BLIP binding. Further, the FMO-PPI analysis indicates that $F142^{BLIP}$ forms a strong π−π stacking interaction with $Y105^{TEM1}$ and π-amide stacking with $N170^{TEM1}$. Moreover, the sidechain of $Y105^{TEM1}$ is involved in additional hydrophobic contacts with $K74^{BLIP}$. SDM of $R243A^{TEM1}$ and $D49A^{BLIP}$ result in loss of TEM1-BLIP formation. This can be explained by the fact that $R243^{TEM1}$ and $D49^{BLIP}$ form an ionic interaction and a water bridge via HOH687. The residues reported to have no effect on TEM1-BLIP binding in SDM experiments (*Supporting Information* Table S1) were also not identified as hotspots by FMO-PPI protocol.

Class Interferon: IFNα2 in complex with IFNAR2

Interferons are cytokines that play an important role in the autoimmune response. They are exploited in the treatment of multiple sclerosis and in different kinds of cancer [53]. Their signaling is mediated by the IFNα2 - IFNAR2 receptor complex (Figure 2A).

We used FMO-PPI to explore PPIs on the IFNα2 - IFNAR2 interface. FMO-PPI results (Table 2, heat map plot Figure 2B and PIEDA results *Supporting Information* Figure S4) were in the agreement with the reported SDM data (*Supporting Information* Table S2).

It was reported that R33A$^{IFN\alpha2}$ had the largest detrimental effect on formation of IFNα2-FNAR2 complex among all the mutated residues [45]. According to FMO-PPI, R33$^{IFN\alpha2}$ forms 5 intermolecular PPIs with residues of IFNAR2: T44, I45, M46, K48 and E50 (Figure 2C).

Table 2: FMO-PPI results and published SDM (see *Supporting Information* Tables S1-S4).

System: Protein A – Protein B	Number of hotspots detected by SDM: Proteins A/B	Number of hotspots detected by FMO-PPI: Proteins A/B (as in SDM)	Number of hotspots from FMO-PPI involved in intermolecular PPI: Proteins A/B	Number of hotspots from FMO-PPI involved in intramolecular PPI: Proteins A/B	Number of hotspots from FMO-PPI involved in water bridges: Proteins A/B
TEM1-BLIP	5/4	36/33 (5/4)	26/21	20/19	12/10
IFNα2-IFNAR2	12/n.a.	22/13 (12/0)	10/11	19/9	9/6
PTH1R-ePTH	7/n.a.	47/31 (7/0)	27/16	37/*	4/3
LIMK1-Cofilin-1	3/3	20/19 (3/3)	16/11	15/17	n.d.

n.a.: not annotated in literature. n.d.: the crystal structure did not contain any water molecules. *ePTH is a polypeptide and its secondary structure is a simple α-helix, so its intramolecular PPIs were not considered.

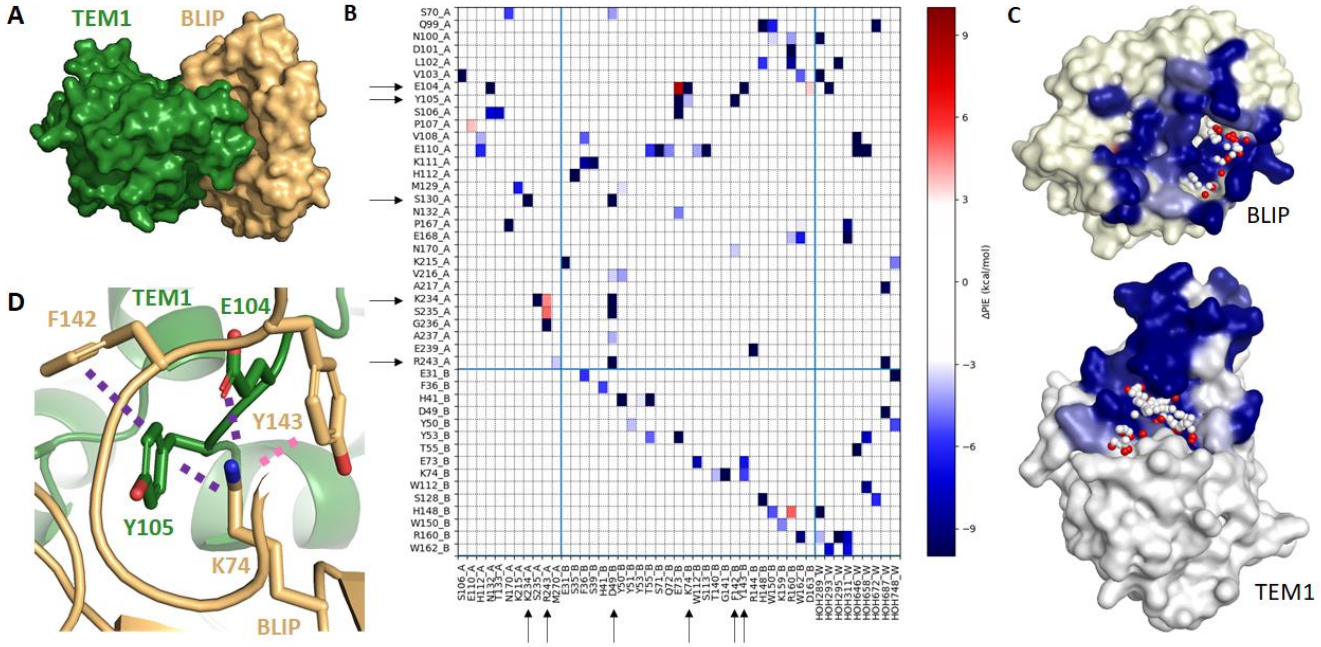


Figure 1. FMO-PPI results for the crystal structure of TEM1-BLIP (PDB entry 1JTG) [44]: (**A**) TEM1 (surface colored in green) in complex with BLIP (surface colored in light-orange) (**B**) Heat map plot representing PPIs detected by FMO-PPI. Each box represents one interaction. Boxes are colored by spectrum according to their ΔPIE values (in kcal/mol) from dark blue (ΔPIE ≤ -10 kcal/mol, strong attraction) to red (ΔPIE ≥ 10 kcal/mol, strong repulsion). PPIs in the range of -3 kcal/mol < ΔPIE < 3 kcal/mol are represented by a white box. Arrows indicate residues that were also detected as hotspots by SDM. (**C**) Mapping of the hotspots on the surfaces of the disassembled proteins. BLIP (on the top) was rotated by 180° on the horizontal axis and translated to expose the interfaces of the two proteins. Each interface residue is colored by its highest ΔPIE value and according to color scheme in panel B. Surface of residues with PIE values in the range of -3 kcal/mol < ΔPIE < 3 kcal/mol or none interface residues are colored in white for TEM1 and light-yellow for BLIP. Potential ligand-binding pockets were identified by Site Finder (see *Methods* section)

and marked by spheres (white spheres represents hydrophobic area and red polar). (**D**) PPIs formed by $K74^{BLIP}$ with surrounding residues. The intramolecular and intermolecular PPIs are shown as pink and purple dashed lines, respectively.

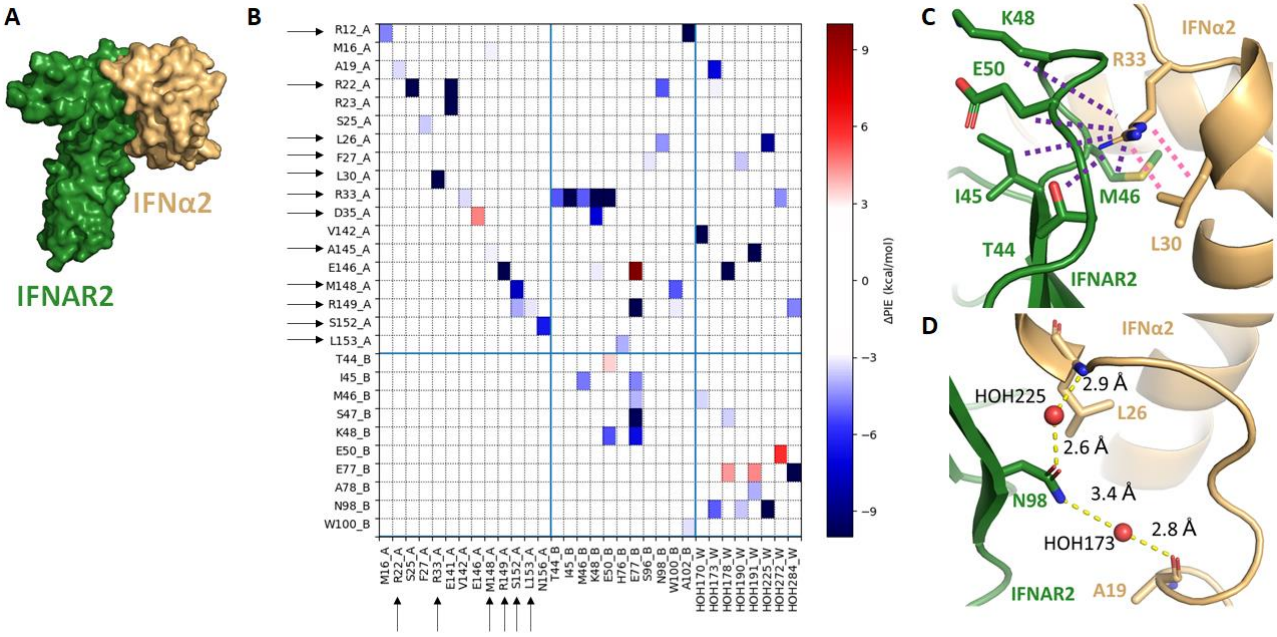


Figure 2. FMO-PPI results for crystal structure of IFNα2 - IFNAR2 (PDB entry 3S9D)[45]: (**A**) IFNAR2 (surface colored in green) in complex with IFNα2 (surface colored in light orange) (**B**) Heat map plot representing PPIs detected by FMO-PPI. Each box represents one interaction. Boxes are colored by spectrum according to their ΔPIE values (in kcal/mol) from dark blue (ΔPIE ≤ -10 kcal/mol, strong attraction) to red (ΔPIE ≥ 10 kcal/mol, strong repulsion). PPIs in the range of -3 kcal/mol < PIE < 3 kcal/mol are represented by a white box. Arrows indicate residues that were also detected as hotspots by SDM. (**C**) PPIs formed by $R33^{IFNα2}$ with surrounding residues. The intramolecular PPIs are shown as pink dashed lines and the intermolecular ones as purple dashed lines. (**D**) Water bridges identified by FMO-PPI, which mediate PPIs between residues of IFNAR2 and IFNα2, are displayed as yellow dashed lines.

SDM of $L30A^{IFNα2}$ has also a dramatic effect on the IFNα2-IFNAR2 affinity. FMO-PPI analysis showed that $L30^{IFNα2}$ forms strong electrostatic (via the backbone) and hydrophobic (via the sidechain) intramolecular PPI with $R33^{IFNα2}$ (Figure 2C and *Supporting Information* Figure S4A), suggesting that the major role of $L30A^{IFNα2}$ is to stabilize the bioactive conformation of $R33^{IFNα2}$ and, by doing so, indirectly affects the IFNα2 - IFNAR2 binding affinity.

Both FMO-PPI and SDM highlighted the importance of $M148^{IFNα2}$ and $A145^{IFNα2}$ for complex formation. The $M148^{IFNα2}$ forms a hydrophobic, intermolecular PPI with $W100^{IFNAR2}$ and an intramolecular PPI with $A145^{IFNα2}$, suggesting that $A145^{IFNα2}$ works as stabilizer of bioactive conformation of $M148^{IFNα2}$ and, in doing so, indirectly affects complex stability.

Water molecules also play a key role in stabilizing this complex. For example, $N98^{IFNAR2}$ forms two water bridges with HOH173 and HOH225. These water bridges mediate intramolecular PPIs formed between $N98^{IFNAR2}$ with various residues: $L26^{IFNα2}$, that was identified as hotspot by both FMO-PPI and SDM, and $A19^{IFNα2}$ (Figure 2D and *Supporting Information* Table S2). [45]

The residues reported to have no effect on IFNα2-IFNAR2 binding in SDM experiments (*Supporting Information* Table S2) were also not identified as hotspots by FMO-PPI protocol.

Class GPCR: PTH1R in complex ePTH

G-protein coupled receptors (GPCRs) have enormous physiological and biomedical importance and are involved in a wide range of diseases. It is, therefore, not surprising that 475 drugs (~34% of all drugs approved by the US Food and Drug Administration (FDA)) act on this protein family [54]. Many biological functions of peptides are mediated through GPCRs. However, the design of peptide drugs targeting GPCRs remains challenging [55]. Further development of peptide drugs depends on availability of structural information and the understanding of the interactions formed between peptide and the GPCR receptor. FMO-PPI can be helpful tool for this purpose.

ePTH is a peptide which is an engineered version of the parathyroid hormone (PTH) that regulates calcium homeostasis and used to treat osteoporosis. [46].ePTH forms a complex with PTH1R (parathyroid hormone 1 GPCR recep-

tor) via its extracellular domain and the transmembrane helices (TM) (Figure 3A). We used FMO-PPI to explore PTH1R-ePTH complex and results (Table 2, heat map plot Figure 3B and PIEDA results *Supporting Information* Figure S5) in agreement with the reported experimental SDM data (*Supporting Information* Table S3).

SDM of F184^{PTH1R} to alanine resulted in the abolition of PTH1R-ePTH binding. According to FMO-PPI, F184^{PTH1R} forms face-to-edge π-stack (Figure 3C) intermolecular interaction with W14^{ePTH} (Figure 3B) and intramolecular PPI interactions with L187^{PTH1R}.

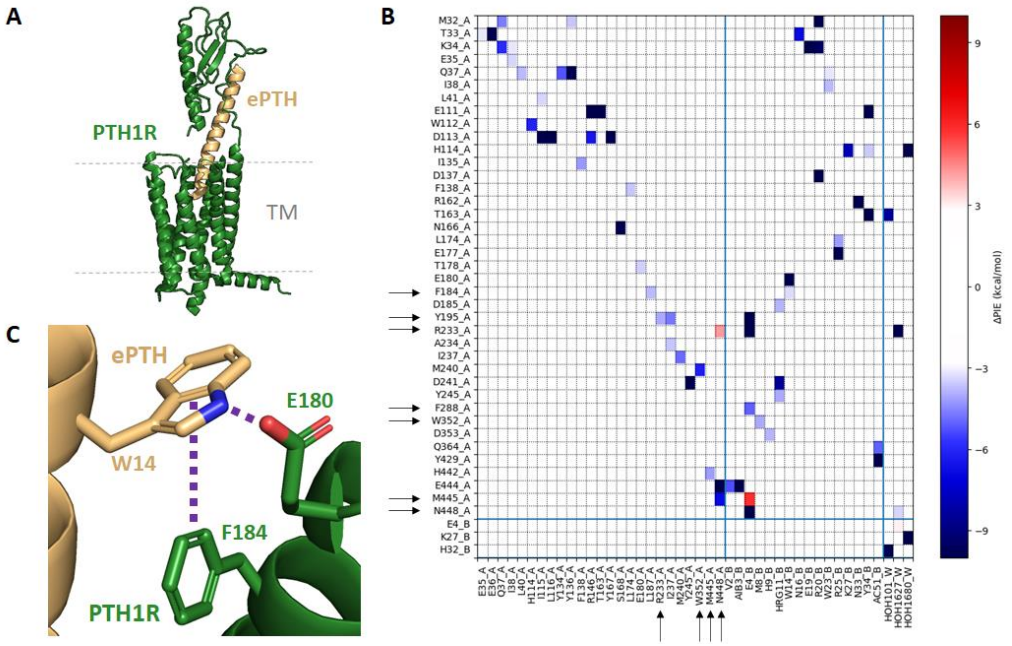


Figure 3. FMO-PPI results for crystal structure of PTH1R - ePTH (PDB entry 6FJ3)[46] (**A**) PTH1R (ribbon colored in green) in complex with peptide ePTH (ribbon colored in light orange) (**B**) Heat map plot representing PPIs detected by FMO-PPI. Each box represents one interaction. Boxes are colored by spectrum according to their ΔPIE values (in kcal/mol) from dark blue (ΔPIE ≤ -10 kcal/mol, strong attraction) to red (ΔPIE ≥ 10 kcal/mol, strong repulsion). PPIs in the range of -3 kcal/mol < ΔPIE < 3 kcal/mol are represented by a white box. Arrows indicate residues that were also detected as hotspots by SDM. (**C**) Intermolecular PPIs, formed by W14 ^{ePTH} with surrounding residues of PTH1R shown as purple dashed lines.

Moreover, alanine mutations of Y195^{PTH1R}, R233^{PTH1R}, F288^{PTH1R} or N448^{PTH1R} resulted in a statistically significant decrease in measured IC_{50}. According to FMO-PPI, this is due to the loss of their intermolecular interactions with E4^{ePTH}. Y195^{PTH1R} is also involved in two intramolecular interactions with R233^{PTH1R} (polar) and Ile237^{PTH1R} (hydrophobic), and E180^{PTH1R} forms a hydrogen bond with W14^{ePTH} (Figure 3C).

L244^{PTH1R} and W352^{PTH1R}, which is in the extracellular loop 2 (ECL2), form hydrophobic interactions with each other and with M8^{ePTH}, while M445^{PTH1R} is involved in weak hydrophobic interactions with Aib3^{ePTH} (below -3 kcal/mol) and an electrostatic repulsion with E4^{ePTH}.

Further, SDM experiments have shown that mutation to Ala of L232^{PTH1R} or V235^{PTH1R} did not significantly affect peptide binding [56]. These experimental evidences are in agreement with FMO-PPI results that correctly predicted that these residues are not hotspots.

Class Kinase: LIMK1 in complex with Cofilin-1

Kinases are one of the major drug targets [57], regulating almost every cellular process by switching on and off other proteins via their phosphorylation. Due to their crucial role, kinases are often involved in several kinds of cancers and thus, their inhibitors are usually used in chemotherapy. [58] Kinase substrate recognition is based on protein-protein interactions. LIM domain kinase 1 (LIMK1) is a potential drug target for the prevention of amyotrophic lateral sclerosis [59]. LIMK1 plays a crucial role in the regulation of actin

dynamics, by binding to (and phosphorylating) the actin depolymerizing factor cofilin-1 (Figure 4A) [47].

We used FMO-PPI to explore LIMK1-Cofilin-1 complex (Table 2, heat map plot Figure 4B and PIEDA results *Supporting Information* Figure S6) and results agree with reported experimental SDM data (*Supporting Information* Table S4). SDM highlighted several important residues that affect LIMK1-Cofilin-1 affinity and/or function. For instance, SDM of $M115^{Cofilin\text{-}1}$ to alanine resulted in loss of Cofilin-1 phosphorylation. According to FMO-PPI, $M115^{Cofilin\text{-}1}$ forms two hydrophobic intermolecular PPIs with $Y514^{LIMK1}$ and $F559^{LIMK1}$. Double mutation $D549K^{LIMK1}$ and $D551K^{LIMK1}$ lead to the same effect which can be explained by the strong ionic intermolecular PPI between $K112^{Cofilin\text{-}1}$ and $D549^{LIMK1}$ and $D551^{LIMK1}$. FMO-PPI also highlighted importance of the underappreciated (see Introduction section) interactions (Figure 4C): $S119^{Cofilin\text{-}1}$ forms two CH-π interactions with $Y514^{LIMK1}$ and with $F559^{LIMK1}$ and $S119^{Cofilin\text{-}1}$ forms OH—S interaction with $M516^{LIMK1}$.

Moreover, protein-protein interactions play a key role not only in stabilizing the complex, but also in correctly positioning the substrate in the catalytic site. Indeed, Cofilin-1's α-helix 5 (where residues K112, M115 and S119 are located) guides the amino acid $S3^{Cofilin\text{-}1}$ towards the correct orientation for the phosphorylation mechanism. Phosphorylation of $S3^{Cofilin\text{-}1}$ is essential for the regulation of the interaction between Cofilin and actin. FMO-PPI analysis identified interactions between the phosphorylated serine and the kinase residues of the catalytic site. For instance, the phosphorylated $S3^{Cofilin\text{-}1}$ forms an electrostatic repulsion with $D478^{LIMK1}$, which is part of the conserved DFG motif.

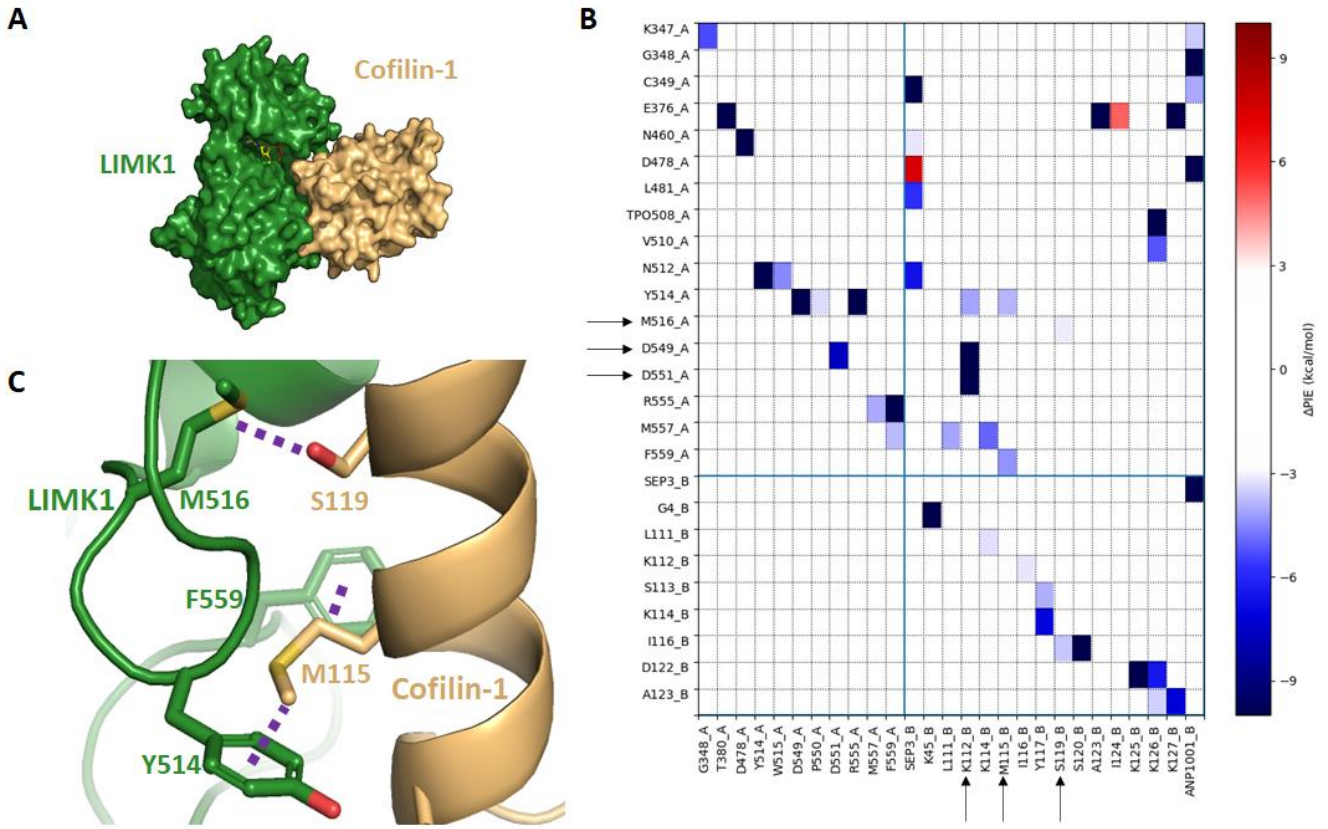


Figure 4. FMO-PPI results for crystal structure of LIMK1 - Cofilin-1 (PDB entry 5HVK) [47]: (**A**) LINK1 (surface colored in green) in complex with Cofilin-1 (surface colored in light orange) (**B**) Heat map plot representing PPIs detected by FMO-PPI. Each box represents one interaction. Boxes are colored by spectrum according to their ΔPIE values (in kcal/mol) from dark blue (ΔPIE ≤ -10 kcal/mol, strong attraction) to red (ΔPIE ≥ 10 kcal/mol, strong repulsion). PPIs in the range of -3 kcal/mol < ΔPIE < 3 kcal/mol are represented by a white box. Arrows indicate residues that were also detected as hotspots by SDM. (**C**) Intermolecular PPIs formed between hotspots of LINK1 (ribbon and carbons of residues colored in green) and Cofilin-1 (ribbon and carbons of residues colored in orange) shown as purple dashed lines.

FMO-PPI guided SBDD

In the second part of our work, we wanted to demonstrate how the information obtained by FMO-PPI can be used for SBDD of PPI modulators. We suggest here a simple and practical approach (Figure 5) to achieve this goal. The first step in this protocol is to define the protein-protein interface, followed by FMO-PPI analysis and independent Site Finder search as implemented in MOE [60]. These two parallel steps are required for determining hotspots and potential ligand binding sites (pockets).

The ideal PPI-ligand binding site should contain at least one hotspot. A ligand that would bind to this site and interact with hotspot residues will have a higher chance of affecting protein-protein binding. This information can be efficiently applied in VS (virtual screening) to search for the new ligands that can target these interface sites and/or in advanced SBDD phases, such as hit-to-lead or lead optimization. To demonstrate how this approach works in a real-world example, we retrospectively applied it in two SBDD drug discovery cases.

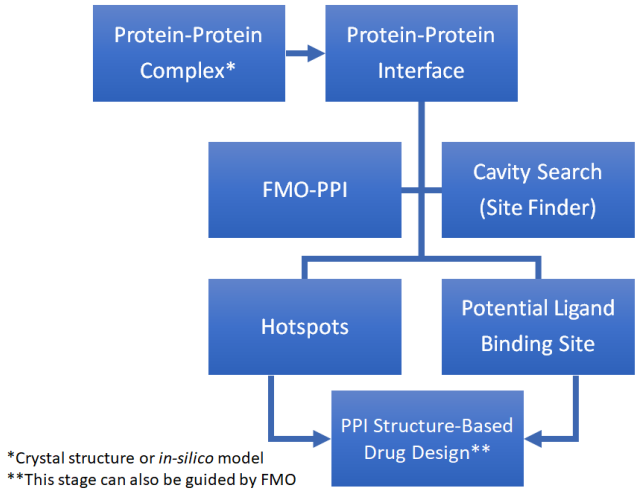


Figure 5. Workflow for structure-based drug design of PPI modulators (SBDD-PPI).

Case study 1: Design of compound 4 (BAY-293) as inhibitor of KRAS-SOS1 interaction

KRAS is a GTPase that is activated by SOS1 (Son of Sevenless 1, the most-studied guanine nucleotide exchange factor [61]). KRAS-SOS1 complex is frequently involved in various cancers. Inhibition of its formation by a small drug-like molecule is an attractive strategy for anticancer treatments. Compound **4** (BAY-293, Figure 6C) is a new KRAS-SOS1 inhibitor.[48] Here we retrospectively demonstrate how compound **4** could have been designed if FMO-PPI approach as descried in Figure 5 were applied.

We started with the FMO-PPI analysis of KRAS-SOS1 interface (PDB code 6EPL), which identified numerous hotspots (Figure 6A). We mapped these hotspots on the surfaces of the two proteins (Figure 6B). In parallel, we used Site Finder [60] to explore the interface of KRAS and SOS1 and search for potential ligand-binding pockets. We found multiple pockets and one of these pockets on the interface of SOS1 (site 1, marked in red square in Figure 6B) was particularly interesting as it contained 3 hotspots $V875^{SOS1}$, $N879^{SOS1}$ and $Y884^{SOS1}$.

Hillig and coworkers [48] reported that two independent and parallel high-throughput screenings (HTS) against SOS1 lead to the identification of compounds **1** (Figure 6C, EC_{50} in the µM range) and **2** (Figure 6D, IC_{50} = 320 nM) [48]. Crystal structures of compounds **1** and **2** with SOS1 (Table 1) confirm that both these compounds bind site 1. We analyzed these structures with FMO and found 5 interactions between compound **1** and SOS1, including with the hotspot residue Y884 (Figure 6D). Regarding compound **2**, FMO detected 9 interactions with SOS1, including with 3 hotspot residues ($V875^{SOS1}$, $N879^{SOS1}$ and $Y884^{SOS1}$) and 2 via water molecules (Figure 6D). Engagement of these 3 hotspots of SOS1 in protein-ligand interactions instead of in PPIs with KRAS can explain why compound **2** acts as inhibitor of KRAS-SOS1 complex.

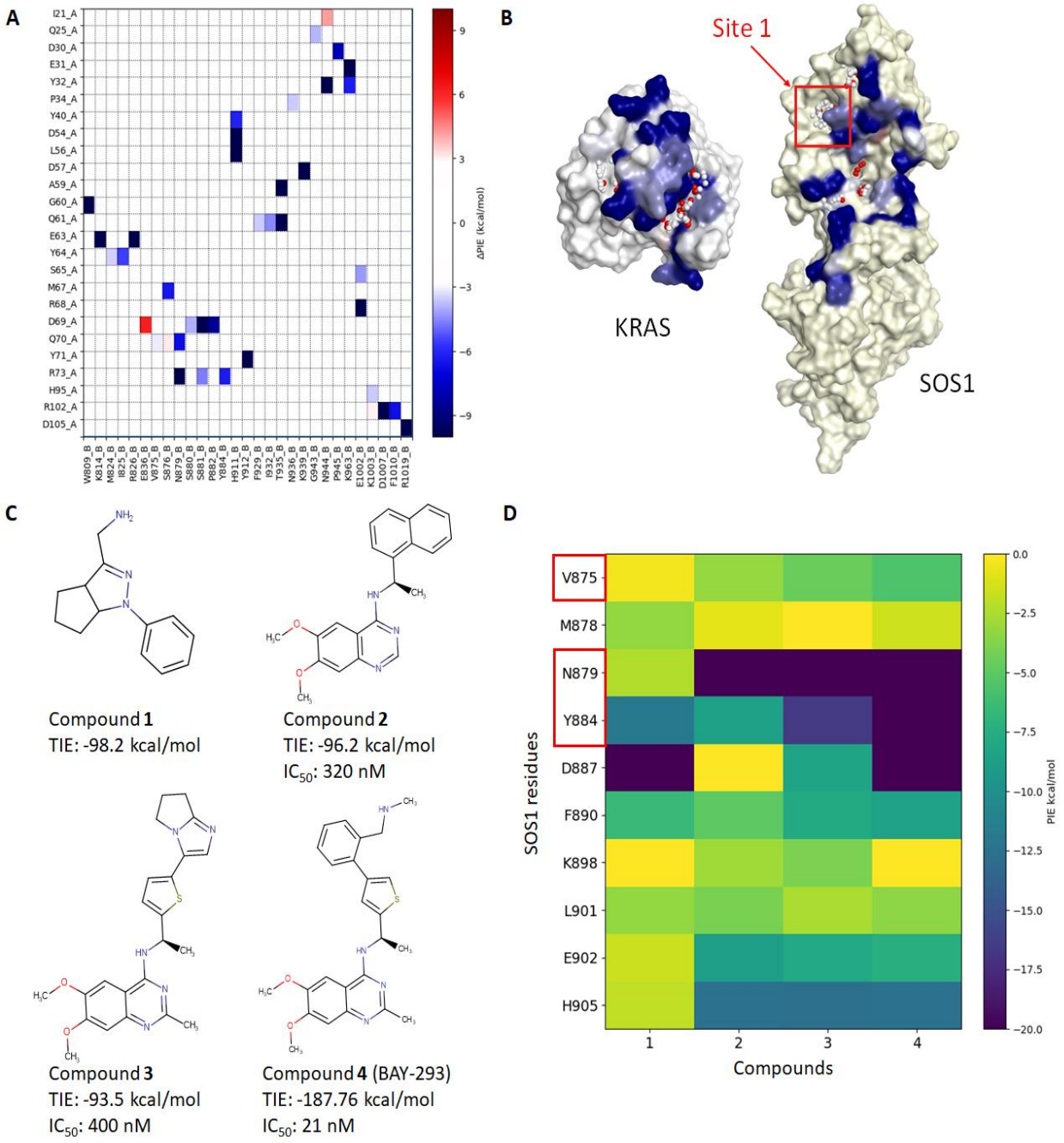
Compound **3** (Figure 6C) is a synthetic hybrid of compounds **1** and **2**. FMO analyses of the crystal structure of SOS1 with compound **3** detected 9 interactions (Figure 6D), including interaction with the hotspots V875, N879 and Y884. Compound **3** is a new chemical matter, however no improvement in IC_{50} was observed compared to compound **2**. This can be explained by the fact that compound **3** also forms just 9 interactions with SOS1 like compound **2** and no improvement in TIE (total interaction energy calculated by FMO) was observed.

After several SAR (structure-activity relationship) iterations conducted by Hillig and coworkers [48], compound **4** was designed (BAY-293, IC_{50} = 21 nM, Figure 6E). FMO-analysis of the crystal structure of SOS1 with compound **4** (Table 1) detected 11 interactions (Figure 6D, compare to just 9 interactions of compounds **2** and **3**) including new salt bridge with $D887^{SOS1}$ (Figure 6E) and TIE was -187.76 kcal/mol (compared to -96.2 and -93.5 kcal/mol of compounds **2** and **3**). This FMO outcome can explain why compound **4** had 21-fold improvement in IC_{50} compared to compounds **2** and **3**. Compound **4** also binds site 1 of SOS1 and interacts with hotspot residues V875, N879 and Y884 (Figure 6E), preventing these residues from forming interactions with KRAS. These evidences explains why compound **4** is a potent inhibitor of KRAS-SOS1 complex formation [48]. SDM of $N879^{SOS1}$ to alanine resulted in loss of KRAS-SOS1 binding [48].

We were also able to identify additional pockets like site 2 (*Supporting Information* Figure S8B) that can be targeted to design PPI inhibitors or molecular glues (small molecules that stabilise protein-protein complexes). For instance, in another case study [62], a series of fragments was identified targeting site 2 on the SOS1 surface (*Supporting Information* Figure S8B). Therefore, these fragments could be used as starting point for the SBDD-PPI program.

With respect to the KRAS pockets identified by Site Finder together with FMO-PPI, there are many ligands reported in literature [63, 64] that can bind to them, affecting KRAS activity or the binding of KRAS to effector proteins (*Supporting Information* Figure S8C-D).

In this case, we demonstrated how FMO-PPI approach, as described in Figure 5, integrated with Site Finder and followed by FMO-based small molecule SBDD, can be a powerful tool in discovery and lead-optimization of novel PPI modulators.



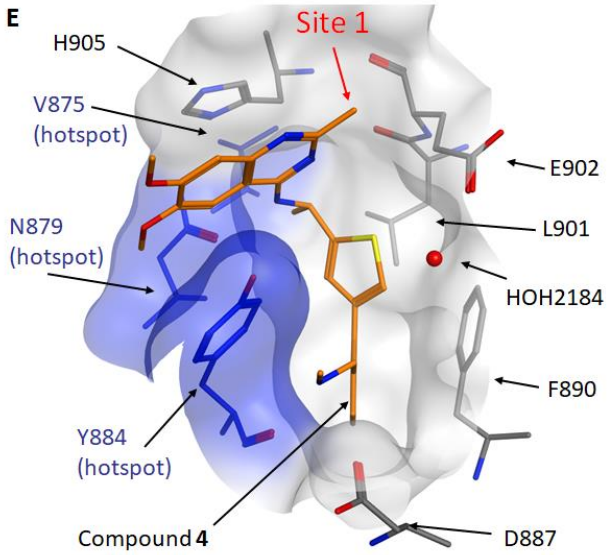


Figure 6. (**A**) Heat map representing intermolecular PPIs detected by FMO-PPI for SOS1 in complex with KRAS (PDB entry 6EPL) [48]. Boxes are colored by spectrum according to their ΔPIE values (in kcal/mol) from dark blue (ΔPIE ≤ -10 kcal/mol, strong attraction) to red (ΔPIE ≥ 10 kcal/mol, strong repulsion). PPIs in the range of -3 kcal/mol < ΔPIE < 3 kcal/mol are not shown. (**B**) Mapping of the hotspots on the surfaces of the disassembled proteins. SOS1 (left) was rotated by 180° on the vertical axis and translated to expose the interfaces of the two proteins. Each interface residue is colored by its highest ΔPIE value and according to the color scheme in panel **A**. Potential ligand-binding pockets were identified by Site Finder and marked by spheres (white spheres represent hydrophobic area and red spheres the polar one). (**C**) 2D structures of compounds **1-4**, experimental data extracted from literature, compared together with the TIE (total interaction energy) values that were calculated by FMO for each SOS1-ligand complex. (**D**) Heat map plot showing PIE values calculated by FMO for protein-ligand interactions (hotspots are marked in red boxes). (**E**) Compound **4** – SOS1 complex (PDB entry 5OVI) [48]. Hotspot residues are highlighted in blue and compound **4** is represented in orange sticks for carbon atoms.

Case study 2: Why is LVY (lenalidomide) critical for CK1α - CRBN binding?

CRBN is a part of CRL4 (Cullin-4-RING E3 ubiquitin ligase) complex. E3 ubiquitin ligases recognize their substrates through a short sequence of amino acids which are crucial for the protein-protein interaction. Binding of thalidomide and its derivatives to CRL4 induces the degradation of proteins of interests, such as CK1α kinase [65] and increases their efficacy against multiple myeloma cells.[66]

No direct binding of CK1α to CRBN has been experimentally detected in the absence of the small drug-like molecule LVY (lenalidomide) or its analogs [49]. The mechanism of action of LVY and its effect on CRBN-CK1α is not fully understood yet [49]. We used FMO-PPI to rationalize why the CRBN-CK1α complex cannot be formed without LVY.

We applied FMO-PPI to analyze the crystal structure of CRBN-LVY-CK1α complex (PDB code 5FQD [49], Figure 7A and 7B,). According to FMO-PPI, LVY forms 10 interactions (Figure 7B) with the CRBN-CK1α complex (9 + 1 respectively) and it is critical for CRBN and CK1α binding for the following reasons:

LVY stabilizes the bioactive conformation of CK1α by forming 1 critical interaction with $G40^{CK1\alpha}$. The $G40^{CK1\alpha}$ is located on the pinhead of the conserved loop of CK1α (loop 1; β-hairpin loop between $L33^{CK1\alpha}$ and $V45^{CK1\alpha}$ of CK1α).

Loop 1 is critical for CRBN-CK1α binding because 5 of its residues; $I37^{CK1\alpha}$, $T38^{CK1\alpha}$, $N39^{CK1\alpha}$, $G40^{CK1\alpha}$ and $E41^{CK1\alpha}$ forming intermolecular PPIs with CRBN (Figure 7B). According to SDM [49], mutation of residues $I37^{CK1\alpha}$ and $N39^{CK1\alpha}$ (located on loop 1) into alanine, prevents CRBN-CK1α binding even in the presence of LVY [49]. FMO-PPI detected that $I37^{CK1\alpha}$ forms two intermolecular PPIs with $N351^{CRBN}$ and $H353^{CRBN}$. $N39^{CK1\alpha}$ forms two intermolecular PPIs with $H397^{CRBN}$ and $W400^{CRBN}$, and two intramolecular PPIs with other loop 1 residues ($N36^{CK1\alpha}$ and $E41^{CK1\alpha}$).

Due to the central position of $G40^{CK1\alpha}$ in loop 1, an interaction with this residue controls the conformation of the entire loop 1. According to FMO, LVY forms a CH-π interaction with $G40^{CK1\alpha}$ (Figures 7C and 7D) and this interaction stabilizes the bioactive conformation of loop 1, allowing its other residues to interact with CRBN. Mutation of $G40N^{CK1\alpha}$ abolishes CRBN-CK1α binding [49], because asparagine (instead of glycine) clashes with LVY. This demonstrates that a change in even one hotspot could have a dramatic effect on protein-protein binding.

LVY also stabilizes the bioactive conformation of CRBN interface by forming 9 interactions with its interface residues. It simultaneously interacts with both rigid side (β-sheets) and flexible side (loops) of the CRBN binding pocket

and, by doing so, rigidifies the bioactive conformation of CRBN interface (Figures 7C and 7D). FMO-PPI detected that LVY forms 3 interactions with β-sheets residues: W386^{CRBN}, W400^{CRBN} and F402^{CRBN} and 6 interactions with loops residues: N351^{CRBN}, P352^{CRBN}, H353^{CRBN} E377^{CRBN}, H378^{CRBN} and W380^{CRBN} (Figures 7C and 7D). These FMO-PPI results agree with SDM experiments[49] that outlined the crucial role of H353^{CRBN}. H353^{CRBN} forms strong interaction with the LVY and two intermolecular PPIs with K18$^{CK1\alpha}$ and I37$^{CK1\alpha}$.

We compared the crystal structures of CRBN with and without LVY and observed a large conformational change of the flexible side of the CRBN pocket (*Supporting Information* Figure S10A) as result of LVY binding. However, the binding of CK1α to CRBN-LVY complex did not affected CRBN structure and only light shift in the position of the phthalimide ring of LVY was observed (*Supporting Information* Figure S10B). These experimental evidences support the original FMO-PPI based hypothesis that LVY stabilizes the bioactive conformation of CRBN interface.

In summary, LVY is critical for CRBN-CK1α binding because it stabilizes the bioactive conformation of both proteins by simultaneously interacting with them and, thus, acting as molecular 'glue' (Figures 7C and 7D).

DISCUSSION

PPI-focused modulators have become promising drug discovery targets and the focus of increasing attention for the development of novel therapeutics. The design of such drugs is highly dependent on the availability of structural data and accuracy of computational tools. The use of experimental methods such as SDM to identify PPI hotspots is an expensive, laborious, and lengthy process that is not always feasible for every protein-protein complex. These experimental technologies have proven difficult to align with the typical timescales of drug discovery programs. Here, we present a new computational approach, FMO-PPI, which provides an alternative means of supporting SBDD-PPI in real times and obviates the need of running expensive and long SDMs.

In our work, we have demonstrated that FMO-PPI is able to detect all hotspot residues that have previously been reported in the published SDM experiments. The residues reported to have no effect on protein-protein binding in SDM experiments were also not identified as hotspots by FMO-PPI protocol. FMO-PPI was able to detect hotspots in addition to those reported in the published SDM literature. Unfortunately, due to the lack of experimental data for these residues, we were unable to determine whether these hotspots are true or false positives.

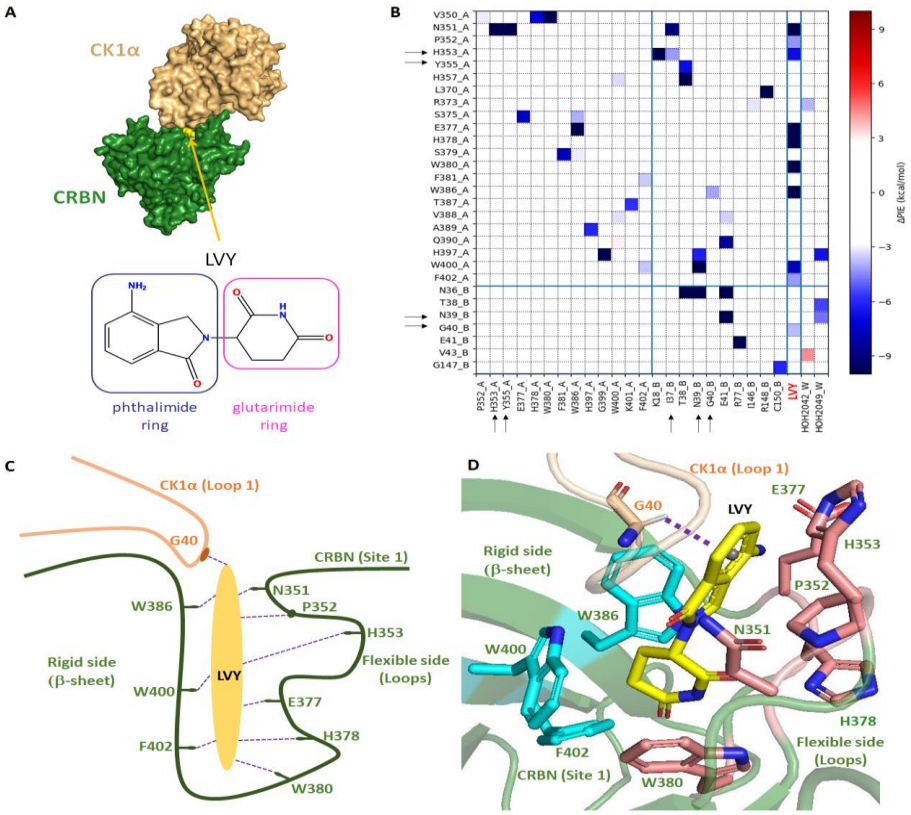


Figure 7. (**A**) Crystal structure of CRBN (surface colored in green) in complex with CK1α (surface colored in light orange) and LVY (lenalidomide, shown as yellow spheres) (PDB entry 5FQD) [49]. (**B**) Heat map representing PPIs detected by FMO-PPI. Each box represents one interaction. Boxes are colored by spectrum according to their ΔPIE values (in kcal/mol) from dark blue (ΔPIE

≤ -10 kcal/mol, strong attraction) to red (ΔPIE ≥ 10 kcal/mol, strong repulsion). PPIs in the range of -3 kcal/mol < ΔPIE < 3 kcal/mol are represented by a white box. The interactions of LVY with CRBN-CK1α are marked with red font. Arrows indicate residues that were also detected as hotspots by SDM. (**C**) Scheme description of the CRBN-CK1α-LVY binding mode with the key interactions detected by FMO. (**D**) CRBN-CK1α-LVY binding mode. Only residues detected by FMO as important for interactions with LVY (yellow sticks) are shown. CK1α residues are colored in orange. CRBN residues located on loops are colored in pink, on β-sheet as cyan, and the key interaction with $G40^{CK1\alpha}$ is shown as purple dashed line.

FMO-PPI analysis provides four major categories of information about the PPIs formed between any two interacting proteins: (a) the list of PPIs that each interface residue forms with its neighboring amino acids, (b) how strong these PPIs are in kcal/mol, (c) their chemical nature (electrostatics or hydrophobic) and (d) their PPIs category e.g.: inter-, intra- molecular or water-bridges. Comparable advantages were also reported by Lim H et al.[67,79] with their method 3D-SPIEs.

FMO/PIEDA calculations across all 6 protein-protein complexes (Figure 8A) reveals the importance of hydrophobic (dispersion) contributions to protein-protein affinity. Our results have shown that, on average, approximately 20% of PPIs have mainly hydrophobic nature (Figure 8B). We were also able to detect other underappreciated PPIs, including non-classical hydrogen bonds, CH-π and OH-S, highlighting the importance of these in the establishment and maintenance of the interaction between two protein partners.

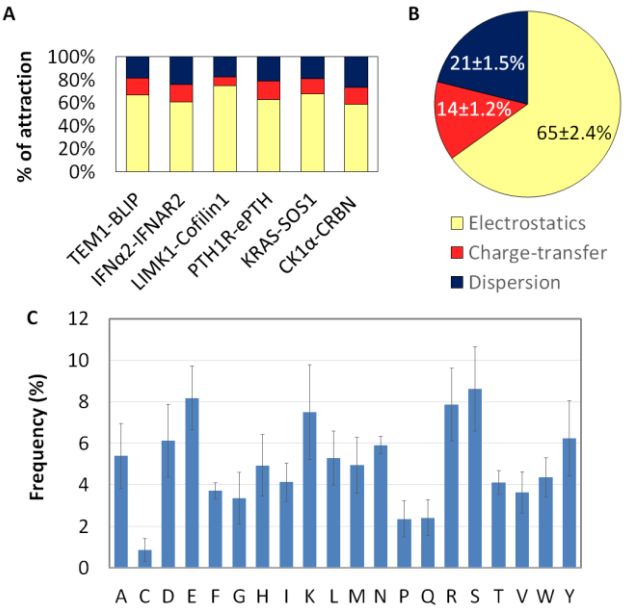


Figure 8. Relative percentages of electrostatics (yellow), charge transfer [51] and dispersion (blue) energy terms are displayed for each system (**A**) and as averaged values across all the systems under consideration (**B**). The sum of these terms constitutes the so-called "attraction energy". (**C**) Average frequency of each amino acid at the PPI as detected by FMO-PPI across all systems. Error bars represent the standard error of mean.

Primarily FMO has been applied to the exploration of hotspots, with particular focus on the intermolecular interactions between proteins, taking into account the distance between the residues and their interaction energies, such as calculated with 3D-SPIEs method [67, 68]. However as research progressed, FMO-PPI has highlighted not only the intermolecular interactions that play a key role in protein-protein binding but also the intramolecular PPIs and water bridges. The same observation has previously been reported by Anzaki S. et. al [69] and Okuwaki K. et al. [70]

On average, approximately 50% of all PPIs are intramolecular. Interestingly, most of the hotspots identified by SDM and FMO are simultaneously involved in more than one category of PPI (intermolecular, intramolecular and/or water bridges, Table 2). These hotspots play a dual role by directly and indirectly stabilizing the protein-protein complex. They directly interact with a hotspot of the partner protein whilst simultaneously stabilizing the bioactive conformation of neighboring hotspot by forming intramolecular interactions with it. This link between intramolecular and intermolecular PPIs may also explain why two proteins strive to bind each other: as the formation of the protein-protein complex progresses, there is an increase in the number of intra- and intermolecular PPIs and, thus, potentially an increase in the stability of the secondary structure of each individual protein.

The significance of intramolecular PPIs for protein-protein binding as highlighted by FMO-PPI were also validated experimentally. For example, according to FMO-PPI, $L30^{IFN\alpha2}$ (in the above study), does not engage in intermolecular PPIs with any residue of IFNAR2 but forms just two intramolecular PPIs with $R33^{IFN\alpha2}$ (Figure2C). Mutation of $L30^{IFN\alpha2}$ into alanine resulted in the elimination of IFNα2 and IFNAR2 binding. These insights suggest that some of the intramolecular PPIs in a protein-protein complex ($L30^{IFN\alpha2}$ in the example provided here) solely exert an indirect effect on protein-protein affinity, by stabilizing the bioactive conformation of a neighboring hotspot ($R33^{IFN\alpha2}$).

FMO-PPI has also highlighted the important role of water molecules in enhancing the stability of the protein-protein complex by forming water-bridges between hotspots of two proteins, increasing the number of intermolecular PPIs between them.

It has been shown that mutation of a single amino acid can trigger a cascade of structural changes that affects the entire PPI network [8]. FMO-PPI provides a detailed molecular explanation for this. FMO-PPI is always applied to current (static) snapshot of the system. If the structure of the system has changed, for example due to a mutation, the FMO-PPI can be applied to the new snapshot and the results can be compared with original (WT) structure. In this way, the dynamics of the system and the change in PPIs can be analyzed by FMO-PPI. For example, mutation $F142A^{BLIP}$ (in

complex with TEM1, PDB code 1S0W) [71] induces a significant conformational change of the neighboring amino acids (*Supporting Information* Figure S3): $Y105^{TEM1}$ rotates away and the distance between salt bridge residues $E104^{TEM1}$ and $K74^{BLIP}$ increases from 2.68Å in the WT complex to 4.45Å in mutated structure. FMO-PPI is able to detect the residues that play a key role in the formation of the WT complex without needing to take into account the structure change upon mutation. However, FMO-PPI cannot predict these structural changes, which need to be addressed either experimentally (using crystallography or other methods for determining atomic structures of proteins) or computationally (using protein modelling and molecular dynamics).

According to our FMO-PPI analyses serine, arginine, lysine and glutamate are the residues frequently appearing as hotspots (Figure 8C), suggesting that polar amino acids are naturally predisposed to form strong PPIs.

The information produced by FMO-PPI can be used to design SDM experiments and to support SBDD-PPI programs in ways that have not been possible previously. For example, the identification of hotspots and the quantification of their interaction energies in the PPI can be mapped on the top of the molecular surface of the complex and combined with Site Finder to improve the design of interface ligand binding sites.

We retrospectively applied FMO-PPI protocol (Figure 5) to two SBDD-PPI cases. The practicality and efficiency of this approach is enhanced by the generation of structural information for the protein-protein interface, which is essential to drive SBDD-PPI. This approach can inform the VS, hit-to-lead or lead optimization stages of SBDD-PPI and help in design of PPI inhibitors or glues. FMO-PPI can be also applied to protein-peptide complexes as demonstrated for PTH1R-ePTH system.

In some cases, the experimental structure of the protein-protein complex is not available. However, alternative computational methods (such as Rosetta [72], MolFit [73], HADDOCK [74], MDockPP [75] and other software participating to the CAPRI challenge [76, 77]) together with information from the AlphaFold database [78] can provide predicted models of the protein-protein complex, which could be further analyzed with FMO-PPI to establish the key interactions for experimental validation and the design of potential PPI modulators.

FMO calculations focus solely on the enthalpic contributions of the hotspots to PPIs TIE (total interaction energy) and do not include entropy, long distance electrostatic PPIs (due to the truncated system) and energy terms that are related to structural changes, which can also affect protein-protein binding affinity. These additional energy terms are obviously included in experimentally measured free energy of binding ($\Delta\Delta G$). Therefore, although there is qualitative agreement with experimental findings, there is no direct quantitative correlation between FMO-PPI TIEs and experimentally measured $\Delta\Delta G$ values. However, Lim et al.[94] observed quantitative correlation between experimentally measured and calculated free energies when FMO-DFTB was combined with machine learning algorithms.

A similar phenomenon is observed for protein-ligand binding. Apart from the protein-ligand TIE calculated in this study, the binding event might be driven by different energy terms including direct enthalpic contributions, entropy, solvation and the "strain energy" of the ligand's bioactive conformation [31]. Despite the fact that not all these factors are accounted for we expected to see some correlation between measured and calculated binding affinities as observed in the case of SOS1 inhibitors.[48] These inhibitors belong to 3 related chemical subseries and their FMO calculated TIEs correlate well with their experimentally obtained pIC_{50} values ($R^2 = 0.60$, see *Supporting Information* Figure S7E). Some improvement in correlation can be observed if FMO TIE values are combined with the number of rotatable bonds ($R^2 = 0.67$, see *Supporting Information* Figure S7F). Number of rotatable bonds in a way mimics the entropic contributions. Experimentally measured binding affinity and computed TIE usually correlate well in specific cases where TIE values are calculated for the same chemical series of small molecules that have a similar chemical structure (usually taken from the same SAR) and where the entropic contribution is almost identical. In these cases, the affinity depends almost exclusively on enthalpy [34, 79, 80].

CONCLUSIONS

In this work, we propose a new QM protocol for the analysis of PPIs and SBDD of PPI modulators where FMO calculations are performed using differential PIEs. This protocol allows a detailed molecular understanding of not only intermolecular but also intramolecular interactions and water bridges – elements that are crucial for the formation of the protein-protein complex. In the current work, we focused solely on hotspots located on the protein-protein interface that are able to be targeted by PPI modulators. However, non-interface residues from deeper layers of proteins can also contribute through the formation of long-distance electrostatic PPIs.

We have shown that FMO-PPI successfully identifies PPI hotspots, generating: i. the list of PPIs that are formed by each residue with neighboring residues at the protein-protein interface, ii. the strength of these interactions in kcal/mol, iii. their chemical nature (electrostatics or hydrophobic) and iv. whether these PPIs are involved in inter-, intra-molecular interactions or in water-bridges. We have also shown that hydrophobic and underappreciated interactions play a key role in PPIs, and that FMO is able to detect and quantify these interactions in an accurate way. Our findings show that FMO-PPI results are consistent with the results of SDM experiments and can be used to rationalize these with a structure-based approach.

FMO-PPI is a new computational method that opens an alternative approach for SBDD-PPI. We confirmed

this by showing two examples of the retrospective application of FMO-PPI to the design of PPI inhibitors for the KRAS-SOS1 complex and molecular glues for the CRBN-LVY-CK1α complex. In each case, FMO-PPI, together with Site Finder, would have guided the identification of suitable pockets for the design of PPI modulators, identifying crucial hotspots at the interface and assisting the optimization of hit fragments.

FMO-PPI is a pragmatically driven approach that, in two relatively simple FMO calculations, can obtain information on every interface residue and its individual contribution to overall PPI TIE, including its role and chemical nature. This protocol does not aim to mimic experimental alanine scans because: (1) the mutation into alanine (or into any other residue) can result in structural change and information on the nature of this change is rarely available for every mutated residue; (2) reproducing an alanine scan computationally requires a large number of FMO calculations (as number of residues in each protein) and extensive calculation time, making the method impractical for short iterations of SBDD. However, whilst it is not practical to apply FMO to all experimentally mutated residues, the application of the FMO-PPI protocol integrated with structure optimization methods can be highly efficient for identifying key residues to be mutated for SDM experiments.

METHODS

Structure Preparation

Hydrogen atoms were added to the crystal structures at physiological pH (7.4) with the Protonate 3D [81] tool implemented in MOE (Molecular Operating Environment, Computational Chemistry Group, version 2020.09) [60]. The purpose of the Protonate3D function is to assign ionization states and position hydrogens in proteins, ligands and solvents given 3D coordinates.

The protonated structures were subjected to constrained minimization to optimize the interaction geometries required for accurate FMO calculations. Small errors in the positions of atoms depend on the B-factor of the atom and overall resolution of the structure, and could translate in large deviations in energy terms [82]. In this work we used a constrained minimization procedure with the semi empirical Amber10:EHT force-field [83, 84] implemented in MOE, where each atom was allowed to deviate up to 0.5 Å from its original position in the crystal structure.

Definition of the protein-protein interface

The protein-protein interface was detected with the Protein Contacts tool implemented in MOE. Two protein surface residues are in contact if the distance between (any of their atoms) them is less than ≤ 4.5Å. Interface residues of a protein are the residues that contact with residues from the interacting proteins. Only interface residues were included in the FMO calculations.

FMO-PPI calculation protocol

FMO is a quantum mechanical method which divides the system into fragments [25, 50]. Both amino acids and water molecules represent individual fragments. The detailed description of the fragmentation strategy and the basic methodology can be found in the published reviews [25, 27, 50], including detailed mathematical formulation that are beyond the scope of this manuscript.

The FMO method [50] was applied using FMO code [32] as embedded in GAMESS (General Atomic and Molecular Electronic Structure System version Jun2020a) [85], which is a general *ab initio* quantum chemistry package. In our calculations, we used the MP2 method (2nd order Møller-Plesset perturbation theory [86]) with the 6-31G* basis set and polarizable continuum model (PCM) [87] solvation. We have elected to use 6-31G* instead of 6-31G** as a reasonable compromise between accuracy and efficiency: PIE, the focus of the present report, is minimally affected by the addition of polarization functions, whereas computational time is significantly increased when using 6-31G**. C-PCM was used at the level of FMO/PCM, with the charge renormalization ICOMP=2, IDISP=0, default van-der-Waals radii and partial screening [88]. By performing QM computations on fragments one can achieve high computational speed. The FMO method has been efficiently parallelized for high performance computer clusters.

In this work, we propose a new FMO protocol to calculate the pair interaction energy [45] (see *Supporting Information* Figure S1) between two proteins. Instead of mutating each single residue to Gly and running FMO on each of them separately, we calculated all the differential energy terms between wild-type (WT) and mutants simultaneously, by computing the WT full-atom pairwise interaction energy [45] and the backbone-backbone PIE for each residue pair in two independent calculations. To compute the backbone-backbone PIE, we deleted the sidechains from the WT full-atom complex, obtaining a system that includes only Gly residues for the proteins, water molecules and ligands (if any).

Given residues *i* and *j*, their pairwise interaction energy (ΔPIE_{ij}) is calculated by deducting the PIE_{ij} of the backbone (PIE_{ij}^{bb}) from the PIE of full-atom system (PIE_{ij}^{fa}), according to the following equation (1):

$$\Delta PIE_{ij} = PIE_{ij}^{fa} - PIE_{ij}^{bb} \quad (1)$$

This method excludes backbone-backbone interaction energies from the total PIE_{ij} between two residues and keeps all interactions for side chains. The rationale behind this choice is that backbone-backbone interactions cannot be experimentally tested with SDM and, thus, used for validation of the method. Moreover, they would introduce more background noise in the analyses of the pair interaction energies. As an alternative to our FMO-PPI protocol, it may be

possible to use higher-order FMO, such as FMO4 [89], to define interactions between the side chains of two proteins or to use partition analysis based on FMO2-DFTB [90].

The fragmentation in FMO [91] results in residue fragments shifted by a carbonyl group relative to conventional residues, as driven by the accuracy considerations for fragmentation. This does not create problems in discussing the interactions of residue fragments except for the case when the carbonyl is the group that interacts. In this case, the interaction energy needs to be manually curated and assigned to the correct residue. The structure was examined to determine such cases and the interaction energies involving carbonyl groups were assigned to the residue to which the carbonyl belongs, not to the residue fragment it is assigned to in FMO. Thereby, the interactions reported in this work correspond to the actual residues, not residue fragments.

Analysis of FMO-PPI results

FMO outputs from the full-atom and backbone systems were processed with an internal script to extract the pairwise energy terms and calculate the PIE_{ij} values. Water bridges were also considered, whereas solvent interactions with single amino acids were discarded. Contacts were then visualized and analyzed in MOE with a SVL script provided by CCG [92].

Finally, significant PPIs with PIE_{ij} below -3 or higher than 3 kcal/mol were plotted as heat map with Python (version 2.7.9), together with the ratio of electrostatics (PIE_{ij}^{es}) and its sum with the dispersion term (PIE_{ij}^{disp}):

$$(2) \qquad f_{ij} = \Delta PIE_{ij}^{es} / (\Delta PIE_{ij}^{es} + \Delta PIE_{ij}^{disp})$$

This factor distinguishes between electrostatic (ratio above 0.5) and hydrophobic (ratio below 0.5) interactions. Duplicate intermolecular interactions were removed, leaving the bottom-left corner of the heat map empty.

Hotspots were mapped on the surface of the structure complexes, using an SVL script provided by CCG, which colored the residues according to the PIE values. For each residue we used the highest PIE value that belonged to this amino acid among its inter-molecular interactions. In this way, hotspots can be easily visualized.

Identifying potential ligand binding site

We used Site Finder, that is implemented in MOE[60], to search for potential interface ligand binding sites. This tool does not consider any results from FMO but uses only the 3D coordinates of the protein target or protein-protein complex.

Site Finder is a geometric method based on the calculation of alpha spheres, which are collections of 3D points in the site. The relative position and accessibility of the amino acids are taken into consideration together with a rough classification of chemical type. Regions that present a tight atomic packing and which are either not too exposed to the solvent or inaccessible are identified. The sites are then classified as either hydrophobic or polar and alpha spheres are calculated. Finally, sites are ranked according to the Propensity for Ligand Binding (PLB) score [93], which is based on the composition of amino acids in the pocket.

DATA AND SOFTWARE AVAILABILITY

Protein structures in this manuscript have been downloaded from Protein Data Bank (PDB, see Table 1). The FMO protocol was applied using FMO code (https://www.msg.chem.iastate.edu/gamess/index.html) as embedded in open source GAMESS (General Atomic and Molecular Electronic Structure System version Jun2020a) which is a general *ab initio* quantum chemistry package. Site Finder, Protein Contacts and Protonate 3D tools implemented in MOE (Molecular Operating Environment, Computational Chemistry Group, version 2020.09). The full description of these tools appears in *Methods* section of this manuscript.

ASSOCIATED CONTENT

Supporting Information includes general FMO workflow, FMO-PPI and PIEDA heat maps plots, tables with experimental SDM data extracted from literature and table with total number of interface residues.

AUTHOR INFORMATION

Corresponding Author

Dr. Alexander Heifetz, email address: Alexander.Heifetz@evotec.com

ACKNOWLEDGMENTS

A.H. and A.T.N would like to acknowledge the support of European Commission Horizon 2020 – EINFRAEDI-01-2018 (H2020 RIA 823712) CompBioMed2 (http://www.compbiomed.eu/). We thank Andrew Henry, Nadia Li and Guido Kirsten from CCG Support for providing us with SVL scripts to visualize and analyze the FMO-PPI results in MOE. We thank Dr. Inaki Morao, Dr. Tahsin Kellici and Dr. William R. Pitt for their insightful feedback. We thank Benjamin C. B. Symons for running the FMO calculations for *Supporting Information* Figure S7E.

ABBREVIATIONS

FMO, Fragment Molecular Orbitals method; PPIs, Protein-protein interactions; SDM, site directed mutagenesis; GPCR, G protein-coupled receptor; TM, transmembrane helix; ECL, extracellular loop; PDB, Protein Data Bank;

SBDD, structure-based drug discovery; PIE, Pairwise Interaction Energy.

REFERENCES

1. Snider, J.; Kotlyar, M.; Saraon, P.; Yao, Z.; Jurisica, I.; Stagljar, I., Fundamentals of protein interaction network mapping. *Molecular Systems Biology* **2015**, 11, 848.

2. Koh, G. C. K. W.; Porras, P.; Aranda, B.; Hermjakob, H.; Orchard, S. E., Analyzing Protein–Protein Interaction Networks. *Journal of Proteome Research* **2012**, 11, 2014-2031.

3. Lu, H.; Zhou, Q.; He, J.; Jiang, Z.; Peng, C.; Tong, R.; Shi, J., Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials. *Signal Transduction and Targeted Therapy* **2020**, 5, 213.

4. White, A. W.; Westwell, A. D.; Brahemi, G., Protein–protein interactions as targets for small-molecule therapeutics in cancer. *Expert Reviews in Molecular Medicine* **2008**, 10, e8.

5. Blazer, L. L.; Neubig, R. R., Small Molecule Protein–Protein Interaction Inhibitors as CNS Therapeutic Agents: Current Progress and Future Hurdles. *Neuropsychopharmacology* **2009**, 34, 126-141.

6. Luck, K.; Kim, D.-K.; Lambourne, L.; Spirohn, K.; Begg, B. E.; Bian, W.; Brignall, R.; Cafarelli, T.; Campos-Laborie, F. J.; Charloteaux, B.; Choi, D.; Coté, A. G.; Daley, M.; Deimling, S.; Desbuleux, A.; Dricot, A.; Gebbia, M.; Hardy, M. F.; Kishore, N.; Knapp, J. J.; Kovács, I. A.; Lemmens, I.; Mee, M. W.; Mellor, J. C.; Pollis, C.; Pons, C.; Richardson, A. D.; Schlabach, S.; Teeking, B.; Yadav, A.; Babor, M.; Balcha, D.; Basha, O.; Bowman-Colin, C.; Chin, S.-F.; Choi, S. G.; Colabella, C.; Coppin, G.; D'Amata, C.; De Ridder, D.; De Rouck, S.; Duran-Frigola, M.; Ennajdaoui, H.; Goebels, F.; Goehring, L.; Gopal, A.; Haddad, G.; Hatchi, E.; Helmy, M.; Jacob, Y.; Kassa, Y.; Landini, S.; Li, R.; van Lieshout, N.; MacWilliams, A.; Markey, D.; Paulson, J. N.; Rangarajan, S.; Rasla, J.; Rayhan, A.; Rolland, T.; San-Miguel, A.; Shen, Y.; Sheykhkarimli, D.; Sheynkman, G. M.; Simonovsky, E.; Taşan, M.; Tejeda, A.; Tropepe, V.; Twizere, J.-C.; Wang, Y.; Weatheritt, R. J.; Weile, J.; Xia, Y.; Yang, X.; Yeger-Lotem, E.; Zhong, Q.; Aloy, P.; Bader, G. D.; De Las Rivas, J.; Gaudet, S.; Hao, T.; Rak, J.; Tavernier, J.; Hill, D. E.; Vidal, M.; Roth, F. P.; Calderwood, M. A., A reference map of the human binary protein interactome. *Nature* **2020**, 580, 402-408.

7. Smith, M. C.; Gestwicki, J. E., Features of protein–protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert Reviews in Molecular Medicine* **2012**, 14, e16.

8. Jubb, H. C.; Pandurangan, A. P.; Turner, M. A.; Ochoa-Montaño, B.; Blundell, T. L.; Ascher, D. B., Mutations at protein-protein interfaces: Small changes over big surfaces have large impacts on human health. *Progress in Biophysics and Molecular Biology* **2017**, 128, 3-13.

9. Thirumal Kumar, D.; Sneha, P.; Uppin, J.; Usha, S.; George Priya Doss, C. Investigating the influence of hotspot mutations in protein–protein interaction of IDH1 homodimer protein: a computational approach. In *Advances in Protein Chemistry and Structural Biology*, Donev, R., Ed.; Academic Press: 2018; Vol. 111, pp 243-261.

10. Hashemi, Z. S.; Zarei, M.; Fath, M. K.; Ganji, M.; Farahani, M. S.; Afsharnouri, F.; Pourzardosht, N.; Khalesi, B.; Jahangiri, A.; Rahbar, M. R.; Khalili, S., In silico approaches for the design and optimization of interfering peptides against protein–protein interactions. *Frontiers in Molecular Biosciences* **2021**, 8, 669431.

11. Wells, J. A., Systematic mutational analyses of protein-protein interfaces. In *Methods in Enzymology*; Academic Press: 1991; Vol. 202, pp 390-411.

12. Guvench, O.; MacKerell, A. D., Jr., Computational Fragment-Based Binding Site Identification by Ligand Competitive Saturation. *PLOS Computational Biology* **2009**, 5, e1000435.

13. Brenke, R.; Kozakov, D.; Chuang, G.-Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S., Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics* **2009**, 25, 621-627.

14. Bhattacharya, S.; Vaidehi, N., Differences in Allosteric Communication Pipelines in the Inactive and Active States of a GPCR. *Biophysical Journal* **2014**, 107, 422-434.

15. Nivedha, A. K.; Tautermann, C. S.; Bhattacharya, S.; Lee, S.; Casarosa, P.; Kollak, I.; Kiechle, T.; Vaidehi, N., Identifying Functional Hotspot Residues for Biased Ligand Design in G-Protein-Coupled Receptors. *Molecular Pharmacology* **2018**, 93, 288-296.

16. Bissantz, C.; Kuhn, B.; Stahl, M., A medicinal chemist's guide to molecular interactions. *J Med Chem* **2010**, 53, 5061-84.

17. Ozawa, T.; Okazaki, K.; Kitaura, K., CH/π hydrogen bonds play a role in ligand recognition and equilibrium between active and inactive states of the β2 adrenergic receptor: An ab initio fragment molecular orbital (FMO) study. *Bioorganic & Medicinal Chemistry* **2011**, 19, 5231-5237.

18. Lu, Y.-X.; Zou, J.-W.; Wang, Y.-H.; Yu, Q.-S., Substituent effects on noncovalent halogen/π interactions: Theoretical study. *International Journal of Quantum Chemistry* **2007**, 107, 1479-1486.

19. Gallivan, J. P.; Dougherty, D. A., Cation-π interactions in structural biology. *Proceedings of the National Academy of Sciences* **1999**, 96, 9459-9464.

20. Johnston, R. C.; Cheong, P. H.-Y., C–H···O non-classical hydrogen bonding in the stereomechanics of organic transformations: theory and recognition. *Organic & Biomolecular Chemistry* **2013**, 11, 5057-5064.

21. Raha, K.; Peters, M. B.; Wang, B.; Yu, N.; Wollacott, A. M.; Westerhoff, L. M.; Merz, K. M., The role of quantum mechanics in structure-based drug design. *Drug Discovery Today* **2007**, 12, 725-731.

22. Tong, Y.; Mei, Y.; Li, Y. L.; Ji, C. G.; Zhang, J. Z. H., Electrostatic Polarization Makes a Substantial Contribution to the Free Energy of Avidin−Biotin Binding. *Journal of the American Chemical Society* **2010**, 132, 5137-5142.

23. Popov, P.; Peng, Y.; Shen, L.; Stevens, R. C.; Cherezov, V.; Liu, Z.-J.; Katritch, V., Computational design of thermostabilizing point mutations for G protein-coupled receptors. *eLife* **2018**, 7, e34729.

24. Yu, N.; Li, X.; Cui, G.; Hayik, S. A.; Merz, K. M., 2nd, Critical assessment of quantum mechanics based energy restraints in protein crystal structure refinement. *Protein science : a publication of the Protein Society* **2006**, 15, 2773-84.

25. Fedorov, D. G.; Kitaura, K., Extending the power of quantum chemistry to large systems with the fragment molecular orbital method. *The journal of physical chemistry. A* **2007**, 111, 6904-14.

26. Fukuzawa, K.; Tanaka, S., Fragment molecular orbital calculations for biomolecules. *Current Opinion in Structural Biology* **2022**, 72, 127-134.

27. Fedorov, D. G.; Kitaura, K., Pair interaction energy decomposition analysis. *J Comput Chem* **2007**, 28, 222-37.

28. Phipps, M. J. S.; Fox, T.; Tautermann, C. S.; Skylaris, C.-K., Energy decomposition analysis approaches and their evaluation on prototypical protein–drug interaction patterns. *Chemical Society Reviews* **2015**, 44, 3177-3211.

29. Hatada, R.; Okuwaki, K.; Mochizuki, Y.; Handa, Y.; Fukuzawa, K.; Komeiji, Y.; Okiyama, Y.; Tanaka, S., Fragment Molecular Orbital Based Interaction Analyses on COVID-19 Main Protease − Inhibitor N3 Complex (PDB ID: 6LU7). *Journal of Chemical Information and Modeling* **2020**, 60, 3593-3602.

30. Hengphasatporn, K.; Wilasluck, P.; Deetanya, P.; Wangkanont, K.; Chavasiri, W.; Visitchanakun, P.; Leelahavanichkul, A.; Paunrat, W.; Boonyasuppayakorn, S.; Rungrotmongkol, T.; Hannongbua, S.; Shigeta, Y., Halogenated Baicalein as a Promising Antiviral Agent toward SARS-CoV-2 Main Protease. *Journal of Chemical Information and Modeling* **2022**, 62, 1498-1509.

31. Heifetz, A.; Chudyk, E. I.; Gleave, L.; Aldeghi, M.; Cherezov, V.; Fedorov, D. G.; Biggin, P. C.; Bodkin, M. J., The Fragment Molecular Orbital Method Reveals New Insight into the Chemical Nature of GPCR–Ligand Interactions. *Journal of Chemical Information and Modeling* **2016**, 56, 159-172.

32. Alexeev, Y.; Mazanetz, M. P.; Ichihara, O.; Fedorov, D. G., GAMESS as a free quantum-mechanical platform for drug research. *Current topics in medicinal chemistry* **2012**, 12, 2013-33.

33. Tanaka, S.; Watanabe, C.; Honma, T.; Fukuzawa, K.; Ohishi, K.; Maruyama, T., Identification of correlated inter-residue interactions in protein complex based on the fragment molecular orbital method. *Journal of Molecular Graphics and Modelling* **2020**, 100, 107650.

34. Chudyk, E. I.; Sarrat, L.; Aldeghi, M.; Fedorov, D. G.; Bodkin, M. J.; James, T.; Southey, M.; Robinson, R.; Morao, I.; Heifetz, A. Exploring GPCR-Ligand Interactions with the Fragment Molecular Orbital (FMO) Method. In *Computational Methods for GPCR Drug Discovery*, Heifetz, A., Ed.; Springer New York: New York, NY, 2018, pp 179-195.

35. Heifetz, A.; Morao, I.; Babu, M. M.; James, T.; Southey, M. W. Y.; Fedorov, D. G.; Aldeghi, M.; Bodkin, M. J.; Townsend-Nicholson, A., Characterizing Interhelical Interactions of G-Protein Coupled Receptors with the Fragment Molecular Orbital Method. *Journal of Chemical Theory and Computation* **2020**, 16, 2814-2824.

36. Fukuzawa, K.; Kato, K.; Watanabe, C.; Kawashima, Y.; Handa, Y.; Yamamoto, A.; Watanabe, K.; Ohyama, T.; Kamisaka, K.; Takaya, D.; Honma, T., Special Features of COVID-19 in the FMODB: Fragment Molecular Orbital Calculations and Interaction Energy Analysis of SARS-CoV-2-Related Proteins. *Journal of Chemical Information and Modeling* **2021**, 61, 4594-4612.

37. Hwang, S.; Baek, S.-H.; Park, D., Interaction Analysis of the Spike Protein of Delta and Omicron Variants of SARS-CoV-2 with hACE2 and Eight Monoclonal Antibodies Using the Fragment Molecular Orbital Method. *Journal of Chemical Information and Modeling* **2022**, 62, 1771-1782.

38. Heifetz, A.; Sladek, V.; Townsend-Nicholson, A.; Fedorov, D. G. Characterizing Protein-Protein Interactions with the Fragment Molecular Orbital Method. In *Quantum Mechanics in Drug Discovery*, Heifetz, A., Ed.; Springer US: New York, NY, 2020, pp 187-205.

39. Kim, J.; Lim, H.; Moon, S.; Cho, S. Y.; Kim, M.; Park, J. H.; Park, H. W.; No, K. T., Hot Spot Analysis of YAP-TEAD Protein-Protein Interaction Using the Fragment Molecular Orbital Method and Its Application for Inhibitor Discovery. *Cancers* **2021**, 13, 4246.

40. Ozono, H.; Ishikawa, T., Visualization of the Interfacial Electrostatic Complementarity: A Method for Analysis of Protein–Protein Interaction Based on Ab Initio Quantum Chemical Calculations. *Journal of Chemical Theory and Computation* **2021**, 17, 5600-5610.

41. Heifetz, A.; James, T.; Southey, M.; Morao, I.; Aldeghi, M.; Sarrat, L.; Fedorov, D. G.; Bodkin, M. J.; Townsend-Nicholson, A., Characterising GPCR–ligand interactions using a fragment molecular orbital-based approach. *Current Opinion in Structural Biology* **2019**, 55, 85-92.

42. Grünberg, R.; Leckner, J.; Nilges, M., Complementarity of Structure Ensembles in Protein-Protein Binding. *Structure* **2004**, 12, 2125-2136.

43. Xu, D.; Tsai, C. J.; Nussinov, R., Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Engineering, Design and Selection* **1997**, 10, 999-1012.

44. Lim, D.; Park, H. U.; De Castro, L.; Kang, S. G.; Lee, H. S.; Jensen, S.; Lee, K. J.; Strynadka, N. C. J., Crystal structure and kinetic analysis of β-lactamase inhibitor protein-II in complex with TEM-1 β-lactamase. *Nature Structural Biology* **2001**, 8, 848-852.

45. Thomas, C.; Moraga, I.; Levin, D.; Krutzik, Peter O.; Podoplelova, Y.; Trejo, A.; Lee, C.; Yarden, G.; Vleck, Susan E.; Glenn, Jeffrey S.; Nolan, Garry P.; Piehler, J.; Schreiber, G.; Garcia, K. C., Structural Linkage between Ligand Discrimination and Receptor Activation by Type I Interferons. *Cell* **2011**, 146, 621-632.

46. Ehrenmann, J.; Schöppe, J.; Klenk, C.; Rappas, M.; Kummer, L.; Doré, A. S.; Plückthun, A., High-resolution crystal structure of parathyroid hormone 1 receptor in complex with a peptide agonist. *Nature Structural & Molecular Biology* **2018**, 25, 1086-1092.

47. Hamill, S.; Lou, Hua J.; Turk, Benjamin E.; Boggon, Titus J., Structural Basis for Noncanonical Substrate Recognition of Cofilin/ADF Proteins by LIM Kinases. *Molecular Cell* **2016**, 62, 397-408.

48. Hillig, R. C.; Sautier, B.; Schroeder, J.; Moosmayer, D.; Hilpmann, A.; Stegmann, C. M.; Werbeck, N. D.; Briem, H.; Boemer, U.; Weiske, J.; Badock, V.; Mastouri, J.; Petersen, K.; Siemeister, G.; Kahmann, J. D.; Wegener, D.; Böhnke, N.; Eis, K.; Graham, K.; Wortmann, L.; von Nussbaum, F.; Bader, B., Discovery of potent SOS1 inhibitors that block RAS activation via disruption of the RAS–SOS1 interaction. *Proceedings of the National Academy of Sciences* **2019**, 116, 2551-2560.

49. Petzold, G.; Fischer, E. S.; Thomä, N. H., Structural basis of lenalidomide-induced CK1α degradation by the CRL4CRBN ubiquitin ligase. *Nature* **2016**, 532, 127-130.

50. Fedorov, D. G.; Nagata, T.; Kitaura, K., Exploring chemistry with the fragment molecular orbital method. *Phys Chem Chem Phys* **2012**, 14, 7562-77.

51. Heifetz, A.; Trani, G.; Aldeghi, M.; MacKinnon, C. H.; McEwan, P. A.; Brookfield, F. A.; Chudyk, E. I.; Bodkin, M.; Pei, Z.; Burch, J. D.; Ortwine, D. F., Fragment Molecular Orbital Method Applied to Lead Optimization of Novel Interleukin-2 Inducible T-Cell Kinase (ITK) Inhibitors. *Journal of Medicinal Chemistry* **2016**, 59, 4352-4363.

52. Eiamphungporn, W.; Schaduangrat, N.; Malik, A. A.; Nantasenamat, C., Tackling the Antibiotic Resistance Caused by Class A β-Lactamases through the Use of β-Lactamase Inhibitory Protein. *International Journal of Molecular Sciences* **2018**, 19, 2222.

53. Borden, E. C.; Sen, G. C.; Uze, G.; Silverman, R. H.; Ransohoff, R. M.; Foster, G. R.; Stark, G. R., Interferons at age 50: past, current and future impact on biomedicine. *Nature Reviews Drug Discovery* **2007**, 6, 975-990.

54. Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E., Trends in GPCR drug discovery: new agents, targets and indications. *Nature Reviews Drug Discovery* **2017**, 16, 829-842.

55. Davenport, A. P.; Scully, C. C. G.; de Graaf, C.; Brown, A. J. H.; Maguire, J. J., Advances in therapeutic peptides targeting G protein-coupled receptors. *Nature Reviews Drug Discovery* **2020**, 19, 389-413.

56. Gensure, R. C.; Shimizu, N.; Tsang, J.; Gardella, T. J., Identification of a Contact Site for Residue 19 of Parathyroid Hormone (PTH) and PTH-Related Protein Analogs in Transmembrane Domain Two of the Type 1 PTH Receptor. *Molecular Endocrinology* **2003**, 17, 2647-2658.

57. Cohen, P.; Cross, D.; Jänne, P. A., Kinase drug discovery 20 years after imatinib: progress and future directions. *Nature Reviews Drug Discovery* **2021**, 20, 551-569.

58. Ayala-Aguilera, C. C.; Valero, T.; Lorente-Macías, Á.; Baillache, D. J.; Croke, S.; Unciti-Broceta, A., Small Molecule Kinase Inhibitor Drugs (1995–2021): Medical Indication, Pharmacology, and Synthesis. *Journal of Medicinal Chemistry* **2022**, 65, 1047-1131.

59. Salah, E.; Chatterjee, D.; Beltrami, A.; Tumber, A.; Preuss, F.; Canning, P.; Chaikuad, A.; Knaus, P.; Knapp, S.; Bullock, A. N.; Mathea, S., Lessons from LIMK1 enzymology and their impact on inhibitor design. *Biochemical Journal* **2019**, 476, 3197-3209.

60. Chemical Computing Group ULC, S. S. W., Suite #910, Montreal, QC, Canada, H3A 2R7 *Molecular Operating Environment (MOE), 2019.01*, 2020.09; 2021.

61. Vigil, D.; Cherfils, J.; Rossman, K. L.; Der, C. J., Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? *Nature Reviews Cancer* **2010**, 10, 842-857.

62. Winter, J. J. G.; Anderson, M.; Blades, K.; Brassington, C.; Breeze, A. L.; Chresta, C.; Embrey, K.; Fairley, G.; Faulder, P.; Finlay, M. R. V.; Kettle, J. G.; Nowak, T.; Overman, R.; Patel, S. J.; Perkins, P.; Spadola, L.; Tart, J.; Tucker, J. A.; Wrigley, G., Small Molecule Binding Sites on the Ras:SOS Complex Can Be Exploited for Inhibition of Ras Activation. *Journal of Medicinal Chemistry* **2015**, 58, 2265-2274.

63. Patricelli, M. P.; Janes, M. R.; Li, L.-S.; Hansen, R.; Peters, U.; Kessler, L. V.; Chen, Y.; Kucharski, J. M.;

Feng, J.; Ely, T.; Chen, J. H.; Firdaus, S. J.; Babbar, A.; Ren, P.; Liu, Y., Selective Inhibition of Oncogenic KRAS Output with Small Molecules Targeting the Inactive State. *Cancer Discovery* **2016**, 6, 316-329.

64. Kessler, D.; Gmachl, M.; Mantoulidis, A.; Martin Laetitia, J.; Zoephel, A.; Mayer, M.; Gollner, A.; Covini, D.; Fischer, S.; Gerstberger, T.; Gmaschitz, T.; Goodwin, C.; Greb, P.; Häring, D.; Hela, W.; Hoffmann, J.; Karolyi-Oezguer, J.; Knesl, P.; Kornigg, S.; Koegl, M.; Kousek, R.; Lamarre, L.; Moser, F.; Munico-Martinez, S.; Peinsipp, C.; Phan, J.; Rinnenthal, J.; Sai, J.; Salamon, C.; Scherbantin, Y.; Schipany, K.; Schnitzer, R.; Schrenk, A.; Sharps, B.; Siszler, G.; Sun, Q.; Waterson, A.; Wolkerstorfer, B.; Zeeb, M.; Pearson, M.; Fesik Stephen, W.; McConnell Darryl, B., Drugging an undruggable pocket on KRAS. *Proceedings of the National Academy of Sciences* **2019**, 116, 15823-15829.

65. Krönke, J.; Fink, E. C.; Hollenbach, P. W.; MacBeth, K. J.; Hurst, S. N.; Udeshi, N. D.; Chamberlain, P. P.; Mani, D. R.; Man, H. W.; Gandhi, A. K.; Svinkina, T.; Schneider, R. K.; McConkey, M.; Järås, M.; Griffiths, E.; Wetzler, M.; Bullinger, L.; Cathers, B. E.; Carr, S. A.; Chopra, R.; Ebert, B. L., Lenalidomide induces ubiquitination and degradation of CK1α in del(5q) MDS. *Nature* **2015**, 523, 183-188.

66. Krönke, J.; Udeshi, N. D.; Narla, A.; Grauman, P.; Hurst, S. N.; McConkey, M.; Svinkina, T.; Heckl, D.; Comer, E.; Li, X.; Ciarlo, C.; Hartman, E.; Munshi, N.; Schenone, M.; Schreiber, S. L.; Carr, S. A.; Ebert, B. L., Lenalidomide Causes Selective Degradation of IKZF1 and IKZF3 in Multiple Myeloma Cells. *Science* **2014**, 343, 301-305.

67. Lim, H.; Chun, J.; Jin, X.; Kim, J.; Yoon, J.; No, K. T., Investigation of protein-protein interactions and hot spot region between PD-1 and PD-L1 by fragment molecular orbital method. *Scientific Reports* **2019**, 9, 16727.

68. Lim, H.; Baek, A.; Kim, J.; Kim, M. S.; Liu, J.; Nam, K.-Y.; Yoon, J.; No, K. T., Hot spot profiles of SARS-CoV-2 and human ACE2 receptor protein protein interaction obtained by density functional tight binding fragment molecular orbital method. *Scientific Reports* **2020**, 10, 16862.

69. Anzaki, S.; Watanabe, C.; Fukuzawa, K.; Mochizuki, Y.; Tanaka, S., Interaction energy analysis on specific binding of influenza virus hemagglutinin to avian and human sialosaccharide receptors: Importance of mutation-induced structural change. *Journal of Molecular Graphics and Modelling* **2014**, 53, 48-58.

70. Okuwaki, K.; Akisawa, K.; Hatada, R.; Mochizuki, Y.; Fukuzawa, K.; Komeiji, Y.; Tanaka, S., Collective residue interactions in trimer complexes of SARS-CoV-2 spike proteins analyzed by fragment molecular orbital method. *Applied Physics Express* **2021**, 15, 017001.

71. Reichmann, D.; Rahat, O.; Albeck, S.; Meged, R.; Dym, O.; Schreiber, G., The modular architecture of protein–protein binding interfaces. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, 102, 57-62.

72. Chaudhury, S.; Berrondo, M.; Weitzner, B. D.; Muthu, P.; Bergman, H.; Gray, J. J., Benchmarking and Analysis of Protein Docking Performance in Rosetta v3.2. *PLOS ONE* **2011**, 6, e22477.

73. Kowalsman, N.; Eisenstein, M., Inherent limitations in protein–protein docking procedures. *Bioinformatics* **2007**, 23, 421-426.

74. de Vries, S. J.; van Dijk, M.; Bonvin, A. M. J. J., The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols* **2010**, 5, 883-897.

75. Huang, S.-Y.; Zou, X., MDockPP: A hierarchical approach for protein-protein docking and its application to CAPRI rounds 15-19. *Proteins* **2010**, 78, 3096-3103.

76. Lensink, M. F.; Nadzirin, N.; Velankar, S.; Wodak, S. J., Modeling protein-protein, protein-peptide, and protein-oligosaccharide complexes: CAPRI 7th edition. *Proteins: Structure, Function, and Bioinformatics* **2020**, 88, 916-938.

77. Fernández-Recio, J.; Abagyan, R.; Totrov, M., Improving CAPRI predictions: Optimized desolvation for rigid-body docking. *Proteins: Structure, Function, and Bioinformatics* **2005**, 60, 308-313.

78. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D., Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, 596, 583-589.

79. Lim, H.; Jin, X.; Kim, J.; Hwang, S.; Shin, K. B.; Choi, J.; Nam, K.-Y.; No, K. T., Investigation of Hot Spot Region in XIAP Inhibitor Binding Site by Fragment Molecular Orbital Method. *Computational and Structural Biotechnology Journal* **2019**, 17, 1217-1225.

80. Paciotti, R.; Agamennone, M.; Coletti, C.; Storchi, L., Characterization of PD-L1 binding sites by a combined FMO/GRID-DRY approach. *Journal of Computer-Aided Molecular Design* **2020**, 34, 897-914.

81. Labute, P., Protonate3D: assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins* **2009**, 75, 187-205.

82. Jazayeri, A.; Dias, J. M.; Marshall, F. H., From G Protein-coupled Receptor Structure Resolution to Rational Drug Design. *The Journal of biological chemistry* **2015**, 290, 19489-95.

83. Gerber, P. R.; Muller, K., MAB, a generally applicable molecular force field for structure modelling in

medicinal chemistry. *Journal of computer-aided molecular design* **1995**, 9, 251-68.

84. Cerutti, D. S.; Swope, W. C.; Rice, J. E.; Case, D. A., ff14ipq: A Self-Consistent Force Field for Condensed-Phase Simulations of Proteins. *Journal of chemical theory and computation* **2014**, 10, 4515-4534.

85. Barca, G. M. J.; Bertoni, C.; Carrington, L.; Datta, D.; De Silva, N.; Deustua, J. E.; Fedorov, D. G.; Gour, J. R.; Gunina, A. O.; Guidez, E.; Harville, T.; Irle, S.; Ivanic, J.; Kowalski, K.; Leang, S. S.; Li, H.; Li, W.; Lutz, J. J.; Magoulas, I.; Mato, J.; Mironov, V.; Nakata, H.; Pham, B. Q.; Piecuch, P.; Poole, D.; Pruitt, S. R.; Rendell, A. P.; Roskop, L. B.; Ruedenberg, K.; Sattasathuchana, T.; Schmidt, M. W.; Shen, J.; Slipchenko, L.; Sosonkina, M.; Sundriyal, V.; Tiwari, A.; Galvez Vallejo, J. L.; Westheimer, B.; Włoch, M.; Xu, P.; Zahariev, F.; Gordon, M. S., Recent developments in the general atomic and molecular electronic structure system. *The Journal of Chemical Physics* **2020**, 152, 154102.

86. Fedorov, D. G.; Kitaura, K., Second order Moller-Plesset perturbation theory based upon the fragment molecular orbital method. *The Journal of chemical physics* **2004**, 121, 2483-90.

87. Fedorov, D. G.; Kitaura, K.; Li, H.; Jensen, J. H.; Gordon, M. S., The polarizable continuum model (PCM) interfaced with the fragment molecular orbital method (FMO). *Journal of Computational Chemistry* **2006**, 27, 976-985.

88. Fedorov, D. G., Solvent Screening in Zwitterions Analyzed with the Fragment Molecular Orbital Method. *Journal of Chemical Theory and Computation* **2019**, 15, 5404-5416.

89. Watanabe, C.; Fukuzawa, K.; Okiyama, Y.; Tsukamoto, T.; Kato, A.; Tanaka, S.; Mochizuki, Y.; Nakano, T., Three- and four-body corrected fragment molecular orbital calculations with a novel subdividing fragmentation method applicable to structure-based drug design. *Journal of Molecular Graphics and Modelling* **2013**, 41, 31-42.

90. Fedorov, D. G., Partition Analysis for Density-Functional Tight-Binding. *The Journal of Physical Chemistry A* **2020**, 124, 10346-10358.

91. Nakano, T.; Kaminuma, T.; Sato, T.; Akiyama, Y.; Uebayasi, M.; Kitaura, K., Fragment molecular orbital method: application to polypeptides. *Chemical Physics Letters* **2000**, 318, 614-618.

92. Scientific Vector Language (SVL) source code provided by Chemical Computing Group ULC, S. S. W., Suite #910, Montreal, QC, Canada, H3A 2R7 *CSV contact analysis*, 2021.

93. Soga, S.; Shirai, H.; Kobori, M.; Hirayama, N., Use of Amino Acid Composition to Predict Ligand-Binding Sites. *Journal of Chemical Information and Modeling* **2007**, 47, 400-406.

94. Lim, H.; Jeon HN.; Lim, S.; Jang, Y.; Kim, T.; Cho, H.; Pan, JG.; No, KT. Evaluation of protein descriptors in computer-aided rational protein engineering tasks and its application in property prediction in SARS-CoV-2 spike glycoprotein. *Computational and structural biotechnology journal* **2022**, Jan 1; 20:788-98.