

Computer- and NMR-aided design of small-molecule inhibitors of the Hub1 protein

Ewa Surmiak^{1,*}, Katarzyna Kubica¹, Atilio Reyes Romero², Jacek Plewka³, Ismael Rodriguez¹, Katarzyna Magiera-Mularz¹, Alexander Dömling² and Tad A. Holak^{1,*}

¹Department of Organic Chemistry, Faculty of Chemistry, Jagiellonian University, Gronostajowa 2, 30-387 Krakow, Poland

²University of Groningen, Department of Drug Design, A. Deusinglaan 1, 9713 AV Groningen, The Netherlands

³Malopolska Centre of Biotechnology, Jagiellonian University, Gronostajowa 7a, 30-387 Kraków, Poland

*Corresponding author(s). E-mail(s): ewa.surmiak@uj.edu.pl; holak@chemia.uj.edu.pl

Abstract

Protein-protein interactions play a key role in cell homeostasis and physiological functions of the organisms. Consequently, their malfunction leads to diseases such as cancer, metastasis and neurodegeneration. The Hub1/Snu66 interaction is responsible for controlling alternative splicing through non-covalent binding to the HIND (Hub1-Interacting Domain) domain of the Snu66 spliceosomal protein. To better understand how Hub1 works in living organisms, we conducted a study to find small molecules that have an affinity for the Snu66 binding site of Hub1. The *in silico* investigation was based on global peptide modeling followed by high-resolution protein-peptide refinement. Docking analysis of nearly 200 molecules in the binding pocket of Hub1 is also described. Finally, Nuclear Magnetic Resonance fragment-based screening was used to confirm our findings.

Keywords: Protein-protein interactions, anti-cancer therapy, nuclear magnetic resonance, protein-peptide docking, small-molecule inhibitors

Introduction

Human Hub1 (known also as UBL5; Ubiquitin-Like Protein 5) is a small protein consisting of 73 amino acids. The yeast orthologue is called Hub1 and in mammals, Beacon. Hub1s are

widely expressed and belong to a large group of ubiquitin-like proteins. They are extremely conserved across phylogeny and despite low residue similarity, are homologs to ubiquitin [1]. Hub1, like other ubiquitin-like modifiers, was reported to interact with several target proteins. The mechanism of these interactions is unique and remains largely unexplored.

There are two classes of the ubiquitin-like proteins [2]: one called ubiquitin-domain proteins (UDPs), consisting of proteins which harbor a ubiquitin fold within their polypeptide chain but do not form covalent conjugates, while the other group is ubiquitin-like modifiers (UBLs), which conjugate their target protein by covalent binding with an iso-peptide formation between their carboxyl at the C-terminus and the ϵ -amino groups of lysine residues of acceptor proteins [3].

Hub1, in contrast to other UBLs, lacks a characteristic di-glycine motif (GG motif) at its C-terminal tail. Instead, it possesses a di-tyrosine motif (YY motif), which seems to be crucial to Hub1's function and its ability to conjugate with other proteins [4]. In addition, the Hub1 polypeptide chain lacks an exposed unstructured C-terminal tail, which is critical for other ubiquitin-like modifiers for their attachment. Because Hub1 seems to possess features of both ubiquitin-like protein classes (UBLs and UDPs), it was classified as a novel protein modulator, which acts through the formation of tight, non-covalent interactions with target proteins.

The studies by Mishra et al. (2011) and Ammon et al. (2014) show that both yeast Hub1 and human UBL5 are involved in alternative splicing by non-covalent binding to the HIND (Hub1-interacting domain) of the N-terminal domain of the Snu66 (Small Nuclear Ribonucleoprotein) spliceosomal protein. Alternative splicing of pre-messenger RNA diversifies gene products in eukaryotes and is guided by factors that enable spliceosomes to recognize a particular splice site. There is an assumption that Hub1/UBL5 modifies spliceosomes so that they can use certain noncanonical 5' splice sites [5]. Because Hub1 interacts with Snu66 on the opposite side to the other ubiquitin-binding domains (UBDs), it makes this interaction unique (Mishra et al., 2011). Since malfunctions in alternative splicing can cause multiple diseases like hypercholesterolemia, premature aging, neurodegenerative diseases, as well as cancer, finding small-molecule antagonists that can inhibit the Hub1/Snu66 interactions could not only allow understanding the mode of action of Hub1 in living organisms but could also begin novel approaches to anticancer therapy in the future [2, 6].

Till now there are no reports in the literature about small molecules that bind to Hub1. Finding such moieties that can inhibit protein-protein interactions (PPIs) [7] is an extremely challenging task and requires the right choice of methods. In this study, the fragment-based NMR (Nuclear Magnetic Resonance Spectroscopy) screening aided by molecular modeling has

been chosen as a tool for finding small-molecule compounds that can interact with the Hub1 binding site.

The typical fingerprint of protein-protein interactions is embodied in flat, featureless surfaces [8] with weak intermolecular contacts spread across them. One of the most reliable strategies in finding small molecules that can interact with protein surface is a rational design of a peptide that can antagonize the target by maintaining the same linear motif. Thanks to such an approach designed peptide should remain selective, which is an important step in further optimization and therapeutics development. Depending on *a priori* structural knowledge of the target interactions, several categories of protein/peptide docking were classified: template-based e.g. GalaxyPepDock [9], PepComposer [10]; local e.g. DynaDock [11], HADDOCK [12]; and global docking e.g. PIPER-FlexPepDock [13], CABS-dock [14], ClusPro PeptiDock [15]. Peptide docking may be assisted by ancillary tools for coarse refinement of peptide-protein models [16] or prediction of the peptide/binding site [17]. A result of this discovery of active compounds is based on virtual screening [18, 19] and/or rational reasoning of experimental data [20].

Another widely used method is fragment-based screening. It is currently one of the most popular approaches in drug discovery [21, 22]. It relies on identifying chemical fragments that have weak affinities towards the biological target and are subsequently modified into more potent and selective molecules. This approach has allowed access to a completely new area of the chemical space, which is the biggest advantage of fragment screening over the high-throughput screening (HTS) that has been widely used over decades and is based on preexisting compound sets. Additionally, fragment-based screening requires smaller libraries compared to HTS, and higher hit rates are observed. Small fragments will be, by definition, weak binders, therefore have to be screened at high concentrations. Hence, choosing an appropriate screening methodology is another important issue. A wide spectrum of different techniques, which were adopted for fragment-based screening is described in the scientific literature, one of which comprises nuclear magnetic resonance techniques [23–25]. In comparison with other screening methodologies, NMR is capable of detecting the binding of small-molecule compounds to macromolecular targets over an extraordinary broad affinity range from covalent to millimolar. A unique feature of NMR is its robust capability to detect weak intermolecular interactions, which are extremely important in fragment-based screening. [26, 27]. With the application of this method, it is possible to have structural information that can identify the binding site or binding modes of the fragments [22, 28].

In this study, we decided to use a workflow based on global docking, all atom reconstructions from the C-alpha representation, and high-resolution protein-peptide refinement with the CABS-dock molecular docking and FlexPepDock high-resolution protein-peptide refinement which are well documented and already tested in cases like GPCR proteins [29]. Docking analysis of almost 200 molecules in the binding pocket of Hub1 has been carried out. Finally, NMR fragment-based screening was used to confirm these findings

Results

Peptide design and analysis of their affinity towards Hub1 binding side

Interface analysis

Snu66 (HINDI domain)/Hub1 interface was analyzed with PDBePISA to identify interactions that are crucial for complex creation. We defined that protein complex has an interface area of 496.1 Å² which includes 66.7% (12 amino acids) of the total residues of 18. Based on amino acid Δ^iG values shown in **Table S1**, we observed that seven amino acids (Thr12, Ile15, Lys18, Leu19, Met21, and Ile24) form a hydrophobic interface with Hub1 surface while only two (Arg16, Pro23) and hydrophilic one. Moreover, Arg16 of Snu66 is involved in an ionic contact with Asp22 of Hub1, while Leu19, Met21, and Ile15 form Van der Waals interactions with Hub1, corresponding to a 63.8% of the total BSA (Buried Surface Area). Due to the lack of structural coordinates in the original crystal structure, Lys22 was excluded from the constraints. Finally, we retained Thr12, Ile15, Arg16, Lys18, Leu19, Met21, and Ile24 for the restrained conformation in the CABS-dock coarse-grained simulations. It is also visible that Leu7 and Ile9 significantly contributed to the binding affinity with a total Gibbs energy solvation of 0.75 kcal/mol and a buried surface area of 115.13 Å².

Peptide design

Based on the Hub1/Snu66 co-crystal structures [39] and our interface analysis we designed four peptides that derived from HIND domain of Snu66 protein and could show affinity towards Hub1 binding side. All of them had a IREKL sequence in common. The first two derivatives contained apart from IREKL sequence **Peptide 1** an addition of two **Peptide 2** and five residues **Peptide 3** on the less structured C-terminal end. The fourth one **Peptide 4** was longer by 3 and 2 amino acids respective on the N- and C- terminal ends (**Fig. 1**).

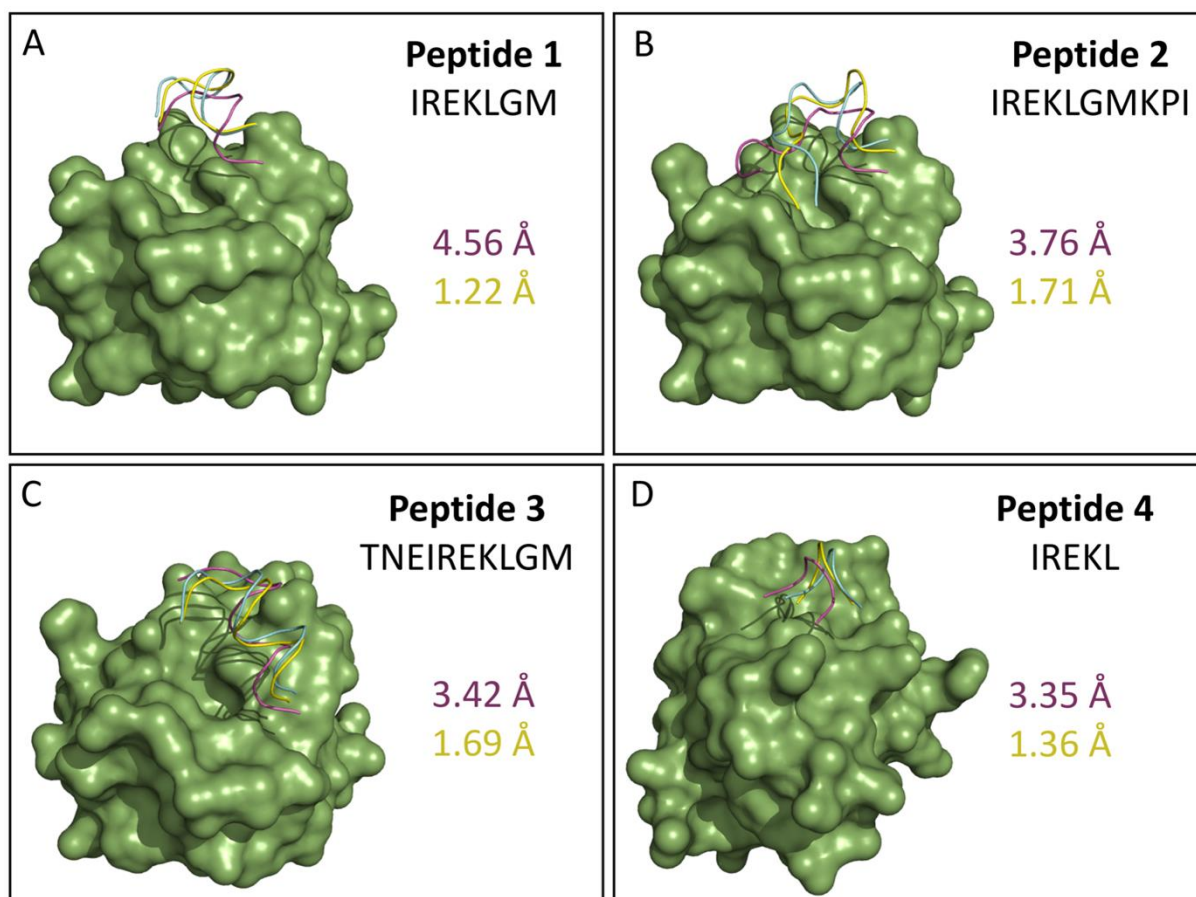


Fig. 1. Docking poses of the peptides **1** (Panel A), **2** (Panel B), **3** (Panel C), and **4** (Panel D) on Hub1 (green surface). The reference crystal structure is represented as cyan ribbons, while the yellow and magenta colors indicate the Rosetta FlexPepDock and the CABS-dock top10 -scored models, respectively. LRMS values are calculated using CA only.

Short peptide modeling

Further investigation was focused on peptide modeling. The main objective of this step was to prove if designed peptides have the potential to bind to Hub1 binding side. The computational settings used for CABS-dock peptide modeling are shown in **Table S2**, while the L-RMS , I-RMS , and Fnat values are presented in **Table S3**, **Table S4**, and **Table S5** respectively. Overall, 57% of the final top10 models resulted in an inadequate accuracy with a CA RMSD (Root Mean Square Deviation) > 5.0 Å (**Table S6**).

Initially, the CAPRI ranking scores were “acceptable” for peptides **1**, **2**, and **3** and “incorrect” for peptide **4**. Nevertheless, further inspection within the collected 10.000 trajectories resulted in better poses with lower RMSD and CABS-dock energy scores (**Fig. S1**). CAPRI ranking scores improved to "medium" for peptides **1**, **2**, **4** and “acceptable” for peptide **3**. With these poses, we then proceeded to the all-atom reconstruction and finally, high-resolution refinement with constraints using FlexPepDock. Among the top 10 poses, we chose

the first, most accurate one based on the CAPRI rank. The refined structures and the Rosetta energy score are presented in **Fig. 1** and **Fig. 2**, respectively.

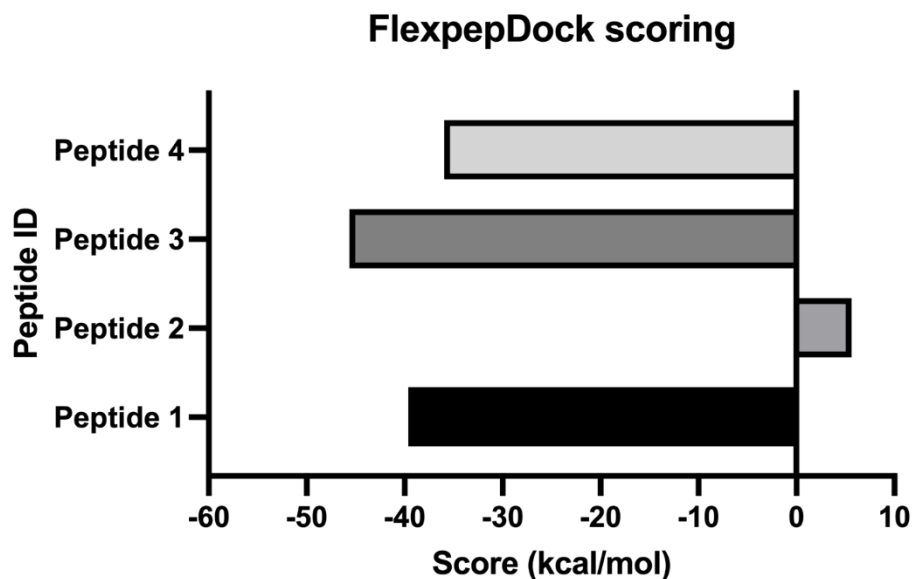


Fig. 2. Bar plot of the Rosetta energy score (x-axis) calculated with FlexPepDock high-resolution refinement of the four peptides (y-axis).

Based on our data, we have observed that peptide **3** (-45.6 kcal/mol) retains the highest score among all peptides, followed by peptide **1** (-39.6 kcal/mol), peptide **4** (-36.0 kcal/mol), and peptide **2** (5.6 kcal/mol). All these poses were ranked as “medium”.

We noticed that peptide **2**, while providing the crystal structure interactions calculated with Scorpion, shows a score with a positive value, most likely due to a clash between the residues not at the interface between Hub1 and Snu66 but within the same peptide. Indeed, peptide **2** elongating toward the C-terminal end forms an antiparallel reentrant loop with the alpha-helix with the possibility of contact. This clash, therefore, could occur during modeling between GMKPI (C-terminal loop) and IREKL (alpha helix) residues, respectively. Despite this, the best-refined score held a value of -1.792 kcal/mol but with a consequent reduction in CAPRI rank from "medium" to "acceptable". Nevertheless, peptide **2** remained the lowest scored peptide among other ones.

We also analyzed the tight network of molecular interactions of peptide **3** with Hub1. As can be seen in **Fig. S2**, the docking pose clearly preserves the ionic bond between Asp22 and Arg16 previously observed with PDBePISA. Other conserved Van Der Waals interactions include one between Ile15 and Leu33 and four between Met21 and Asp26 and Val30.

Finally, additional contacts of the same type were observed between Thr12 and Lys17, Met21 and Lys19, and one hydrogen bond between Arg16 and Cys18.

Peptide NMR analysis

To confirm our theoretical model, we decided to perform NMR-based experiments for peptides **1** and **3** that showed the highest Rosetta score towards the Hub1 binding site. For that purpose, we titrated Hub1 protein with peptides **1** and **3** using a ratio from 1:1 to 1:10 and measured the ^1H NMR spectrum of each solution (**Fig. 3**).

Analysis of superimposed spectra (**Fig. S3**) showed in both cases that there are changes in chemical shifts in low aliphatic ($\delta = 0.04$ ppm) and aromatic ($\delta = 6.5$ ppm) regions. Changes of buffer pH after titration were relatively small being 6.7 and 6.5 for peptide **1** and **3** respectively, which seems to be not relevant and suggest weak interaction between peptides and Hub1 instead.

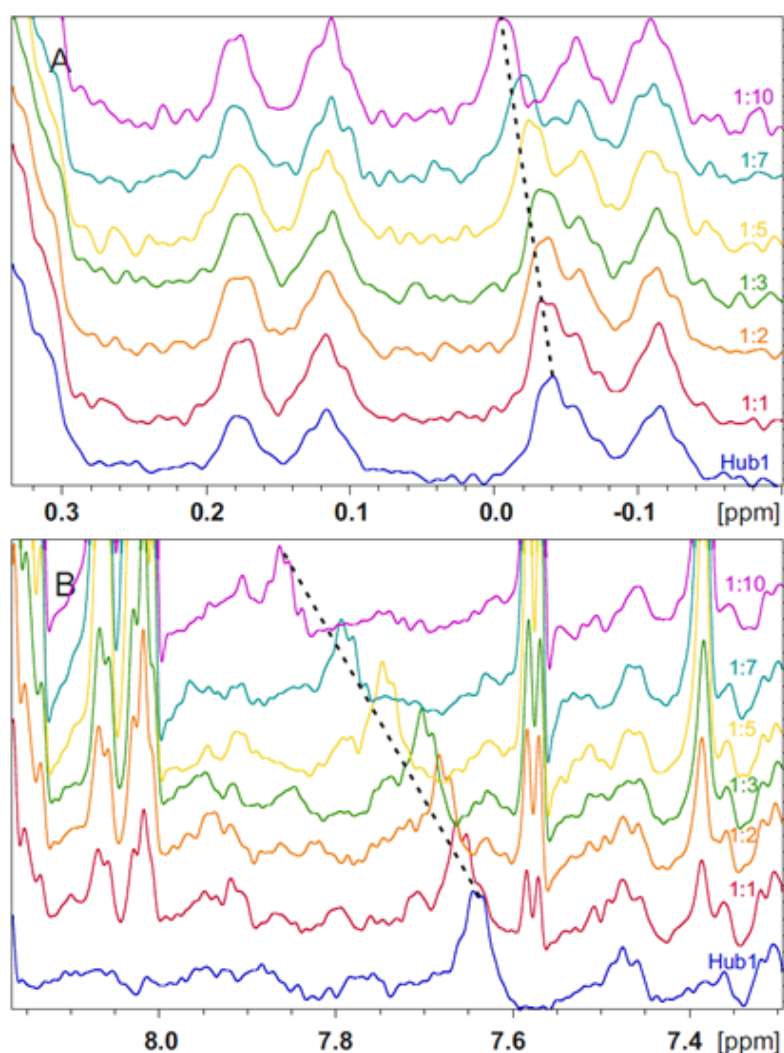


Fig. 3. ^1H NMR titration experiment of Hub1 with peptide **3**. A. Aromatic region of the spectra B. Aliphatic region of the spectra. Color coding: reference Hub1 – blue, 1:1 ratio of Hub1:peptide **3** – red, 1:2 ratio of Hub1:peptide **3** – yellow, 1:3 ratio Hub1:peptide **3** – green, 1:5 ratio Hub1:peptide **3** – orange, 1:7 ratio Hub1:peptide **3** – light blue, 1:10 ratio Hub1:peptide **3** – purple.

Peptide Nano-DSF analysis

To determine if the peptides are interacting with the Hub1 protein and to exclude potential pH changes influence, we performed a nano-DSF experiment. Thanks to that we could see how the presence of the peptides influences the thermal stability of Hub1. Hub1, control of peptide **1** and peptide **3** alone and mixed were measured by Tycho to determine the thermal stability of Hub1 under those conditions. As visible from the first derivative of the intensity profiles (ratio of 330/350nm wavelengths) in **Fig. S4**, Hub1 is characterized by its thermal stability of 64.1°C, while the addition of ca. 100x excess of peptide **1** or peptide **3** decreased Hub1 stability to 52.9 and 57.8°C, respectively. Peptides themselves did not show any higher-order structure in the temperature range between 40 to 95 °C. Described results suggest a weak interaction between the Hub1 protein and both peptides.

Fragment-based screening methods

In silico fragment-based screening of small molecules

Encouraged by our initial findings, we proceeded to the fragment-based screening experiments in an attempt to discover the small-molecule inhibitors of the Hub1 protein. We performed a docking analysis of almost 200 molecules in the binding pocket of Hub1. We have tested the noncommercial library of small-molecular fragments which fulfill the ‘rule of three’: each fragment has a molecular weight lower than 300 Da, partition coefficient (logP) values lower than 3, and each compound contains up to three hydrogen bond acceptors and up to three hydrogen bond donors (**Tables S7-S11**) [40]. One of the used fragments were also single amino acids that were present in peptides **1** and **3** such as leucine, isoleucine, methionine, threonine, and lysine. The compounds were screened using the AutoDock Vina algorithm integrated into the PyRx software. Structures were initially minimized using Open Babel software and docked onto a rigid structure of Hub1 protein from PDB ID: 4PYU. The interactions of the five best picks were assessed using Protein-Ligand Interaction Profiler (PLIP). The best structures (**Fig. 4**) lay within the same cleft on Hub1.

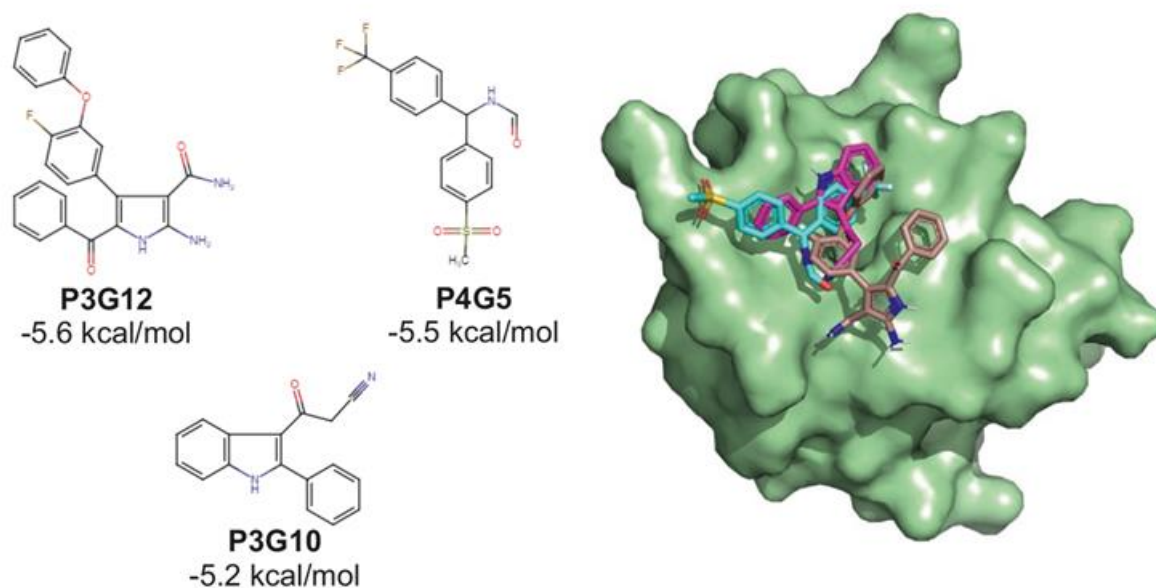


Fig. 4. The *in silico* docking results of the Hub1 protein (PDB ID: 4PYU) presents the top five fragments with the best energy value (Panel A). All structures lay within the same cleft of Hub1 (Panel B). Hub1 is shown as a green surface representation, molecules are shown as a stick representation: **P3G12**– brown, **P4G5**– cyan, **P3G10**– magenta.

We selected two compounds that show the strongest affinity towards Hub1. According to our *in-silico* docking routine both create hydrogen bonds between Lys17 (acceptor) and nitrogen atoms in compounds structure (donor). Additionally, we observed numerous hydrophobic interactions with aliphatic amino acids, mainly with Val30, Leu19, and Leu33. Those interactions remain the same pattern as observed for previously analyzed peptides. **P3G10** lacks a hydrogen bond to Hub1, which explains lower affinity towards targeted protein but its affinity can be explained by multiple hydrophobic interactions with Leu19 and Asp26 (for details see Supplementary Information).

NMR screening of small-molecules

Following the *in silico* analysis, we have performed experimental fragment-based screening of a library of small molecules with the application of NMR techniques. To make the throughput higher, as a first step, the ^1H NMR spectra of the Hub1 protein treated with a 10-fold molar excess of fragments cocktail were recorded. When a perturbation in chemical shifts of the aliphatic region was observed, we examined the possible binding with single compounds using the ^1H - ^{15}N SOFAST-HMQC (Band-Selective Optimized Flip-Angle Short-Transient Heteronuclear Multiple Quantum Correlation) methods. To achieve that, the protein was titrated with a stock solution of small fragments in DMSO (dimethyl sulfoxide) starting from 10-fold

molar excess of the compound. (**Fig. 5, Table S12**). Based on that, 31 fragments were identified as potential hits.

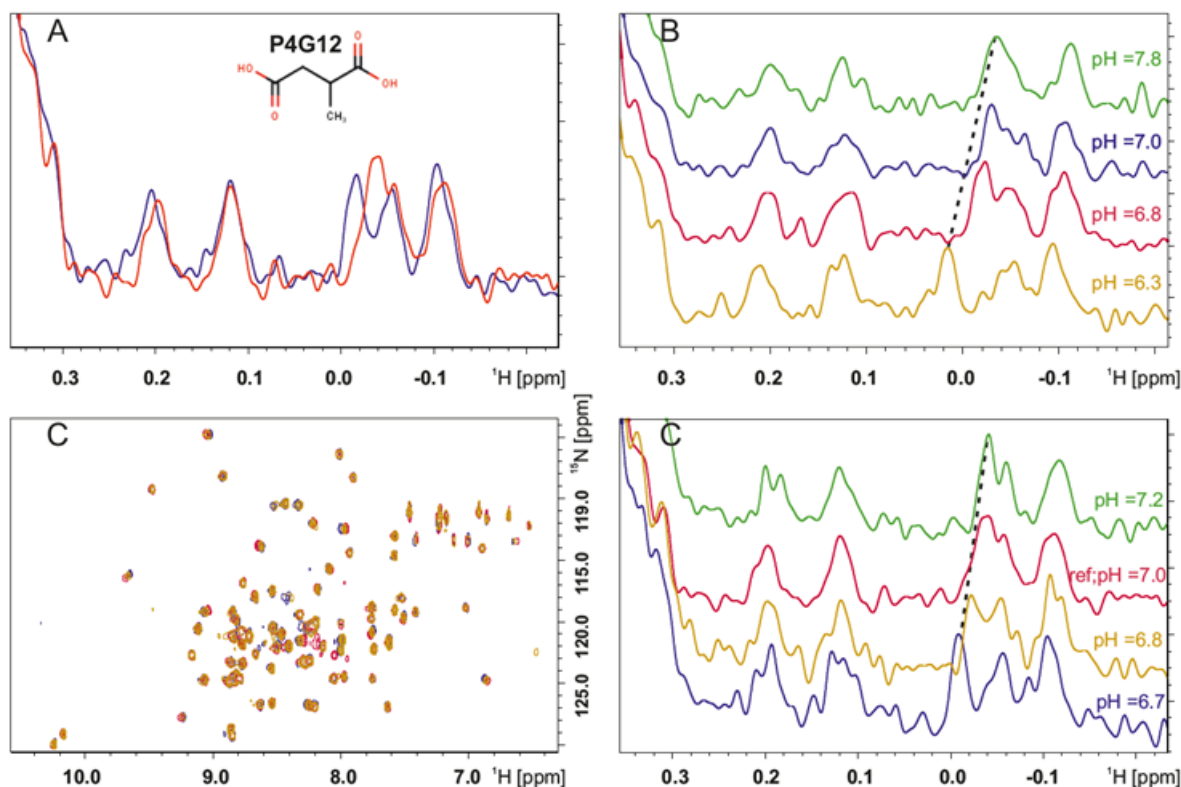


Fig. 5. The NMR experiment of Hub1 protein titrated with small molecular fragments. A. ^1H NMR aliphatic spectrum of reference Hub1 (red) superimposed with the spectrum of Hub1 titrated with 10 fold excess of **P4G12** (blue) indicating a possible interaction of compound **P4G12** with Hub1 protein; B. pH dependence of the aliphatic range of ^1H NMR of Hub1 protein indicating the sensitive nature of Hub1 to even small pH changes: blue – initial pH of the protein in PBS, red – pH =6.8; yellow = 6.3; green = 7.8; C. pH correction of Hub1 with the **P4G12** small fragment. Left corner – ^1H - ^{15}N HMQC spectra; right corner – aliphatic ^1H spectrum, red – reference Hub1, blue – after 1:10 addition of **P4G12**, yellow – pH corrected to 6.8, green – pH corrected to 7.2. Recorded spectra have shown that all changes visible during titration of Hub1 by **P4G12** can be reversed by even small pH changes.

Analysis of pH influence on Hub1 NMR spectrum

Analyzing the hits selected by NMR and comparing them with the compounds identified by *in silico* docking simulation, we noticed that the structures of those compounds are significantly different. The compounds identified in the *in silico* docking simulation possess aromatic rings and amine moieties, whereas the fragments selected in the NMR experiment have carboxylic groups that seem to be crucial to create interaction. Although the results are not as expected from the *in silico* screening, the difference may be explained by the conditions of those two experiments. Docking simulation has used a rigid crystal structure model, whereas NMR

screening is based on the flexible character of the protein in the solution. The similarity of the compound structures selected by NMR screening, especially their acidic characteristics, prompted us to analyze the possibility of the pH influence on the Hub1 protein during titration. The changes in pH during titration were relatively small, similar to those for the tested peptides, being in the range of approx. 0.3-0.4 lower than the initial pH. Such differences typically should not affect chemical shifts in the aliphatic region of the NMR spectrum and all perturbations should result from binding interactions. To confirm that thesis, we adjusted the pH of the Hub1 solution to represent the experiment conditions without the added compounds and it turned out that this protein is highly sensitive to pH changes (**Fig. 5B**). Such sensitivity does not indicate that the obtained results are caused solely by pH changes, as the binding may be independent of it. To investigate what exactly causes perturbations in chemical shifts, we adjusted the pH of the protein titrated with small molecular fragments to the initial one (**P4G9** – 2-hydroxy-2-phenylacetic acid; pH of 6.8 after 1:10 ligand addition, corrected to pH of 7.2; **P4G12** – 2-methylsuccinic acid; pH of 6.7 after 1:10 ligand addition, corrected to pH of 7.2; **Fig. 5C, Fig. S5**). Unfortunately, we observed reversing of changes in spectra during the pH correction, which brought us to the conclusion that disruptions in chemical shifts may be related to the acidic character of the compounds, instead of their interaction with protein

Discussion

To identify small molecule ligands for Hub1, we have used several different methods: interface analysis, peptide modeling, nano-DSF analysis, *in silico* screening, and NMR screening. As a first step, we decided to perform an interface analysis of the interaction between Hub1 and the HIND domain of the Snu66 protein to understand better how to design ligands that would interact with the Snu66 binding site of Hub1. Two peptides were selected and their affinity towards Hub1 was tested via NMR and nano-DSF analysis. Observed changes in the NMR spectrum are not globally abundant over the spectrum. That can indicate weak ionic interaction between the peptide and aspartic acid in the protein binding pocket or maybe result of lack of peptides secondary structure, which would be also visible on the NMR spectra (**Fig. S6**). Similarly results obtained from nano-DSF show weak interaction between the protein and the tested peptide.

Having that knowledge, we extended our research to find small-molecule non-peptide fragments that bind to the Hub1 protein. We performed *in silico* and NMR ^1H - ^{15}N SOFAST-HMQC titration screening of fragments library. Our results showed that all experimentally

selected “hits” had an acidic character and perturbations in chemical shifts were shown to most likely be caused by the pH change. We further confirmed that the Hub1 protein is sensitive to small pH changes, thus perturbations in NMR signals in the aliphatic region that reflects mostly protein side-chain methyl groups may be the result of pH changes but the weak binding of the fragments to Hub1 signals cannot be excluded. The situation we observed although uncommon, is possible due to the specific character of the protein binding pocket. The peaks in the low aliphatic region of the spectra originate from the methyl side-chains of the leucine, isoleucine, alanine, or other aliphatic residues, therefore, should be resistant to pH changes of the solution. However, it may happen that the proximity of the pH-sensitive amino acid such as arginine, could influence changes even in the low aliphatic region. The Hub1 binding pocket is widely occupied by pH-sensitive residues (especially Lys29, Asp22, Asp26) which lies close to the aliphatic amino acids (Leu33 or Val30). Taking this possibility into account the NMR methodology has proven to be insufficient for the fragment base screening toward the Hub1 protein. Although the positive hits in NMR experiments for this protein need to be verified by another method, this method can be used to exclude compounds that are not interacting with the protein surface. If the changes in the protein spectra are not observed over titration with ligands the corresponding compounds are not interacting with protein. We proved that NMR cannot be the only method used to indicate hits in this particular case, but might be useful for excluding fragments that are not binding to Hub1 protein.

Our study led to the discovery of two peptides that showed weak affinity towards the Hub1 binding side which are the first reported small-molecule ligands that interact with Hub1 protein. Although those interactions are too weak to be a base for drug design, we recognize this as a milestone in further optimization and new ligands discovery.

Methods

Protein purification

Hub1 (residues 1-73) was expressed in *E.coli* BL21 (DE3) (New England Biolabs) using the pET28a vector. The protein was either grown in LB or M9 minimal medium at 37 °C and was subsequently induced with 1 M IPTG (isopropyl β -D-1-thiogalactopyranoside) at OD₆₀₀. The bacterial culture was grown overnight at 25 °C. Cells were then harvested by centrifugation (10 min, 5000 x g); the obtained pellet was then dissolved in lysis buffer containing 50 mM NaH₂PO₄, 30 mM NaCl, 10 mM imidazole at pH 8.0 and was subsequently disrupted by sonication (5 times 2 min using a macrotip, output control 8, 80%). The disrupted material was then centrifuged (15 min, 15000 x g) and applied to Ni-NTA (nickel-nitrilotriacetic acid) resin

previously equilibrated with lysis buffer; the column was then washed with buffer containing 50 mM NaH_2PO_4 , 300 mM NaCl, 20 mM imidazole at pH 8.0, and the protein were eluted with buffer containing 50 mM NaH_2PO_4 , 300 mM NaCl, 250 mM imidazole at pH 8.0. The last purification step consisted of size exclusion chromatography (SEC) using a Superdex 75 column (GE Healthcare, Akta Pure) using PBS buffer at pH 7.4.

Nuclear magnetic resonance

NMR measurements were carried out at 300 K on an ultrashielded 600 MHz Bruker AVANCE 3 spectrometer equipped with a liquid nitrogen cryogenic system. Uniform ^{15}N isotope labeling was achieved by the expression of the protein in the M9 minimal media containing $^{15}\text{NH}_4\text{Cl}$ as the sole nitrogen source. The NMR buffer for ^{15}N -labeled protein was PBS pH 7.4 with 5 mM DTT (dithiothreitol). To provide a lock signal, 10% (v/v) of D_2O was added to the samples. Water suppression was carried out using the WATERGATE sequence [30]. Stock solutions of inhibitors used for titration were prepared in DMSO-d_6 . The binding of small molecule fragments towards Hub1 was measured with the 2D SOFAST-HMQC spectra [31] of the ^{15}N -labeled Hub1 recombinant protein in 3-mm NMR capillaries. For the preliminary screening, the mixture of three compounds was added in molar ratio compound to protein 1:10. Then the ^1H and SOFAST-HMQC spectrum was recorded and compared with the reference spectrum of protein Hub1 titrated with a suitable amount of DMSO. When any activity occurred, all compounds which were in the mixture were tested one by one. The individual spectra were recorded in 10-fold excess of inhibitor either by ^1H NMR or SOFAST-HMQC. The ^1H spectra of Hub1 with peptides 1 and 3 (10 mM solution in water) were measured during a 9-step titration routine; SOFAST-HMQC were recorded for 10-fold excess of the peptide. The spectra were processed with TopSpin 3.2 software.

Peptide docking

The apo and bound forms of Hub-1 in complex with Snu66 peptide (HINDI) were fetched from the PDB under the accession codes of 1M94 and 3PLU. These structures were treated as receptors and references for peptide docking. The analysis of the interface was performed with PDBePISA ((Proteins, Interfaces, Structures, and Assemblies) server. [32]) Peptide docking was performed with CABS-dock [14] using DSSP [33, 34] to assign the secondary structure. A licensed version of Modeller 10.1 was employed for optimization of the receptor structure within the top10 models [35]. In addition, the ranking of the docking poses was assigned based on CAPRI (Critical Assessment of Prediction of Interactions) criteria for protein-peptide

systems [36]. These criteria consist of the measurement of IRMS (interface root mean square deviation), LRMS (ligand root mean square deviation) and Fnat (fraction of native contacts). On one hand, the first and second parameters measure the accuracy of backbone atoms and interface residues respectively. On the other hand, the third parameter consists in the number of predicted contacts divided by the number of native contacts. The models from the CABS-dock docking with the lowest RMSD to Snu66 were chosen and the all-atom representation was performed with ca2all.py script. The resulting models were refined and scored with the FlexPepDock server using 100 high-resolution structures for each peptide [16, 37]. The top 10 poses were retained and visually inspected. Finally, the molecular interactions between the refined pose and the receptor were generated with Scorpion (DesertSci Ltd), which identifies and rank hot spot interactions based on the concept of cooperativity binding network [38]. All figures were rendered with Pymol (The PyMOL Molecular Graphics System, Version 2.3.4, Schrödinger, LLC).

Nano DSF

The protein sample was mixed with the peptides and allowed for 15 min incubation at RT before measuring with Tycho (Nanotemper). Hub1 concentration was kept constant at 240 μ M, peptide 1: 25 mM, peptide 3: 19.5 mM.

Declarations

Funding This work was supported by the Preludium grant UMO-2015/17/N/ST5/01942 (for K. K.) founded by the National Science Centre, Poland. This project is funded from the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Skłodowska-Curie Grant Agreement No. 675555, Accelerated Early-stage drug discovery (AEGIS) (for A.R.R.)

References

1. Friedman JS, Koop BF, Raymond V, Walter MA (2001) Isolation of a Ubiquitin-like (UBL5) Gene from a Screen Identifying Highly Expressed and Conserved Iris Genes. *Genomics* 255:252–255. <https://doi.org/10.1006/geno.2000.6439>
2. Hochstrasser M (2009) Origin and function of ubiquitin-like proteins. *Nature* 458:422–429. <https://doi.org/10.1038/nature07958>

3. Jentsch S, Pyrowolakis G (2000) Ubiquitin and its kin: how close are the family ties ? *Trends Cell Biol* 10:335–342
4. McNally T, Huang Q, Janis RS, et al (2003) Structural analysis of UBL5, a novel ubiquitin-like modifier. *Protein Sci* 12:1562–1566. <https://doi.org/10.1110/ps.0382803.Beacon>
5. Masopust M, Weisz F, Bartenschlager H, et al (2014) Porcine ubiquitin-like 5 (UBL5) gene: genomic organization, polymorphisms, mRNA cloning, splicing variants and association study. *Mol Biol Rep* 41:2353–2362. <https://doi.org/10.1007/s11033-014-3089-6>
6. Bedford L, Lowe J, Dick LR, et al (2011) Ubiquitin-like protein conjugation and the ubiquitin – proteasome system as drug targets. *Nat Rev Drug Discov* 10:29–46. <https://doi.org/10.1038/nrd3321>
7. Jones S, Thornton JM (1996) Review Principles of protein-protein interactions. 93:13–20
8. Wells JA, McClendon CL (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* 450:1001–1009. <https://doi.org/10.1038/nature06526>
9. Lee H, Heo L, Lee MS, Seok C (2015) GalaxyPepDock: A protein-peptide docking tool based on interaction similarity and energy optimization. *Nucleic Acids Res* 43:W431–W435. <https://doi.org/10.1093/nar/gkv495>
10. Obarska-Kosinska A, Iacoangeli A, Lepore R, Tramontano A (2016) PepComposer: computational design of peptides binding to a given protein surface. *Nucleic Acids Res* 44:W522–W528. <https://doi.org/10.1093/nar/gkw366>
11. Antes I (2010) DynaDock: A new molecular dynamics-based algorithm for protein-peptide docking including receptor flexibility. *Proteins Struct Funct Bioinforma* 78:1084–1104. <https://doi.org/10.1002/prot.22629>
12. Dominguez C, Boelens R, Bonvin AMJJ (2003) HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* 125:1731–1737. <https://doi.org/10.1021/ja026939x>
13. Alam N, Goldstein O, Xia B, et al (2017) High-resolution global peptide-protein docking using fragments-based PIPER-FlexPepDock. *PLoS Comput Biol* 13:1–20.

- <https://doi.org/10.1371/journal.pcbi.1005905>
14. Kurcinski M, Pawel Ciemny M, Oleniecki T, et al (2019) CABS-dock standalone: A toolbox for flexible protein-peptide docking. *Bioinformatics* 35:4170–4172. <https://doi.org/10.1093/bioinformatics/btz185>
 15. Porter KA, Xia B, Beglov D, et al (2017) ClusPro PeptiDock: Efficient global docking of peptide recognition motifs using FFT. *Bioinformatics* 33:3299–3301. <https://doi.org/10.1093/bioinformatics/btx216>
 16. Raveh B, London N, Schueler-Furman O (2010) Sub-angstrom modeling of complexes between flexible peptides and globular proteins. *Proteins Struct Funct Bioinforma* 78:2029–2040. <https://doi.org/10.1002/prot.22716>
 17. Lavi A, Ngan CH, Movshovitz-Attias D, et al (2013) Detection of peptide-binding sites on protein surfaces: The first step toward the modeling and targeting of peptide-mediated interactions. *Proteins Struct Funct Bioinforma* 81:2096–2105. <https://doi.org/10.1002/prot.24422>
 18. Krüger DM, Glas A, Bier D, et al (2017) Structure-Based Design of Non-natural Macrocyclic Peptides That Inhibit Protein-Protein Interactions. *J Med Chem* 60:8982–8988. <https://doi.org/10.1021/acs.jmedchem.7b01221>
 19. Leffler AE, Kuryatov A, Zebroski HA, et al (2017) Discovery of peptide ligands through docking and virtual screening at nicotinic acetylcholine receptor homology models. *Proc Natl Acad Sci U S A* 114:E8100–E8109. <https://doi.org/10.1073/pnas.1703952114>
 20. Kar RK, Gazova Z, Bednarikova Z, et al (2016) Evidence for Inhibition of Lysozyme Amyloid Fibrillization by Peptide Fragments from Human Lysozyme: A Combined Spectroscopy, Microscopy, and Docking Study. *Biomacromolecules* 17:1998–2009. <https://doi.org/10.1021/acs.biomac.6b00165>
 21. Mashalidis EH, Paweł S, Lang S, Abell C (2013) A three-stage biophysical screening cascade for fragment-based drug discovery. *Nat Protoc* 8:2309–2324. <https://doi.org/10.1038/nprot.2013.130>
 22. Hajduk PJ, Greer J (2007) A decade of fragment-based drug design: Strategic advances and lessons learned. *Nat Rev Drug Discov* 6:211–219. <https://doi.org/10.1038/nrd2220>
 23. Fielding L (2007) NMR methods for the determination of protein-ligand dissociation

- constants. *Prog Nucl Magn Reson Spectrosc* 51:219–242. <https://doi.org/10.1016/j.pnmrs.2007.04.001>
24. Williamson MP (2013) Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc* 73:1–16. <https://doi.org/10.1016/j.pnmrs.2013.02.001>
 25. Harner MJ, Frank AO, Fesik SW (2013) Fragment-based drug discovery using NMR spectroscopy. *J Biomol NMR* 56:65–75. <https://doi.org/10.1007/s10858-013-9740-z>
 26. Dalvit C, Fogliatto G, Stewart A, et al (2001) WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability. *J Biomol NMR* 21:349–359
 27. Babaoglu K, Shoichet BK (2006) Deconstructing fragment-based inhibitor discovery. *Nat Chem Biol* 2:720–723. <https://doi.org/10.1038/nchembio831>
 28. Shuker SB., Hajduk PJ., Meadows RP., Fesik SW (1996) Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* (80-) 274:1531–1534
 29. Badaczewska-Dawid AE, Kmiecik S, Koliński M (2021) Docking of peptides to GPCRs using a combination of CABS-dock with FlexPepDock refinement. *Brief Bioinform* 22:1–9. <https://doi.org/10.1093/bib/bbaa109>
 30. Piotto M, Saudek V, Sklenar V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* 2:661–665. <https://doi.org/10.1007/BF02192855>
 31. Brutscher B, Schanda P (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* 33:199–211. <https://doi.org/10.1007/s10858-005-4425-x>
 32. Krissinel E, Henrick K (2007) Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol* 372:774–797. <https://doi.org/10.1016/j.jmb.2007.05.022>
 33. Touw WG, Baakman C, Black J, et al (2015) A series of PDB-related databanks for everyday needs. *Nucleic Acids Res* 43:D364–D368. <https://doi.org/10.1093/nar/gku1028>
 34. Kabsch W, Sander C (1983) Routine follow up of patients with treated pulmonary tuberculosis. *Biopolymers* 22:2577–2637. <https://doi.org/10.1002/bip.360221211>

35. Šali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779–815
36. Lensink MF, Velankar S, Wodak SJ (2017) Modeling protein–protein and protein–peptide complexes: CAPRI 6th edition. *Proteins Struct Funct Bioinforma* 85:359–377. <https://doi.org/10.1002/prot.25215>
37. London N, Raveh B, Cohen E, et al (2011) Rosetta FlexPepDock web server - High resolution modeling of peptide-protein interactions. *Nucleic Acids Res* 39:249–253. <https://doi.org/10.1093/nar/gkr431>
38. Kuhn B, Fuchs JE, Reutlinger M, et al (2011) Rationalizing tight ligand binding through cooperative interaction networks. *J Chem Inf Model* 51:3180–3198. <https://doi.org/10.1021/ci200319e>
39. Mishra SK, Ammon T, Popowicz GM, et al (2011) Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. *Nature* 474:173–180. <https://doi.org/10.1038/nature10143>
40. Congreve M (2003) A ‘ Rule of Three ’ for fragment-based lead. *Drug Discov Today* 8:876–877