



Review

Virtual Screening of Peptide Libraries: The Search for Peptide-Based Therapeutics Using Computational Tools

Marian Vincenzi [†], Flavia Anna Mercurio [†]  and Marilisa Leone ^{*} 

Institute of Biostructures and Bioimaging, Via Pietro Castellino 111, 80131 Naples, Italy; marian.vincenzi@ibb.cnr.it (M.V.); flaviaanna.mercurio@cnr.it (F.A.M.)

^{*} Correspondence: marilisa.leone@cnr.it

[†] These authors contributed equally to the study.

Abstract: Over the last few decades, we have witnessed growing interest from both academic and industrial laboratories in peptides as possible therapeutics. Bioactive peptides have a high potential to treat various diseases with specificity and biological safety. Compared to small molecules, peptides represent better candidates as inhibitors (or general modulators) of key protein–protein interactions. In fact, undruggable proteins containing large and smooth surfaces can be more easily targeted with the conformational plasticity of peptides. The discovery of bioactive peptides, working against disease-relevant protein targets, generally requires the high-throughput screening of large libraries, and *in silico* approaches are highly exploited for their low-cost incidence and efficiency. The present review reports on the potential challenges linked to the employment of peptides as therapeutics and describes computational approaches, mainly structure-based virtual screening (SBVS), to support the identification of novel peptides for therapeutic implementations. Cutting-edge SBVS strategies are reviewed along with examples of applications focused on diverse classes of bioactive peptides (i.e., anticancer, antimicrobial/antiviral peptides, peptides blocking amyloid fiber formation).

Keywords: bioactive peptides; PPIs; drug discovery; virtual screening; anticancer peptides; antiviral peptides



Citation: Vincenzi, M.; Mercurio, F.A.; Leone, M. Virtual Screening of Peptide Libraries: The Search for Peptide-Based Therapeutics Using Computational Tools. *Int. J. Mol. Sci.* **2024**, *25*, 1798. <https://doi.org/10.3390/ijms25031798>

Academic Editor: Jesus Vicente De Julián Ortiz

Received: 22 December 2023

Revised: 26 January 2024

Accepted: 30 January 2024

Published: 1 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Protein functions are primarily dictated by interactions with other proteins or biomolecules [1]. Interestingly, an ensemble of hundreds of thousands of protein–protein interactions (PPIs) composes the human interactome, and the aberrant modulation, in particular inhibition, of key PPIs within this ensemble can be associated with different diseases, ranging from cancer to neurodegenerative pathologies [2,3]. Since large regions mediate PPIs, and they are usually not provided with well-defined pockets, grooves or clefts, their targeting by small molecules represents a very challenging drug development route [2,3]. Both natural and synthetic peptides represent valuable alternatives to small molecules in the development of therapeutic agents able to inhibit PPIs [2]. In fact, due to the different chemical–physical properties with respect to small molecules, peptides can better adapt to the large interaction surfaces of proteins [2–4]. In addition, the possibility of improved peptide ADME (i.e., absorption, distribution, metabolism, excretion) profiles represents a further justification for the interest of drug development research in peptides targeting PPIs [2,3]. Peptides with a small molecular weight, high flexibility, and minimal toxicity can potentially constitute a new class of biopharmaceuticals [3]. On the other hand, peptide-based drug development, which requires analysis by high-throughput (HT) approaches of peptide–protein interactions, represents an expensive and time-consuming process that can benefit from *in silico* methods [1]. Therefore, this review is focused on the cutting-edge *in silico* approaches that could be implemented to discover peptide modulators of PPIs with potential therapeutic effects. The applications of computational methods centered around diverse classes of peptides (including anticancer, antiviral, self-assembling peptides) are also discussed.

1.1. Targeting PPIs with Peptides

The availability of 3D protein structures, principally protein–protein complexes, is crucial to developing peptides targeting PPIs [1,2]. In this framework, Nuclear Magnetic Resonance (NMR), X-ray crystallography, and cryo-Electron Microscopy (cryo-EM) are the most frequently applied techniques to obtain the structures of single interactors or their complexes [1,2]. More precisely, X-ray crystallography is rather suitable for retrieving data for large globular domains (alone or associated with each other) whereas, NMR can also be exploited to study proteins involved in the formation of transient and weak complexes and obtain dynamic information [1,2]. Moreover, structures with near-atomic resolutions can be determined by combining low-resolution structural techniques [e.g., cryo-Electron Microscopy (cryo-EM) and Small-Angle X-ray Scattering (SAXS)] with X-ray crystallography, NMR, Förster Resonance Energy Transfer (FRET), or Mass Spectroscopy (MS) techniques, along with *in silico* approaches [2]. The reference database for 3D structures obtained by all these techniques is the Protein Data Bank (PDB) [3]. With a 3D structure in hand, a detailed analysis of the structural features characterizing the protein interfaces can be carried out, and the gained structural insights can be employed for the design of peptides able to interfere with the PPI under analysis. In the peptide drug design field, *in silico* approaches to predicting protein–peptide complexes offer some advantages if compared to experimental methods, although they still present diverse challenges to overcome [1]. A significant issue, which should be taken into account when trying to computationally model protein–peptide interactions, relies on the variety of conformations that can be explored by both proteins and peptides following mutual binding (the so-called induced fit) and the difficulties in properly ranking the *in silico* predicted solutions for a specific protein–peptide complex [1]. In general, it is not easy to guess *in silico* which protein fragment will bind a protein partner and establish which conformation the selected peptide region will exactly assume upon complex formation [1].

However, the combination of *in silico* and experimental methodologies is the most common strategy in drug discovery as it can exploit an initial generation of hypotheses by computational tools, thus greatly speeding up the overall peptide identification process and reducing monetary costs. *In silico* strategies work better when at least the 3D structures of isolated target proteins are available and have been experimentally determined, although, recently, the AlphaFold2 approach has tremendously revolutionized the world of molecular modeling by improving prediction of 3D structures [5–7]. Nevertheless, recent advances have been made for the *de novo* design of new proteins through deep learning approaches, including denoising diffusion probabilistic models (DDPMs) [8]. For example, the recently described “RoseTTAFold diffusion” (RFdiffusion) relies on a generative model of protein backbones. It can be very powerful for the design of protein monomers, protein interactors, oligomers with precise symmetries, and for scaffolding enzyme active pockets and symmetric motifs when trying to build therapeutic proteins or those able to bind metals [8].

When the identification of bioactive peptides targeting certain PPIs is negatively affected by the unavailability of 3D structures, phenotype- and target-oriented approaches often constitute valuable solutions [2]. The phenotype-oriented approach [9] may include screening natural peptides or synthetic libraries, the members of which are synthesized on a solid support. This approach offers the chance to simultaneously evaluate a considerable number of peptide variants. The phage display technology, that relies on screening peptides located on the surface of filamentous bacteriophages expressing them, represents another alternative to identify bioactive peptides [2,4,10]. Regarding target-oriented approaches, PEPscan is a valuable method to identify peptide ligands by scanning consecutive and overlapped peptide fragments that derive from one of the two proteins participating in the PPI of interest [2]. PEPscan relies on peptide arrays that can be obtained through the SPOT technique and offers access to a wide variety of peptide sizes (generally, lengths range from 5 to 30 amino acids) and formats [2,11].

The pepATTRACT webserver can be a valuable *in silico* tool that resembles the experimental PEPscan approach. However, pepATTRACT requires knowledge of the 3D

structure of the target protein to be used, along with peptide sequences, as input to perform blind docking runs [12].

Another crucial point when searching for novel peptide modulators of PPIs is the biophysical characterization of peptide–protein and protein–protein interactions, including the determination of binding affinities and stoichiometries. To this aim, many biophysical techniques can be implemented, but, among the most widely used, there are Isothermal Titration Calorimetry (ITC) [13], Surface Plasmon resonance (SPR) [14], BioLayer Interferometry (BLI) [15], and MicroScale Thermophoresis (MST) [16]. For weak binding and under precise conditions, NMR techniques can be supportive to analyze protein–ligand complexes better and can even be used to estimate the dissociation constants [17–21].

ITC is very useful for determining binding stoichiometry, equilibrium constants, and the variation in enthalpy (ΔH) and entropy (ΔS) associated with molecular recognition, whereas the rates of association (k_{on}) and dissociation (k_{off}) characterizing PPIs can be determined by SPR [2].

Once a peptide modulator of a certain PPI is identified, it is important to chemically optimize it. The first crucial step consists of the improvement of the interaction affinity. The peptide segments extracted from the whole protein structure organization tend to be disordered in solution but might assume a more ordered bioactive conformation when in complex with the interaction partner [22]. Thus, forming a protein–peptide complex can often be associated with a high entropic cost due to the absence of a stable conformation in the peptide-unbound state and the rise of a more constrained structure in the peptide-bound form. Therefore, strategies for optimizing peptides targeting PPIs should include those modifications that can induce lower entropy costs and/or higher enthalpic energies of binding. To reach such a goal, it can be useful to constrain the peptide architecture and eventually favor the rise in the unbound form of certain secondary structure elements, for example, through cyclization, which could be achieved by the formation of amide bonds, disulfide bridges, or the generation of stapled peptides [2,23–26]. Nevertheless, as introduced before, one of the difficulties of targeting PPIs is the extension of the interacting surfaces, and macrocyclization is described as a promising solution to make peptides better suited to adapt to large surfaces and consequently work efficiently as PPI inhibitors [24]. Nevertheless, chemical modifications can also be employed to improve the pharmacological properties of model peptides [2]. In fact, during peptide optimization, particular attention must be given to peptide cell permeability and resistance to protease degradation, and many medicinal chemistry approaches have already been proposed to improve such properties (see below) [27–33].

The following section will describe the different features characterizing peptides as therapeutics and strategies to improve peptide drug-like characteristics.

1.2. Brief Overview of Peptide Therapeutic Potential: Major Drawbacks and Optimization Routes

The use of peptides to develop original therapeutics is supported by their potentially desirable features, including high binding affinities and large selectivity, that can be connected to the relatively big sizes with respect to small compounds and, consequently, the capacity to establish an increased number of intermolecular interactions with the target proteins [34,35].

The hormone insulin is a peptide made up of 51 amino acids. The extraction of insulin from the pancreas of animals and its first employment in the treatment of diabetes back in 1920 represents a milestone in the field of peptide-based therapeutics [34–36]. Another example of a therapeutic peptide isolated from natural sources is the adrenocorticotrophic hormone (ACTH) that is effective in treating different pathologies of the endocrine system [34]. Interestingly, bioactive peptides have also been identified in exotic reservoirs, such as venoms from arthropods and cephalopods [34].

Over the years, tremendous improvements have been achieved concerning peptide synthesis methodologies and automation, favoring research in the field of peptide therapeutics. Peptides started to attract attention in 1963 after Merrifield introduced solid-phase

peptide synthesis (SPPS) [33,35,36]. SPPS laid the foundation for the development of an automatic peptide synthesizer by joining both amino acid coupling and deprotection in a single reactor, thus enormously simplifying the production of diverse peptide sequences [33]. Later in the 1980s, the generation of larger pure peptides became possible thanks to the emergence of recombinant technology [36]. Another leap forward in the field of therapeutic peptides was reached with the development of strategies to increase molecular weight by the conjugation of peptides to lipids, bigger proteins, and polyethylene glycol. Such conjugated systems paved the way to enhanced renal clearance and plasma circulation times with respect to unmodified peptides that are instead quickly (i.e., within minutes) cleared from plasma [36].

Later, the progress of combinatorial synthetic libraries and the introduction of in vitro display technologies (e.g., phage display, yeast display, mRNA display, ribosome display, and DNA display) drove the identification of original peptides for a larger set of targets [35]. Intriguingly, display techniques create a sort of bridge between the phenotype (i.e., peptide) and genotype (RNA/DNA coding sequence), thus allowing peptide selection based on affinity while genetic material can be recovered and amplified through PCR (Polymerase Chain Reaction) or upon infection of host cells by phage virions [36].

As regards peptides from the natural reservoir (also named “native” peptides), a major issue that can be encountered in employing them as therapeutic agents is the disadvantageous absorption, distribution, metabolism, and excretion (ADME) profile [34,35]. Indeed, peptides generally do not respect the so-called rule of 5 established for small molecules by Lipinski (i.e., molecular weight < 500 Da, number of H-bond donors ≤ 5 , number of H-bond acceptors ≤ 10 , and a partition coefficient $\log P \leq 5$) and thus are not likely to passively cross membranes and show good oral bioavailability [35]. Diverse studies have demonstrated how the characteristic peptide flexibility could contribute to low passive permeability, thus suggesting the introduction of structural constraints to improve this unfavorable peptide property [35]. Among the solutions proposed to decrease flexibility, the introduction of modified residues (e.g., α -methyl, α,β -dehydro, and β -substituted amino acids) and the substitution of amide bonds with isosteres (e.g., CH_2O , CH_2CH_2 , CH_2NH , CSNH , CH_2S) represent only a few examples [37]. Cyclization is another strategy proposed to favor a lowering of the peptide flexibility and is therefore indicated as a promising tool to enhance peptide permeability [10,35,37,38].

A further tactic that could increase peptide permeability consists of the N-methylation of backbone nitrogen atoms, which favors an increase in the *cis* peptide bond population by decreasing the energy difference between the *cis* and *trans* configurations. Furthermore, this modification also reduces the number of NH backbone groups available to form H-bonds and, consequently, the energy loss related to the desolvation needed for membrane crossover. The capacity of peptides to pass through biological membranes can be improved by conjugation to the so-called cell-penetrating peptides (CPPs), which are generally made up of 5–30 residues and characterized by the presence of basic residues (i.e., arginines and lysines) [35].

Another issue related to peptide therapeutics is the poor resistance to protease digestion. In this context, one of the proposed strategies to increase peptide stability consists of the substitution of L-amino acids, which characterize “native” peptides and can be recognized by proteases with non-natural D-amino acids [10,33,35]. The N-methylation, as well as other N-alkylation modifications, by favoring the *cis* configuration of peptide bonds, can not only improve permeability, as mentioned above, but also stability. The β -amino acids (i.e., β^2 and β^3) and γ -amino acids possess additional carbon atoms in the backbone and are proposed as a possible route to make peptides more resistant to protease attack [33,35]. The replacement of peptide bonds with isosteres (e.g., azapeptides and peptoids) represents a promising route to mimic the transition state of peptide bond cleavage, and thus, can be considered a way to inhibit proteases [2,35]. Additional modifications are recognized as stability enhancers and include the capping of the N-terminus, deamination, the extension of the N- and C-extremities, and the insertion of moieties able to simulate

disulfide bonds [36]. The peptide's resistance to protease and its half-life can be improved as well by linkage to elements providing a sort of shielding effect for enzymatic cleavage, such as polyethylene glycol and proteins like the Fc (Fragment crystallizable) domains [35].

Similarly, conjugation to the proopiomelanocortin-derived peptide sequence (i.e., NSSSSGSSGAGQ) and lipids (e.g., C14/16/18 fatty acids) has been proven to be effective in improving the plasma stability [33–35,39].

As briefly mentioned before, cyclization is a common strategy to improve peptide stabilities by inducing the formation of secondary structure elements [33,35,38,40]. The stapling of peptides is a widely employed approach that can not only be implemented to enhance peptide resistance to proteases but also cell permeability and target affinity through the introduction of covalent crosslinks that can favor the rise of helical structures [33,36]. Covalent crosslinks can be achieved through a variety of approaches, for example, by the formation of lactam bridges between the side chains of acidic residues (i.e., aspartic acid or glutamic acid) and lysines or by replacing certain residues with cysteines or homocysteines, that can be exploited to obtain disulfide bonds, whereas, several biselectrophilic linkers are available and can be used to introduce a clip by reacting with cysteine residues on a model peptide [33]. In addition, hydrocarbon-stapled peptides can be achieved by an olefinic linkage, obtained through click chemistry, between the side chains of α,α -disubstituted non-natural amino acids, that carry olefinic moieties and are introduced in the $i,i+4$ or $i,i+7$ positions of the chosen peptide sequence [41].

The inclusion in a peptide fragment of residues with D-configuration (mainly D-Proline) might favor turn formation and consequently stabilize β -sheet structural organization; the "D-Proline-L-Proline" dipeptide motif can be exploited as well to stabilize β -hairpins [33]. In addition, the formation of β -sheets and β -strands can be induced by macrocyclization or even the introduction of amyloid-like fragments, which are known to self-associate by forming extensive β -sheet layers [33].

Although macrocyclic peptides can be considered valid therapeutic candidates for their improved affinity for a target and stability, they do not possess a proper balance of rigidity and flexibility and also lack the appropriate solubility required to move between aqueous and lipophilic environments [40]. To ameliorate this undesirable feature of macrocyclic peptides, it can be convenient to introduce, inside smaller cyclic peptides, residues provided with N-methylation and D-configuration. Such small cyclic peptides can be more suitable for oral dosing as they can be characterized by better metabolic stabilities with respect to macrocyclic peptides [40].

The bioavailability of peptide drugs, which is highly reduced by enzymatic cleavage occurring in the intestinal tract, can be enhanced by coformulations with inhibitors of protease activity. However, this approach may present side effects leading to digestion- and pancreas-related problems. It might be convenient to employ carriers that enhance peptide absorption; these conveyors include a variety of chemical molecules that can operate through several mechanisms, such as the opening of tight junctions and modulation of membrane fluidity and mucus viscosity [33,35]. Among the carriers, we can find chelating agents like ethylenediaminetetraacetic acid (EDTA) and citric acid, short fatty acids (i.e., molecules with an amphipathic nature) like sodium caprylate, sodium lauryl sulphate, or sodium taurocholate (representing an example of bile acid) [35].

A further route to improve peptide oral bioavailability and pharmacodynamic features consists of employing mucoadhesive polymeric systems (e.g., sodium carboxymethyl cellulose, polyacrylic acid, polyethylene oxide, methyl cellulose, etc...) that become attached to the intestinal epithelium, thus improving absorption at the mucosal membrane [35,39].

Nowadays, the targeting of intracellular space with peptide therapeutics remains a great challenge [38,42,43]. From a historical perspective, peptide drug discovery was initially (through the 1970s) centered around receptor targets; later (1980s–1990s), remarkable accomplishments were achieved regarding the discovery of peptides showing good cellular permeability and capacity to regulate intracellular protein targets like the macrocyclic natural peptide product cyclosporin, and diverse groups of cell-penetrating peptides (CPPs)

including the synthetic ones obtained by conjugation with the transactivating transcriptional activator (TAT). To describe in depth the methods to improve peptide permeability and delivery inside cells is beyond the goal of this work, as many reviews have been centered on the topic [42–45]. However, it is worth noting that several experimental and computational screening instruments to investigate peptide permeability and support the development of intracellularly targeted peptide therapeutics are presently available. For instance, there exist methods that implement cellular monolayers to analyze transcellular permeability; computational/biophysics tools comprise those focused on exposed polar surface features, the comparative analysis of intramolecular and solvent H-bonding by NMR and mass spectrometry (MS) analyses, the evaluation of octanol/water phase partitioning to establish energy-independent translocation, and the analysis of the radius of gyration, which represents a sort of substitute characteristic for the molecular weight in classes of molecules that do not conform to Lipinski's rule of 5 [38,46].

To generate more drug-like agents, bifunctional systems, exploiting advantageous features of both peptides and small molecules, can be designed and evaluated with the support of *in silico* tools [47,48]. Indeed, once selected, a peptide with an established high affinity for a specific protein target and for which the 3D structure of the protein–peptide complex is available can be split into two fragments provided with a reduced (or not even appreciable) interaction affinity for the protein target. Next, reactive groups can be inserted at one terminal side of each fragment to allow for the exploitation of “click chemistry” to combine the two peptide segments with small molecules. In the end, screening can be conducted with a plethora of small molecules to find out which moieties are able to substitute one of the two peptide segments and restore the high affinity of the starting peptide [47].

The following sections will be focused on the diverse computational approaches for the development of peptides able to hamper crucial PPIs involved in pathological conditions. Several examples of *in silico* identified peptides or peptide-based molecules that can be considered promising starting points to develop therapeutic agents will be highlighted.

2. Identifying *In Silico* Novel Bioactive Peptides: Methodological Aspects

A variety of virtual screening approaches can be applied for the identification of novel potential bioactive peptides. Much attention needs to be given to peptide libraries to be implemented in such computational screenings, as they need to be designed by considering key peptide features that might influence bioactivity and drug-likeness [including, but not limited to, cell permeability, aggregation tendency, and stability (See Section 1.2)].

2.1. Design of Virtual Peptide Libraries

The construction of virtual peptide libraries for specific drug discovery applications can be supported by *in silico* tools to predict peptide bioactivity and drug-like characteristics. However, a mixed computational and experimental approach, which was undertaken to analyze bioactivity in a group of peptides derived from goat casein hydrolysate, highlighted that *in silico* bioinformatics tools to predict bioactivity have some limitations in terms of the range of activities that can be associated with peptide sequences and should only be used in combination with *in vitro* and/or *in vivo* tests. The disagreement in the results from diverse *in silico* predictors is also another potential drawback [49].

Moreover, numerical descriptive vectors (NDVs) for peptide sequences might be considered when planning which peptides are to be included in a library. Indeed, amino acids' physicochemical features can be exploited to obtain, through principal component analysis (PCA), NDVs for peptide sequences [50]. NDVs possess a length equal to peptide sequences, and each entry in an NDV corresponds to an amino acid. NDVs can be employed for quantitative structure–activity relationship (QSAR) analyses of peptide groups and for the prediction of peptide residues representing hot spots [50].

A variety of virtual peptide libraries can be generated through diverse approaches depending on specific peptide properties that need to be satisfied to address the scientific problem under investigation, and a few examples will be provided below.

First of all, to support virtual screening approaches and speed up the computational identification of bioactive peptides, Prasastry and Istyastono provided the structures of 168,400 peptides (di-, tri-, and tetra-peptides) that were generated from all possible combinations of the 20 natural amino acids. The structures were reported in both a simplified molecular-input line-entry system (SMILES) and three-dimensional (3D) formats suitable for diverse molecular docking pipelines [51].

In addition, a powerful instrument for the fast computational analysis of peptide libraries is PDAUG (Peptide Design and Analysis Under Galaxy). Indeed, PDAUG includes several tools for the development of peptide libraries, data visualization, the calculation of peptide characteristics, modeling, and links to databases from which to retrieve peptide sequences. For instance, PDAUG includes the routine “amino acid Property Based Peptide Generation” that creates peptide sequences relying on amino acid properties. The “Sequence Based Peptide Generation” routine instead allows for the creation of peptide sequences through three diverse routes: “Random Peptides” (a method that seeks all the probable combinations of 20 natural amino acid residues within the chosen sequence length), “Mutated Peptides” (a method that, at specific positions, substitutes the existing residues with the residual 19 amino acids) and “Sliding Window Peptides” (this tool uses a protein sequence as input and randomly produces peptide segments based on a sliding gap and fragment dimension) [52].

Ad hoc-generated peptide libraries can be implemented for the design of PPIs by preserving specificity that generally characterizes naturally occurring interactions. In this context, a clever strategy made use of a library composed of 1536 peptides where 32-mer sequences were generated by including characteristic features of the parallel dimeric coiled-coil pattern along with the semi-random introduction of residues in positions crucial for stability and specificity. Computational screening was achieved to identify, based on the predicted T_m (melting temperature) values, eight peptides that could interact together by creating four heterospecific PPIs [53]. A combination of experimental techniques was, in the end, used to validate the capacity of this protocol to predict the formation of heterospecific PPIs. This computational approach provides information that can be applied in different fields of protein science, including the discovery of therapeutic peptides able to specifically inhibit PPIs relevant to certain pathological conditions [53].

Degradation by proteases is one of the challenges that is faced when attempting to develop therapeutic peptides (See Section 1.2). A valid solution relies on the screening of peptide libraries containing D-amino acids; however, experimental strategies based on such an approach are generally time consuming and expensive due to the costs associated with the synthesis of peptides containing the unnatural D-amino acids. In silico screening campaigns with virtual libraries of D-peptides can be planned to ameliorate issues associated with the experimental methods. An interesting approach in this field is exploited by the web server finDr (<https://findr.biologie.uni-freiburg.de/>) that supports the identification of the D-peptide interactors of target proteins. finDr works similarly to a mirror-image phage display where the peptides are assayed in the L-configuration while the target is produced in the D-configuration. If an L-peptide interacts with the D-form of a protein, its mirror images (D-peptide) will be able to interact in a similar manner to the same natural protein in the L-configuration [54]. Starting from this concept, the virtual library employed by finDr consists of helical peptide fragments (12 amino acid long) extracted from Protein Data Bank (PDB) entries. Such a library is screened against the chosen protein target in the D-configuration through Mirror-Image Virtual Screening (MIVS). This leads to a prediction by molecular-docking of L-peptide ligands, the configuration of which can be inverted to obtain D-peptides interacting with the naturally occurring L-protein [54].

The bioactivity of short peptides can be modulated by their ability to self-assemble into supramolecular structures. Another interesting study reported on the employment of

a library of 400 dipeptides, representing all possible combinations of natural amino acids, which were subjected to a coarse-grained molecular dynamics approach to predict their capacity to self-associate [55].

Bioactive peptides hold great interest in drug discovery; however, different to small molecules, for which a vast chemical space can be explored, peptide optimization can only rely on the reduced set of natural and commercially available non-natural amino acids [56]. To overcome this limitation, recently, a large virtual library of non-natural and readily synthesizable amino acids was developed to enhance virtual screening approaches and assist the peptide optimization phase. The library was generated starting from the eMolecules database, that includes more than 26 million compounds. This database was first analyzed to select the included α -amino acids. Moreover, the eMolecules database already contains non-natural amino acids, but to expand this group, chemical reactions were simulated based on a few synthetic canonical routes. To overcome selectivity issues, the compounds in the eMolecules database provided with more than a single reactive group were excluded from the approach to avoid several side products. The compounds leading to reaction products containing more than a single α -amino acid group were automatically excluded as their inclusion in peptide sequences could be challenging. The compounds were included in the final library only if certain rules were respected, concerning, for example, the molecular weight, number of chiral centers, number of heavy atoms, formal charges, etc.... Once a target protein–peptide complex has been chosen, the library can be implemented to set up the *in silico* optimization approach. Modeling studies can be conducted to insert the non-natural amino acid members of the library in the crucial positions of the natural peptide ligand, and docking studies can predict improved interaction affinity and specificity [56]. This strategy can support the selection of the most promising peptide ligands to be synthesized and experimentally validated.

Virtual screening represents a powerful approach when a peptide ligand of the target protein has already been identified and validated, but it needs to be optimized. For such computational screenings, the use of libraries of peptides provided with good predicted affinities and specificities for the target protein should be preferred with respect to randomly generated mutant peptide libraries. A library optimization method that relies on experimental binding information or *in silico* produced structural data to guide the best amino acid substitutions at each peptide position has been proposed [57]. Once a protein–peptide target system has been chosen, the method can exploit diverse data, including experimental binding information derived, for example, by SPOT arrays and alanine scanning, if available, or *in silico*-generated scores for peptide mutants if a 3D structure of a protein–peptide complex has been solved [57]. This information is implemented to classify amino acid substitutions into two sets: “required” and “preferred”. “Required” mutations will constantly be incorporated when developing a new library and comprise the amino acids of the starting wild-type peptide sequence along with additional residues, the influence of which on the affinity and specificity is proved by experimental or computational data. The “preferred” group will include substitutions that an optimization algorithm identifies as possible instead [57]. After initial library generation, further analyses and iterative optimization steps can be completed [57].

Another active area of research focuses on the design of cyclic peptide libraries for *in silico* strategies. As explained in Section 1.2, cyclization is another route to enhance the peptide drug-likeness by reducing susceptibility to enzymatic degradation. Golosov and collaborators [58] described the design of peptide libraries containing a macrocyclic arrangement to allow for both permeability and oral exposure [58]. Such a library was built by speculating that a properly chosen N-methylation pattern that could favor passive permeability should stabilize a macrocyclic structural organization provided with a low desolvation penalty for the passage from water to a membrane-like milieu [58]. Different libraries of macrocyclic peptides were designed (i.e., 6-mer, 7-mer, or 8-mer members), and macrocyclization was achieved through a thioether bond, as it could be experimentally easily generated upon a reaction between the side chain of a C-terminal cysteine and an

N-terminal electrophilic group [58]. Moreover, other rules were considered to generate peptides: the presence of a UV chromophore (i.e., D- and L-phenylalanine with or without N-methylation), presence or absence of a proline residue, presence of mimics of the leucine residue (i.e., norleucine and backbone methylated norleucine), as leucine is contained in certain oral natural products and peptides, and a maximum of three or four N-methylated residues [58]. After the initial design, the *in silico* generation of peptide models was achieved along with conformational sampling in an environment characterized by a low dielectric constant to simulate the nonpolar membrane region. Next, transfer-free energies for the passage from an aqueous to a membrane-like background ($\Delta G_{\text{transfer}}$) were estimated for each peptide conformation. This phase was followed by the evaluation of those conformations characterized by both low energy and low $\Delta G_{\text{transfer}}$ values for each virtual peptide. These combined energetic considerations were employed to guess potentially permeable macrocycles (i.e., those provided with low $\Delta G_{\text{transfer}}$ values and conformations with the reduced number of solvent-exposed backbone amide protons). Indeed, the best-selected macrocyclic peptides were synthesized and *in vitro* tests proved the effectiveness of this screening protocol. This approach can be nicely applied for the identification of sets of permeable cyclic peptides characterized by large side-chain diversity [58].

Peptidomimetic macrocycles are very attractive in drug discovery not only for their drug-like character but also for their potential ability to target the most undruggable protein targets. One limitation of drug discovery based on peptidomimetic macrocycles is often related to difficult synthetic routes. Saha and collaborators recently created a computational platform to generate libraries of synthesizable peptide macrocycles deriving from a multistep reaction series [59]. The platform relies on previously established experimental protocols to merge small linear peptides with synthetic scaffolding reagents. The merged bifunctional compounds are next converted into amphipathic macrocycles provided with specific conformations and enhanced pharmacological features. This computational tool is made up of two principal components: the “composite peptide macrocycle generator” (CPMG) and “ConfBuster++”. CPMG can be first employed to create, starting from user-selected building blocks, a library of two-dimensional (2D) macrocycle structures; the conformations of each macrocycle are next produced by ConfBuster++ [59]. In detail, the library can be created starting from an ensemble of building blocks made up of amino acid derivatives having diverse drug-like and conformationally constraining patterns. Linear oligopeptides are produced by CPMG through methodical permutations of building blocks and are then inserted into specific templates. Template-linked oligopeptides are subsequently transformed into macrocyclic arrangements following rules gained from experimental studies and considering estimates for site reactivity. To favor maximal diversity in macrocycles, a filtering strategy can be applied based on the analysis of physical and three-dimensional characteristics. This *in silico* approach represents an attractive tool to be considered before starting very large virtual screening campaigns with macrocyclic peptides/peptidomimetics to favor the selection of peptides that could be more easily synthesized by established routes [59].

When planning the design of a peptide library for virtual screening, it is also important to evaluate the possibility of introducing, within peptide sequences, cell-penetrating motifs. Cell-penetrating peptides attract attention for intracellular delivery as they can work as carriers for different classes of molecules (e.g., small compounds, peptides, and oligonucleotides), thus overcoming their low bioavailability issue. In this context, computational tools to design CPPs can be very supportive [60,61]. The free web tool “CellPPD” (i.e., <http://crdd.osdd.net/raghava/cellppd/>) can be, for example, implemented to predict CPPs with elevated accuracy [60]. The approach behind “CellPPD” relies on the exploitation of different peptide features (e.g., the type of residues in each position, two-residue motives, and physicochemical properties) to develop support vector machine (SVM)-based models through which it is possible to discriminate between CPPs and non-CPPs [60]. The likelihood of false positives is a problem that can be encountered in CPP prediction, and the “AiCPP” approach seems to provide a possible solution [61]. In detail, this is a

deep learning-based method employing a negative control group made up of a vast set of peptide sequences from human reference proteins [61].

In silico tools to generate peptide libraries able to hamper the binding of antigens to antibodies have also attracted attention. Synthetic peptides constitute a good alternative to antibodies that, although holding great therapeutic potential, might be characterized by unfavorable drug-like features. In this field, the “epitope-paratope interaction (EPI)-peptide” designer tool can be considered a promising computational instrument. Its workflow consists of the initial identification of interacting residues within the contact interface of an antibody–antigen complex through the selection protocol “Interface Interacting Residue (I2R)” [62]. Based on the calculated ensemble of all the intermolecular contacts, the 3D structure of the antibody–antigen complex is translated into interface diagrams. In the end, antibody residues, predicted to bind to the target epitope, are exploited by the EPI-Peptide Designer tool to develop libraries of peptides that should work as ligands of the antigen under investigation and thus possibly function as therapeutic agents [62].

Similarly, a virtual screening method to iteratively develop virtual libraries of peptide ligands able to target the Fc portion of the IgG (Immunoglobulin G) antibody will be described in the next paragraph [63].

Instead of employing very large virtual peptide libraries for virtual screening, it can also be convenient to explore the “in silico panning” approach to identify a peptide ligand of a target protein starting from a small library by merging docking studies with genetic algorithms (GAs) through which peptide ligands are iteratively developed [64]. For “in silico panning”, docking analyses of a small peptide virtual library are conducted first so that peptides are ranked according to their docking scores and predicted affinity [64]. Then, the best peptide ligands are evolved through GAs. GAs somehow simulate the genetic evolution of biological systems and include three steps (selection, crossover, and mutation) that are carried out on an ensemble of sequences to produce a novel peptide generation. The process is repeated several times, and during evolution, the docking energies decrease, leading to the identification of optimized peptide ligands. Yagi and collaborators reported on the application of “in silico panning” to identify a non-competitive inhibitor of the water-soluble quinoprotein glucose dehydrogenase starting from a virtual peptide library composed of just 10 tetrapeptides including seven diverse amino acids and performing four rounds of selection [64].

Interestingly, GAs are also key elements of another protocol developed to avoid the molecular mass repetitiveness that can characterize the combinatorial construction of peptide libraries. In detail, multi-objective GAs have been used to generate large libraries by ensuring the maximum possible number of permutations between positions of peptide sequences and a low probability of having library members with identical masses and/or shared sequences [65].

2.2. Virtual Screening Approaches in Brief

Peptides attract attention for their involvement in the regulation of a considerable number of biological processes, as peptide epitopes mediate an array of PPIs [66]. The size and flexibility of peptides allow them to interact with high specificity to diverse receptors, even to those considered undruggable from the perspective of small molecules. In addition, the potential capacity of peptides to interact with large and flat protein surfaces makes them ideal candidates as therapeutics able to inhibit PPIs that are crucial in modulating pathological conditions [66].

Furthermore, certain peptides can self assemble, forming hydrogels, fibers, and a variety of structures that can find several applications in biomedicine, even as supports for tissue regeneration.

Computational routes are surely helpful and convenient for the prediction of bioactive peptides possessing a particular self-assembly ability or capacity to bind with high affinity and specificity to a protein target and block crucial PPIs [66]. In drug discovery, virtual screening can quickly support the identification of the most promising peptide hits that

should be first synthesized and submitted to experimental validation to speed up the overall process and also reduce connected costs.

Virtual screening approaches can be subdivided mainly into two categories: structure-based virtual screening (SBVS), that relies on the complementarity between the ligand and the target's binding pocket (i.e., receptor–ligand docking), and ligand-based virtual screening (LBVS), which instead employs active ligands as prototypes to target protein receptors. LBVS methods are focused on the selection of candidate molecules, considering that a peptide interactor similar to a bioactive peptide ligand has a higher chance of possessing biological activity. A computational route combining both ligand and receptor-based approaches is often a good choice to identify original molecules and reduce the chance of finding false positives [67].

Challenges usually associated with interactions involving peptides include a low amount of available 3D structures and a lack of information on the binding affinities linked often to the transient nature of peptide-based interactions. In addition, the intrinsic flexibility of peptides highly affects the success rate of computational tools for structure and affinity predictions [66].

Molecular docking predicts the structural arrangement of the ligand, as well as its orientation and position at the binding site [68]. Two diverse groups of docking algorithms to predict protein–peptide interactions exist: template-based and template-free [69]. The template-based approach employs the 3D structure of similar complexes as a model to guess the binding mode of a peptide to a protein, and GalaxyPepDock [70] is an example of such a docking tool. The application of template-based docking has certain limitations related to the restricted availability of templates. No template is required for template-free docking, which includes two diverse approaches (i.e., global and local docking) that can be applied to different case studies depending on whether the knowledge of the binding pocket is available or not. Global docking samples the total protein surface to obtain the peptide interaction mode and identify the binding pocket. Instruments to achieve global docking include the webservers pepATTRACT [12] and HPEPDOCK [71].

Local docking relies instead on a user-defined binding pocket around which to look for peptide binding poses. Popular programs for local docking are AutoDock Vina [72,73], GOLD (Genetic Optimisation for Ligand Docking) [74,75], and HADDOCK (High Ambiguity Driven protein-protein DOCKing) [76,77]. A larger list of docking software for protein–peptide complexes can be found in a 2020 work by Weng and colleagues [69].

Three main subgroups of docking approaches exist: rigid, flexible, and semi-flexible [68]. In rigid docking, ligands and proteins represent rigid objects, and sampling is achieved by keeping in account only six degrees of freedom (i.e., three translational and three rotational ones). This approach can be mostly implemented for protein–protein docking, as in this case, the total number of conformational degrees of freedom are too many to be sampled. Instead, in semi-flexible docking, the conformation of the receptor or peptide ligand can undergo some fluctuations during docking runs. Usually, flexibility is contemplated for ligands as they have a lower number of conformational degrees of freedom that can be sampled during docking runs. Such an approach considers the rigid structure of the receptor as the conformation able to interact with the peptide ligand. Flexible docking assumes instead flexibility in both the protein and ligand and is generally based either on an induced fit interaction model or on conformational selection [68].

The DINC (Docking INcrementally) 2.0 and iMolsDock web tools represent valid examples of computational instruments to face the flexibility issue which, as mentioned before, can characterize not only peptide ligands but also protein receptors [78,79].

Hence, protein–peptide complexes can be studied using different pipelines depending on the presence of flexibility in the protein alone, in the peptide alone, or in both [79]. When dealing with flexible ligands rather than small molecules, it needs to be kept in mind that virtual screening by molecular docking might possess limited accuracy for ligands characterized by more than 10 flexible bonds [78]. Interestingly, the protein–peptide docking webserver named DINC 2.0 moves this limit to 25 flexible bonds thanks to a

parallelized meta-docking approach through which large ligands can be docked against target proteins in an incremental manner [78]. Rather than performing a docking of the entire ligand, the complexity is reduced by incrementally docking larger and overlapping sections of the ligand. Indeed, in the first step, the docking of a small portion of the ligand with only six degrees of freedom (DoFs) is achieved, and the best binding modes are selected [78]. Starting from the selected docking poses, the related fragment is amplified through the addition of ligand atoms; afterwards, a new docking round is carried out. Successive docking runs are conducted, and every time, three new DoFs are included and combined with the previously considered fragment for which just three flexible DoFs are maintained [78]. Therefore, at each docking round, independently from the fragment dimension, only six internal DoFs of the ligand are considered. This stage is repeated until the entire structure of the peptide ligand is covered and docked [78].

An induced-fit docking instrument called iMolsDock employs the mutually orthogonal Latin squares (MOLS) sampling method and is able to consider protein flexibility [79]. A recently updated software version allows for enhanced receptor flexibility, better scoring function, and fast calculations. Initially, the MOLS approach was used to perform peptide modeling by relying on the concept that $(S)^k$ conformations exist for a peptide provided with k torsion angles that can assume s diverse values. The conformational space that can be sampled by the peptide derives from the contribution of all possible combinations of torsion angles. The MOLS approach is able to guess the optimal peptide structure with the lowest energy by inspecting the peptide conformational energy landscape. MOLS was then expanded to perform docking. The ligand docking pose can be obtained by the tool iMOLSDOCK through ligand conformational sampling, including rotation and translation. Nevertheless, to keep protein flexibility into account, the search space of the docking routine was extended to include the conformational space accessible to flexible residues in the receptor [79]. The induced-fit docking protocol of iMOLSDOCK assumes that if the conformations of the ligand and flexible residues in the protein rely on k and l torsion angles, respectively, and the interaction pose of the ligand is established through six extra parameters (three related to the position of the ligand and extra three related to its orientation in the protein binding site) then, the space that is sampled will have $k+l+6$ dimensions with a volume equal to $(s)^{k+l+6}$ and every dimension will be sampled s times. In the end, to find the optimal binding pose for the ligand in the protein interaction pocket, a gradient minimization is performed as well [79].

Among the protein–peptide molecular docking instruments, it is worth mentioning CABS-dock. CABS-dock exists as a webserver and a separate program; it represents another approach for docking that can keep into account a substantial contribution from conformational flexibility for both the protein and peptide [80]. CABS-dock does not involve a predefined binding pocket and considers the peptide as fully flexible. Concerning the protein, its backbone can be subjected to small fluctuations and optionally larger movements [80].

For flexible ligand receptors, another appealing docking web tool is HTP SurflexDock. This is a docking instrument aiming to enhance SBVS's success rate by combining two routes. In the first stage, the ensemble docking protocol is implemented to simulate the inherent receptor flexibility. In the second step, the best docking hits can be rescored either through a search of a larger conformational space or by evaluating the interaction-free energy by means of the molecular mechanics (MM)/Poisson–Boltzmann surface area (PBSA) method [81].

PepVis was instead developed with the goal of making available to the scientific community an effective instrument to automate peptide virtual screening by combining several open-source tools to carry out ensemble and flexible docking approaches [82]. Virtual screening through the Graphical User Interface (GUI)-based pipeline PepVis is achieved by exploiting a combination of several docking tools (i.e., AutoDock Vina [72], ZDOCK [83], and AutoDock CrankPep [84]) [82]. In addition, this pipeline relies on two bioinformatics tools for the modeling of peptides (i.e., Modpep [85] and GROMACS (GRONingen

MOlecular Simulation) [86,87]), one for the re-ranking of peptides (i.e., ZRANK2 [88]) and one for the refinement of the receptor–peptide complex (i.e., FlexPepDock [89]) [82]. The advantage of this pipeline is its modular nature, which allows for the insertion of additional bioinformatics tools, thus increasing the information that can be acquired through it [82].

As mentioned earlier, a virtual screening protocol specific for the identification of peptide ligands able to target the Fc portion of the IgG antibody has been developed; such an approach is also very useful for virtual library development [63]. In the first step, a combinatorial library of tetrapeptides was designed by combining the 20 natural amino acids, and the conformations for each peptide were generated so that the 3D structures were available for each library member. The peptide library was screened against the crystal structure of IgG1 Fc by the docking software CmDock (CurieMarieDock) (v. 0.2.0) [63]. In addition, the best predicted ligands (i.e., 100 tetrapeptides) were subdivided into clusters and further analyzed to estimate the residues important for interaction with the antibody and determine the highest occurrence of specific amino acids in each peptide sequence position [63]. Interestingly, the CmDock program identified preferred amino acid positions inside the top-scoring peptides, which were consistent with the current knowledge of peptides binding to the Fc region of IgG. The method can be further implemented to develop a more focused library for the identification of peptide ligands of antibodies with extended sequences [63]. Indeed, the protocol could proceed iteratively and even be used for developing focused libraries against different targets rather than antibodies.

Instead, a kind of inverse virtual screening approach was set up based on the “select and purge” (SP) algorithm by combining sampling by statistical approaches and selection through rigid molecular docking. This virtual route allows for the identification of short peptides able to function as receptors for small compounds when no knowledge about the peptide/small molecule association is available [90]. Briefly, this approach requires the assembly of a library of small molecules by considering different criteria for compound selection [90]. Concerning the small peptides to be investigated, tripeptides deriving from all possible combinations of the 20 natural amino acids can be implemented. Then, docking analyses are conducted to predict the binding affinities for the peptides/small molecule complexes [90]. The resulting data are exploited by the select and purge (SP) algorithm to further assess how amino acid positions and types can influence the ability of tripeptides to form complexes with specific ligands [90]. False positives are eliminated, and the best candidates are chosen based on the unfavorable and favorable (i.e., those stabilizing the complex) pairs of amino acid positions within the analyzed peptide sequences. The algorithm works iteratively, and after the stop point is reached, a set of possible peptide receptors is detected among the investigated dataset [90].

Specific *in silico* pipelines, somehow different to the common docking-based virtual screening protocols described before, were applied to cyclic peptides [91,92]. As mentioned before, constrained bioactive peptides assume particular relevance in drug discovery due to a likely increased specificity for the target and higher resistance against proteases [91,92]. The cyclic PEPtide matching (cPEPmatch) method consists of a fast *in silico* route to identify cyclic peptides that could interact with specific segments located at the interface of protein–protein complexes [92]. This approach is based on the comparison between the backbone of short peptide fragments, derived from the protein–protein binding surface, and the backbone structures of an ensemble of cyclic peptides. When a positive match is found, the corresponding cyclic peptide is taken as a model interactor and used as a template by adjusting its amino acid side chains based on those of the protein–protein complex. For 154 protein–protein interactions that were analyzed against a small library of cyclic peptides for which structures were available, at least one peptide template could be found for a specific region of a protein–protein interface. Interestingly, for most of the predicted cyclic peptide ligands, MD (molecular dynamics) analyses and binding energy evaluations revealed stable protein–cyclic peptide complexes and good binding energy scores, thus indicating the promising effectiveness of the method [92].

It is not easy to establish which protein–protein interface is suitable to be targeted by cyclic peptides. A high-throughput approach and a rapid method for the virtual screening of peptide libraries were set up by Duffy and colleagues [91]. First, combinatorial 3D peptide libraries, including small peptides containing disulfide bridges, were developed *in silico*. Next, diverse members of these libraries were compared and matched with pharmacophore models based on the structures of crucial protein–peptide and protein–protein complexes. These pharmacophores were constructed by analyzing the 3D structures available in the PDB and consisted of short peptides known to bind proteins, linear motifs, and turns characteristic of protein–protein interaction surfaces. In the end, pharmacophoric models were compared with more than 100,000 cyclic peptides in multiple conformations [91]. This pharmacophore screening is a protocol useful to predict not only the PPIs suitable to be targeted but also the cyclic peptide modulators of these optimal PPIs that can be thus prioritized for synthesis and experimental validation [91].

Limitations of Virtual Screening Approaches

SBVS of compound libraries is a widely employed approach at the initial stages of the drug discovery process to support hit-finding campaigns [93]. Indeed, high-throughput virtual screening and molecular docking platforms have become very popular within the scientific community due to their great capacity to analyze large databases of molecular entities with low costs and rapidity [94]. The employment of protein–ligand docking algorithms to identify potential modulators of disease-related proteins has also been supported by the increasing number of high-resolution protein structures being solved by X-ray crystallography and other experimental techniques, the advancements in molecular modeling platforms, and the development of very powerful computer resources [95].

Molecular docking represents the key point in SBVS [96]. Briefly, it includes two principal phases relying on two diverse algorithms. The diverse conformations (i.e., poses) that a ligand can assume in the receptor binding site are predicted by the sampling algorithm. In addition, the binding energies associated with the receptor–ligand poses are predicted by a scoring function. Scoring functions are generally used to achieve the filtering and ranking of the obtained poses in virtual screening, indicating which screened compounds should represent a possible lead [96]. A robust virtual hit selection is particularly important when working with new targets and can rely on a variety of protocols, including scoring, ensemble docking, consensus pose, and ligand efficiency, to enhance the accuracy of docking results [97]. Due to improvements in computer resources and in the dimension of compound libraries, it is possible to screen billions of molecules with computer clusters provided with a modest size [93]. However, the fast exploration of such a large chemical space through VS requires approximations, leading to the under-sampling of potential configurations and incorrect estimates of absolute interaction energies [93]. Consequently, the usage of docking-based VS alone comes with faults. Nevertheless, different laboratories plan VS workflows based on software and techniques they are more familiar with, but also, when software appears simple to use, every practice presents several weaknesses that need to be considered to avoid incorrect results/artifacts [98]. Moreover, the simplification of molecular docking lowers the efficiency of the docking score, leading to a higher chance of obtaining false positives [99].

As briefly mentioned in Section 2.2, one major issue of VS is often related to the static nature of the receptor during rigid docking, where the dynamic features of biological structures are not properly taken into account [96]. For practical reasons, most docking platforms employ a rigid receptor approximation, in which the ligand is considered a flexible entity, while the conformation of the protein is constrained. A few approaches, like ensemble docking, can be instead a better choice as, in this case, an ensemble of receptor conformations allows for the investigation of protein flexibility [99]. Nevertheless, the docking of large peptidomimetics is challenging with ordinary docking protocols (see also Section 2.2) as docking algorithms can be rather inaccurate in predicting the right poses for molecular entities provided with many rotatable bonds. In fact, the addition of each

rotatable bond induces an enlargement of the conformational space to be sampled and a decreased chance to obtain the correct binding pose [96].

The generation of both false positives and false negatives by docking-based VS can be due to limitations in sampling algorithms and flaws in scoring functions; in addition, the need for training groups in diverse algorithms frequently induces a highly target-dependent accuracy [96]. Another drawback is linked to docking poses and scores severely subjected to the ligand input conformation. In fact, subtle variations in the ligand input structure can induce significant divergences in the resultant docked poses [99]. Much care needs to be taken in the initial preparation of the ligands and structures to be submitted to VS [98].

A significant limitation of the docking software is often linked to the inability of scoring functions to correctly predict ligand binding affinities [99]. Consequently, the best docking hits should not be selected just based on energy scores. Docking scores should be used to discard molecules that do not fit the binding/active pockets but not to establish thresholds for biological activities. Another aspect that needs to be crucially considered is that the comparison of docking scores from diverse docking platforms is not reliable, as these scores are strictly linked to the force fields and protocols implemented by each software [99]. Another important aspect of molecular docking is related to the specific treatment of water molecules that are present inside the binding pocket throughout the docking progression, that could induce the erroneous evaluation of the potential interactions between the receptors and ligands [99]. Finally, the capacity of docking software to distinguish between inactive and active molecules is generally largely connected to the employed protein structure and the degree of similarity between the screened compound and a co-crystallized ligand [99].

As mentioned before, a key issue is the failure of scoring functions to properly rank compounds based on their activities towards the chosen target protein [95]. There are several factors that might negatively impact the performance of scoring functions in successfully guessing interactions energies. Empirical functions deliberately have a simple form in order to allow for computational speed, and can often only slightly compensate for the rigid protein approximation. In addition, several penalty terms, including electrostatic and steric clashes and internal ligand constraints, cannot be easily and correctly parameterized. Due to several approximations, scoring functions can be too “soft”, leading to many incorrect hits (i.e., false positives) [95]. Moreover, in the context of scoring functions, it needs to be recognized that entropy and desolvation are problematic to be handled even inside a strict molecular mechanics formalism [95]. The limitations and difficulties of using molecular docking in nutraceutical research have been covered in a recent work by Agu and coworkers, which also discusses the reliability of scoring functions and the requirement for experimental validation [100]. Limitations in docking-based VS to identify compounds targeting SARS-CoV-2 (Severe Acute Respiratory Syndrome-Coronavirus-2) proteins have also been reported in a few works [97,99].

It emerges from the literature analysis that precautions need to be taken before starting HTVS studies to identify biologically active compounds. For example, regardless of the employed docking platform, it is important that VS is conducted through several trials and that docking solutions are visually inspected to observe similarities in binding poses and recognize artifacts [94]. The reproducibility of the docking experiments should be assessed through the analysis and validation of interaction poses referring to a minimum number of trials, and this protocol also ensures the identification of the correct protein–ligand structural topology of binding [94].

If structural data can be retrieved from the Protein Data Bank (PDB), the docking results should be, of course, further validated through comparison to high-resolution experimental structures of protein–ligand complexes [94].

It is also recommended to validate results generated through HTVS further by molecular dynamics simulations, where the dynamics of both the ligand and receptor are allowed, and the stability of the interaction poses can be better estimated [94].

Within the context of VS, re-scoring is something to keep in mind, as this can yield an improved correlation between the docking results and experimental evidence [96]. For the re-scoring of docking solutions, standard molecular dynamics can be coupled to the prediction of binding free energy through MM-GBSA (molecular mechanics with generalized Born and surface area solvation) and MM-PBSA (molecular mechanics/Poisson–Boltzmann Surface Area) [99]. MM-GBSA or MM-PBSA are usually associated with modest computational efforts; they produce more solid results than those provided by most docking scoring functions and have been largely exploited by the scientific community. These computational techniques do not represent tools for screening large libraries of molecular entities but are useful for analyzing the selected docking poses [99]. Coupling molecular docking with MM-GBSA or MM-PBSA re-scoring can be considered a very promising approach to selecting the correct binding poses and ordering a group of ligands according to the binding affinities. Again, the success of these approaches is target structure dependent [99].

The availability of a lot of published data on the protein target increases the chance of achieving a successful VS campaign [96]. The accessibility of the high-resolution X-ray or NMR structures of the receptor is also vital before starting the *in silico* screening, as a reduced performance of VS could be achieved by using homology models [96]. Nevertheless, if possible, the employment of docking-based VS of holo-structures, rather than apo-ones, could further enhance the success of *in silico* approaches leading to the enrichment of promising hits. Exhaustive knowledge of the active and/or allosteric interaction pockets, including a complete picture of the flexible side-chains and their positioning within the binding site, the occurrence of water molecules inside binding pockets, and the exact protonation forms of ionizable residues, will also contribute to the positive outcome of a VS campaign [96].

As VS protocols include several computational techniques, in the end, virtual screening hits, including those validated through MD and binding free energy calculations, represent just predictions that require further *in silico* and experimental testing to validate their activities. *In silico* testing can be achieved by running additional parallel virtual screenings of groups of known active and decoy molecules, if available [98]. Experimental *in vitro* and/or *in vivo* validation of the most promising computational hits must be performed after the proper compound selection has been achieved. The selection of biologically active ligands (such as possible therapeutic agents) by VS might also take ADME and solubility predictions into account [98]. The experimental validation of the predicted active VS hits should not be considered the end point of the drug discovery process as the newly identified compounds might represent just the base for further optimization cycles [98].

Practical guidelines that can be followed for large-scale docking have been reported in an interesting work by Bender et al. [93]. The authors describe the best practices to follow before starting a massive docking-related VS approach and suggest the software-independent controls that are needed to choose the best docking parameters for a specific target receptor and all the required controls to possibly obtain hits with desired activities, once experimentally validated [93].

3. Identifying Novel Bioactive Peptides through *In Silico* Approaches: Applications

The critical role played by *in silico* tools for the discovery of novel bioactive peptides has been proved in different biomedical fields, such as the development of therapeutics against cancer or viral infections, immune system regulators, antidepressants, anxiolytics, analgesics, as well as peptides for the treatment of nicotine addiction, hypertension, Parkinson's disease, and neuropathic pains [101–119].

3.1. General Overview

The first step common to all diverse computational strategies consists of choosing a target protein or PPI for the specific class of bioactive peptides to be identified.

In searching for antihypertensive drugs, the angiotensin I-converting enzyme (ACE) can be employed as a target and the computational screening of peptides deriving from

the hydrolysis of proteins extracted from animal sources is a potential approach [101]. A mixed in vitro–in silico strategy based on ACE has been reported; the experimental protocol included the in vitro hydrolysis of α -lactalbumin, the isolation of fractions with ACE inhibitory action, and the identification within these fractions of diverse peptide sequences that were employed to generate a peptide library. The library was next implemented for computational screening and docking poses for the ACE/peptide complexes, which were analyzed to predict peptides potentially able to hamper ACE activity efficiently. Docking studies were also valuable for distinguishing the competitive peptide inhibitors (i.e., those able to interact with the residues of the active site) from the non-competitive and mixed ones [101]. A diverse strategy to identify the peptide inhibitors of ACE started from the in silico hydrolysis of the protein nebulin from *Larimichthys crocea* to generate tripeptides that were subsequently employed in virtual screening from which the “HGR” (Histidine-Glycine-Arginine) sequence emerged as a promising bioactive agent [102]. Similarly, the “WCW” (Tryptophan-Cysteine-Tryptophan) peptide has been predicted as a possible ACE inhibitor using a fully computational approach based on the virtual screening of a tripeptide library containing 8000 peptides that was assembled considering that di- and tri-peptides, resulting from the enzymatic digestion of food proteins, might work to block ACE activity [103]. Due to the presence of antihypertensive peptides in diverse natural sources and the possibility to exploit them as therapeutics provided with reduced side-effects with respect to synthetic drugs targeting ACE, a web tool was even established through which antihypertensive peptides could be designed, selected by screening, or localized within a protein sequence [104].

Similar approaches were applied to Dipeptidyl-peptidase IV (DPP-IV), which can be considered another appealing therapeutic target as DPP-IV inhibitors can be effective in treating type 2 diabetes. For instance, a combination of peptidomics and computational studies allowed for the establishment of characteristic peptide features important in hampering DPP-IV activity. In this case, the peptide set screened in silico against DPP-IV was derived from a protein extract resulting from the in vitro digestion of Pinto bean [120]. The design of a virtual tripeptide library and virtual screening also led to the prediction of potential peptides inhibiting α -Glucosidase activity for therapeutic applications, as α -Glucosidase is linked to diverse pathological conditions, including diabetes mellitus 2 and obesity [115].

From the above-reported case studies, it is evident that the employment of peptide libraries derived from the in silico/in vitro hydrolysis of specific proteins or protein active fractions from natural extracts is a common strategy that can be employed to set up computational screenings. The in silico protein hydrolysis and virtual screening of the resulting peptides against a human receptor (i.e., the heterodimer made up of taste receptor type 1 member 1 (T1R1)/taste receptor type 1 member 3 (T1R3)) were even conducted to identify umami peptides derived from *O. mykiss* (i.e., rainbow trout) [121].

As mentioned before, selecting a proper target is critical [105,106]. For instance, in silico protocols with the purpose of providing novel potential antidepressants, anxiolytics, and analgesics can be set up based on the kappa opioid receptor (KOR), the γ -Aminobutyric acid (GABA)-A receptors, and the $\alpha 2\delta$ auxiliary subunit of V-gated Ca^{2+} channels (VGCCs) [105,106]. A computational study was planned to identify compounds targeting KOR by including drug design, lead optimization, and molecular dynamics (MD) steps [105]. Starting from the initial virtual screening of a library of 6 million structures, two tripeptides were selected as best hits. The tripeptides were synthesized and assayed in vivo, demonstrating a promising antinociceptive action [105].

One very active research field relies on the application of in silico techniques to identify peptides with anticancer potential (See below) [107–110]. For instance, Caseinolytic protease P (ClpP), Ets-like protein 1 (Elk-1), the programmed cell death protein 1/programmed cell death ligand 1 (PD-1/PD-L1) pathway, the polo-box domain of polo-like kinase 1 (PLK1-PBD) and interleukin-6 (IL6) modulate different processes (e.g., the degradation of misfolded proteins, gene regulation in response to extracellular signals, immune escape,

proliferation, inflammation) linked to cancer onset and progression or acute myeloid leukemia (AML) and represent only a few possible targets [107–111]. The application of in silico methodologies alone or coupled to experimental validation with different direct and competition-type binding assays through an array of biophysical techniques has highlighted characteristic peptide features (e.g., the size and presence of non-natural amino acids) that are needed to appropriately target these specific proteins or related PPIs [107–111]. Another topic related to the anticancer drug discovery field, which is being widely explored in our laboratory through mixed approaches based on the computational screening of diverse in silico-generated peptide libraries and the experimental validation of selected peptide hits, consists of the identification of anticancer peptides/peptidomimetics targeting Sam (Sterile alpha motif)–Sam interactions mediated by the erythropoietin-producing hepatocellular receptor A2 (EphA2) [122–124].

Molecular docking techniques have also supported the identification of immunomodulatory peptides binding the Toll-Like receptor 2 (TLR-2) [112]. Major histocompatibility complex (MHC) proteins assume a crucial role in the field of adaptive immunity. Consequently, their interactions with peptides hold great interest and could be exploited to design vaccines, identify epitopes, and discover novel molecule regulators of the immune system. Doytchinova and collaborators set up an interesting in silico tool to analyze the protein–peptide binding affinity based on an additive quantitative structure–affinity relationship (QSAR) model that could be employed for the identification and optimization of peptides able to bind MHC proteins, and thus highly relevant for studies in the field of immunology [116,117].

A peculiar drug repurposing approach based on virtual screening was implemented for targeting the Calcitonin gene-related peptide receptor (CGRPR), which consists of a heterodimer involved in migraine insurgence and formed by calcitonin receptor-like receptor (CLR) and receptor activity modifying protein type 1 (RAMP1). The study aimed at identifying molecules able to hamper heterodimer formation and the results pointed out that pentagastrin, a synthetic polypeptide analogue of natural gastrin, could be further analyzed by experimental studies [114].

Among the challenges that can be encountered during virtual screening approaches, especially when dealing with peptides, the not always guaranteed availability of an experimentally derived structure for the target protein or complex and the presence of highly flexible regions, particularly within the peptides to be screened, represent two big issues. In this context, it is particularly difficult to target with in silico approaches the nicotinic acetylcholine receptor (nAChR), which is a protein involved in different diseases (e.g., nicotine addiction, Parkinson’s disease, and neuropathic pain). Thus, Leffer et al. proposed the docking algorithm ToxDock, that consists of an in silico tool based on the synergy between ensemble docking and extensive conformational sampling and performs well in the virtual screening of peptide ligands against an $\alpha 4\beta 2$ nAChR homology model [125].

Another research area where in silico strategies are commonly implemented is the search for novel routes to cure Alzheimer’s disease (AD). To treat AD, one possible tactic could rely on the modulation of human serum albumin (HSA) carrier properties to remove as much as possible of the amyloid- β peptide ($A\beta$) from the central nervous system of patients affected by this pathology. Intriguingly, virtual screening strategies have been exploited as well to define HSA variants with the highest affinity for the $A\beta$ peptide and gain information useful for the future development of therapeutics against AD [119].

This paragraph was intended to provide just a brief overview of the variety of settings in which in silico strategies were applied to search for bioactive peptides. Due to the large number of scenarios in which virtual screening and other in silico approaches can be applied, the next paragraphs will be centered more deeply on a few of the most appealing classes of bioactive peptides: anticancer, antimicrobial/antiviral peptides, and peptide inhibitors of fibrillogenesis. This choice is due to the relevance of these peptide groups in drug discovery focused on the most relevant pathologies [i.e., cancer, COVID-19

(Coronavirus disease 2019), AD, and Parkinson's disease] for which the research is still struggling to reach an endpoint.

3.2. Anticancer Peptides

Cancer is characterized by the genetic damage of the cells causing uncontrolled growth and is accounted as one of the principal causes of mortality and morbidity worldwide [126]. According to the International Agency for Research on Cancer of the World Health Organization, breast, prostate, lung, and colorectum cancers are the most frequently diagnosed tumours affecting people of all ages (data from <https://www.iarc.who.int/>, accessed on 25 November 2023 [127]). Over the past few years, many drug discovery campaigns have been aimed at detecting efficacious anticancer agents. However, traditional cancer therapies, such as chemotherapy and radiotherapy, often are not selective to cancer cells or are unsuccessful due to cancer resistance [128]. On the other hand, more cutting-edge molecular targeting therapies, such as genome therapy and immunotherapy, based on an anti-tumoral molecule that is specifically directed towards unhealthy cells, have demonstrated limited effectiveness [129]. Therefore, alternative therapeutic approaches are needed, and in this context, anticancer peptides (ACPs) have great potential due to their high selectivity, their ability to penetrate cells, and the ease of chemical modification [129].

3.2.1. Introduction

ACPs are small peptides often possessing a cationic property which confers them the ability to be selectively toxic to cancer cells characterized by negatively charged membranes. Most ACPs have similar characteristics to antimicrobial peptides (AMPs) because the surfaces of bacterial cells are also negatively charged; indeed, many AMPs are cytotoxic against both bacteria and cancer cells [128]. The structure of ACPs can be characterized by either an α -helical or a β -sheet conformation, but a few peptides with linear structures (lacking specific secondary structure elements) have also been reported. ACPs are categorized into two groups: the peptides belonging to the first group are active against both microbial and cancer cells but are ineffective towards healthy cells; the second group of ACPs includes peptides that are cytotoxic in a non-specific way (active on microbial cells, and healthy and tumoral mammalian cells) [128]. ACPs, like AMPs, can destroy the cells by membrane lysis or pore formation or can implement different routes, such as the disruption or penetration of the mitochondrial and/or nuclear membranes, finally inducing apoptosis. It has been speculated that ACPs can be more selective toward cancer cells not only for the anionic characteristics of cell membranes but also for their different lipid and cholesterol content with respect to normal cells that favor peptide penetration [128].

The fight against cancer also includes the regulation of biological pathways occurring at the cellular level and governing processes such as cell migration or angiogenesis, and in this respect, peptides can also function as interactors and regulators of protein targets with pro-oncogenic functions [130]. Moreover, peptides can be conjugated to anticancer molecules, which cannot penetrate the cells, and be employed as carriers for drug delivery [129]. Although peptides display several convenient characteristics for anticancer therapy, they are susceptible to proteolytic degradation (See Section 1.2). Nevertheless, this aspect can be improved with different modifications such as changes in the backbone chemistry, the incorporation of non-natural amino acids, and cyclization, as described in previous sections [126]. The growing interest in the development of ACPs has led to the creation of specific databases [131] or chemoinformatic and bioinformatic tools devoted to this research field [126,131,132]. In the next section, a few case studies related to the use of molecular modeling-based virtual screening strategies to find ACPs will be reported.

3.2.2. Case Studies Related to Cancer Research

Peptides from plants have several bioactive potentials, as they can show antioxidant, antihypertensive, antimicrobial, and antitumor activities. A computational study explored the possible anticancer properties of a set of antioxidant peptides from plant seeds [133].

The perturbation of ROS (Reactive Oxygen Species) homeostasis is associated with cancer, and antioxidant peptides may intervene in its regulation. The regulation of ROS homeostasis is governed by enzymes with oxidative and antioxidative activities, such as the heme peroxidase MPO (Myeloperoxidase), XO (Xanthine oxidase), and NADPH (Nicotinamide Adenine Dinucleotide Phosphate) oxidase, and by the modulation of antioxidant gene expression. NADPH oxidase is highly expressed in different types of cancer, and its activity is supported by the interaction between p47phox (phox:phagocyte oxidase) and p22phox proteins. Thus, the inhibition of the p47phox/p22phox complex may be explored in anticancer drug discovery. Nrf2 (Nuclear factor erythroid 2-related factor 2) is a transcription factor implied in the activation of genes involved in protection against oxidative stress in cancer cells and its degradation is triggered by interaction with Keap1 (Kelch-like ECH-associated protein 1). Molecules able to hamper the formation of the Nrf2/Keap1 complex could work as anticancer agents by preventing Nrf2 degradation. Due to their importance in ROS homeostasis, MPO, XO, Keap1, and p47phox were chosen as protein targets for the identification of new ACPs. A virtual library of 667 peptides from legumes, cereals, and the seeds of plants, such as oil palm and coconut, with different antioxidant activities was assembled following a detailed peptide search through the literature and the PlantPepDB database [134]. Among the library components, 592 peptides, with a length between 4 and 50 residues and lacking non-natural amino acids, were analyzed with the AntiCP 2.0 tool [135] for anticancer potential, non-toxicity, non-allergenicity, non-hemolyticity. At the end of this screening, five candidates were selected for molecular docking experiments to check their ability to interact with target proteins. The structures of MPO in complex with the compound "7GD" (7-benzyl-1H-[1,2,3]triazolo[4,5-b]pyridin-5-amine) (PDB ID: 6WYD [136]) (Figure 1a), XO in complex with quercetin (PDB ID: 3NVY [137]) (Figure 1b), Keap1 in complex with an Nrf2 peptide (PDB ID: 2FLU [138]) and p47phox in complex with a p22phox peptide (PDB ID: 1WLP [139]) were retrieved from the PDB and the regions of interaction with their ligands were used as binding pockets during docking calculations (Figure 1a,b and Table 1). Peptide 3D structures were built with PEP-FOLD3 [140], and docking runs were performed with Autodock Vina [72] through Webina 1.0.3 [141] and HPEPDOCK [71]. Among the well-predicted peptide hits, "LYSPH" (Table 1), a peptide from cherry seeds, was found to be the most promising one in binding MPO, XO, and Keap1. Instead, p47phox was best targeted in silico by the "PSYLNTPLL" peptide (Table 1) from tomato seeds. The peptides used in the molecular docking were also tested with different computational tools to predict their potential to function as anticancer therapeutics: ML-CCP (machine-learning-based prediction of cell-penetrating peptides) [142], B3Pred (blood-brain barrier penetrating peptides) [143], PlifePred (predicting peptide half-life) [144] and BIOPEP-UWM (Bioactive peptides database—University of Warmia and Mazury) [145]. Among the five tested peptides, "LYSPH" and "PSYLNTPLL" were predicted to be able to penetrate cells and the blood–brain barrier but to be susceptible to GI digestion and to have a plasma half-life lower than other anticancer peptides. Computational alanine scanning mutagenesis performed with BUDE (Bristol University Docking Engine) Alanine Scan [146] suggested that in the complexes made up of MPO, XO, Keap1, and p47phox and selected peptides, Tyr residues play important roles. For example, mutations of Tyr in Ala in "LYSPH" and "PSYLNTPLL" strongly increased the free energies of binding ($\Delta\Delta G$) associated with the peptides in complex with MPO/XO/Keap1, and with p47phox, respectively. Finally, molecular dynamic simulations were run with GROMACS 2020 [87] to analyze the dynamic stability of each selected peptide against MPO, XO, Keap1, and p47phox proteins. In conclusion, the computational work speculated that "LYSPH" could represent a peptide able to block the activity of MPO, and XO enzymes and/or interfere with Keap1–Nrf2 complex formation, whereas the "PSYLNTPLL" peptide was suggested to function as a potential p47phox ligand which could inhibit the interaction of this protein with p22phox and the consequent activation of NADPH oxidase [133]. The real ability of these peptides to interact with target proteins and modulate their cancer-promoting mechanism remains to be unveiled by means of experimental approaches.

Table 1. ACPs mentioned within this paragraph with the respective protein targets and the computational tools implemented for their identification. “Y” or “N” in the column “Exp.” indicates if experimental validation has been or has not been achieved, respectively. References associated with peptide identification are also reported in the last column.

ACP	Protein Target	Computational Tools	Exp.	Ref.
LYSPH PSYLNTPLL	MPO (PDB ID: 6WYD [136]); XO (PDB ID: 3NVY [137]); Keap1 (PDB ID: 2FLU [138]); p47phox (PDB ID: 1WLP [139])	PlantPepDB [134]; AntiCP 2.0 [135]; PEP-FOLD3 [140]; Autodock Vina [72]/Webina 1.0.3 [141]; HPEPDOCK [71]; MLCPP [142]; B3Pred [143]; PlifePred [144]; BIOPEP-UWM [145]; BUDE Alanine [146]; GROMACS 2020 [87]	N	[133]
LARLLT (D4)	Extracellular domain of EGFR (PDB ID: 1NQL [147])	PSCAN 2.2.2 Autodock 3 [148]	Y	[149]
Ebselen-WE	Kinase domain of EGFR (PDB ID: 3W2S [150])	CASTp [151]; ChemsSketch [152]; Autodock Vina [72]	N	[153]
Cyclo- (EIDTVLPTGWVAKRYS)	CTLA-4 (PDB ID: 1I85 [154])	KFC [155]; Phyre2 [156]; FlexPepDock [157]; PyRosetta [158]; NAMD 2.13 [159]; MM-GBSA calculations [160]; BFEE in VMD [161]	Y	[162]

The Epidermal growth factor receptor (EGFR) is a tyrosine kinase receptor expressed on the membranes of epithelial cells. EGFR is activated by different endogenous ligands, such as Epidermal Growth Factor (EGF), and regulates physiological cellular functions. This receptor is overexpressed in several human cancers, including lung, breast, bladder, and ovarian cancers, where it is associated with the growth and progression of tumors, regulating the angiogenesis, invasion, and metastasis of cancer cells. EGFR was the target of a study in which a peptide ligand (D4: LARLLT) was identified starting from a computational approach (Table 1) [149]. First, the X-ray structure of the extracellular region of EGFR was retrieved from PDB (PDB ID: 1NQL [147]), and then, an analysis was conducted with the PSCAN program to find a binding pocket to be targeted by peptide ligands. To this aim, a group of six amino acids surrounding the identified EGFR binding pocket (Q164, C163, S162, E110, E73, and R74) (Figure 1) was implemented to design a library of peptides using the theory of sense and antisense peptide interaction, based on the concept that peptide sequences originating by complementary nucleic acid portions are more prone to specifically make interactions. This method allowed for the design of 132 6-mer peptides that were all docked against the EGFR binding pocket (Figure 1c) with AutoDock3 [148]. Considering the docking energy values, 20 peptides were selected, and 10 of them were synthesized. Finally, the D4 peptide was considered the best choice for further in vitro and in vivo experiments. The peptide was first conjugated with PEG (polyethylene glycol) to insert it into a liposome membrane. The liposomes conjugated with D4 peptide were labelled with the fluorescent probe rhodamine and tested on H1299 cancer cells over-expressing EGFR. These experiments demonstrated that the D4-conjugated liposomes bind and penetrate the H1299 cells by endocytosis. In in vivo fluorescence experiments in mice bearing an H1299 tumor, the peptide accumulated at cancer sites within the mice body, with a circulating half-life of approximately 6 h. In vitro and in vivo experiments were also performed using liposomes conjugated with a scrambled D4 peptide as a negative control that did not show the same D4 specificity to penetrate cancer cells and accumulate in tumor mice regions. In the end, the computational approach led to the identification of the D4 peptide that

could be further employed to develop improved EGFR ligands to be conjugated to delivery systems for cancer treatment [149].

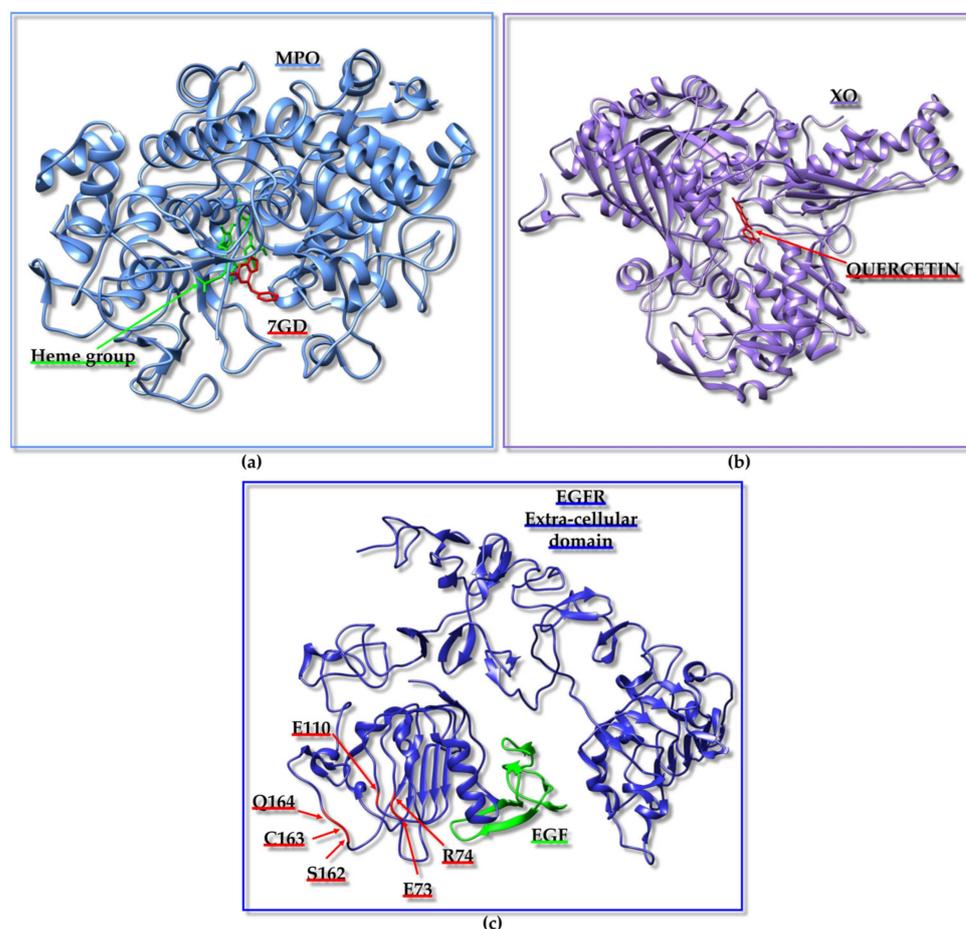


Figure 1. Examples of protein targets used in virtual screening approaches to find ACPs. **(a,b)** Identification of ACPs from plant seeds. **(a)** MPO (turquoise) in complex with the 7GD compound (red) and the heme group (green) (crystal structure with PDB ID: 6WYD [136], chains A and B); **(b)** XO (violet) in complex with quercetin (red) (crystal structure with PDB ID: 3NVY [137], chain C). **(c)** The structure of EGFR extracellular domain (blue) in complex with EGF (green) (crystal structure with PDB ID: 1NQL [147], chains A and B). The backbone of protein residues selected for the design of a virtual peptide library are colored in red, labelled with the one-letter amino acid codes and sequence numbers, and further indicated with arrows.

EGFR was the target of another preliminary computational study in which a few dipeptides conjugated with the compound Ebselen were designed [153] (Table 1). Ebselen is a selenium-containing compound with anti-inflammatory, antioxidant, and cytoprotective activities that has been reported to possess anticancer potential in different tumor types. Ebselen derivatives also present anti-tumoral functions as they have cytostatic and cytotoxic capabilities. The PDB file of the EGFR kinase domain in complex with a small molecule (PDB ID: 3W2S [150]) was used as a target to perform docking studies with three Ebselen dipeptide derivatives (Eb-WD, Eb-WE, Eb-WK). Possible protein binding pockets were searched with the CASTp (Computed Atlas of Surface Topography of proteins) software (v. 3.0) [151], and the 3D structures of the Ebselen-conjugated peptides were obtained with the ChemsSketch software from ACD labs [152]. The docking studies to analyze the interaction between EGFR and different Eb derivatives were carried out with AutoDock Vina [72], and the resulting docking energies were compared with those obtained for the EGFR–doxorubicin complex. Among the screened molecules, Eb-WE (Table 1) showed

an interaction energy to EGFR close to that obtained for the EGFR–doxorubicin complex. Thus, this Ebselen derivative may have anticancer potential, but of course, experimental validation is needed to confirm these computational data [153].

Computational techniques have been used to design peptide inhibitors of the T-cell CTLA4 (Cytotoxic T-Lymphocyte–Associated antigen 4) receptor [162]. CTLA4 is one of the receptors expressed on immune cells and is involved in the immune self-tolerance pathways that are targeted in cancer immunotherapy. Indeed, immune cells can intervene in the tumor microenvironment by recognizing and destroying antigens specific to cancer cells. However, cancer cells can evade these self-tolerance mechanisms by involving inhibitory receptors expressed on immune cells, such as CTLA4, that, together with their ligands, are targeted by immunotherapy. Peptides might work better as inhibitors of these receptors with respect to antibodies that are more expensive and difficult to store. The X-ray structure of the CTLA4 receptor, extracted from the complex with the B7-2 protein (PDB ID: 1I85 [154]) (Table 1), was used as a template to generate peptide sequences by means of *in silico* techniques (including flexible docking and MD simulations). First, the residues mainly involved in the CTLA4/B7-2 complex formation were predicted using the KFC (Knowledge-based FADE and Contacts) web server [155] and, consequently, a B7-2 region to use as a peptide template was identified. The structure of the template peptide was predicted with the Phyre (Protein Homology/analogY Recognition Engine) 2 web server [156] and used in flexible docking runs against CTLA4 through Rosetta FlexPepDock [157]. A few of the protein–peptide models generated during this step were selected and structurally optimized with PyRosetta [158]. Peptide sequences were then head-to-tail cyclized to improve their resistance to protease degradation, and their stabilities in the CTLA4-bound forms were predicted using MD simulations by the NAMD 2.13 program [163] and MM-GBSA binding-free energy calculations [160]. The best-performing protein–peptide complexes were further submitted to an MD free energy calculation in explicit solvent by using the BFEE (Binding Free Energy Estimator) module of VMD (Visual Molecular Dynamics) [161], and the resulting best models were further implemented in flexible docking to achieve an additional optimization cycle.

The peptides were also experimentally tested by means of bio-layer interferometry to confirm their ability to bind CTLA4 and by *in vitro* and *in vivo* experiments in Lewis lung carcinoma (LLC) cells and in orthotopic Lewis lung carcinoma allograft models to verify their capacity to block tumor cell growth. One of the peptides, the cyclic EI-DTVLPTGWVAKRYS, was identified as the most promising immunomodulator in cancer therapy [162].

Another protein involved in the inhibition of the immune response against cancer cells, HVEM (herpesvirus entry mediator), was the target of a study in which linear and cyclic peptides were designed as antagonists of its interaction with BTLA (B- and T-lymphocyte attenuator). Experimental studies and computational techniques were nicely combined within this study. For peptide design, the N-terminal fragment of the gD (glycoprotein D) protein, that binds HVEM and allows for the starting of infections by herpes simplex virus-1 and -2 through a variety of entry mechanisms, was employed. Structure-based drug design and MM-GBSA analysis were implemented as computational tools to select lead compounds to be experimentally validated through diverse experimental analyses [including ELISA (Enzyme-Linked Immunosorbent Assay) and cellular-based reporter assays]. For the best peptide candidates, docking simulations were also performed to predict the protein–peptide binding topology. One of the peptides, provided with a cyclic arrangement and a disulfide bridge (i.e., gD(1–36)(K10C-T29C)), was considered the best candidate to block BTLA inhibitory function and increase the immune response towards cancer cells [164].

In our laboratories, we have long been studying the tyrosine kinase receptor EphA2 and employing computational approaches to identify peptide inhibitors of its interaction with the lipid phosphatase Ship2 (SH2 domain-containing inositol phosphate 5-phosphatase 2) [124]. The EphA2 receptor is overexpressed in several cancer cell types,

with recognized roles in cancer onset and progression. The EphA2 cytosolic Sam domain binds to the Sam domain of the lipid phosphatase Ship2, and this interaction is responsible for the inhibition of EphA2 receptor endocytosis and degradation and is mainly linked to pro-oncogenic outcomes. Different computational-based drug design approaches were implemented to develop virtual peptide libraries focused on Sam domains that were subsequently employed in docking-based virtual screenings to identify peptide inhibitors of the EphA2-Sam/Ship2-Sam complex with possible anticancer activity [122–124].

ACPs can also be identified from virtual peptide libraries with computational techniques not including molecular docking. For instance, a peptide from the milk protein alpha-lactalbumin with cytotoxic activity on A549 lung cancer cells was identified using machine learning methods as a virtual screening approach to identify peptides with relevant physicochemical and anticancer properties [165].

3.3. Antimicrobial/Antiviral Peptides

Antimicrobial resistance (AMR) is a serious warning to worldwide health. The excessive use of antibiotics in farms as an addition to animal feed and in humans to treat bacterial infections are among the key factors contributing to the diffusion of AMR, which mainly affects clinical fields related to surgery, transplantation, and the treatment of inflammations [166]. Viral agents also can develop AMR to vaccines and antiviral therapeutics, and this concern is particularly worrying if associated with the pandemic risk arising from some viral pathogens [167]. The most recent COVID-19 pandemic is a case in which AMR has been among the highest priorities for public health.

These issues point out the necessity for drug discovery campaigns to find new therapeutics able to overcome AMR. Antimicrobial peptides (AMPs), with the ability to perturb pathogen membranes or to regulate pathogen or host protein biological pathways, could be a good source of new antibacterial and antiviral therapeutics.

3.3.1. Introduction

AMPs can be found in all living organisms as they are part of the host defense machinery against pathogens. AMPs can directly or indirectly kill pathogen cells, by modulating immunity pathways in response to pathogen entry into host cells [168]. The biological activities of AMPs depend on their physicochemical properties, such as the net charge, the conformation, and hydrophobicity. These peptides are usually positively charged, with a length between 10 and 50 amino acids, and characterized by a variety of conformations (α -helical, β -sheet, or linear without specific secondary structure elements) [169]. Different models to explain the mechanisms that AMPs employ to disrupt cell membranes have been proposed, such as the barrel-stave, the carpet-like, the toroidal pore, and the detergent-like models [168]. AMPs are generally considered in the context of anti-bacterial treatments, but the Antimicrobial Peptide Database (APD) [170] contains 3569 antimicrobial peptides from six life kingdoms (<https://aps.unmc.edu/>, accessed on 15 November 2023) and 207 of them are classified as antiviral peptides. Viral AMPs can act as well by different mechanisms, such as the inhibition of viral infection by preventing interaction with host cells or by inhibiting the transcription and translation processes of the viral genes [168]. As for ACPs, databases [170,171] and specific computational tools, including machine learning methods [171–174], have been developed to help the scientific community in the drug discovery field and to understand the mechanisms of action of AMPs.

The next paragraphs will focus on several studies related to the search using virtual screening and additional *in silico* tools of antiviral peptide agents against SARS-CoV-2.

3.3.2. Case Studies: Targeting SARS-CoV-2

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has been responsible for 6,985,964 deaths worldwide since the onset of the coronavirus disease 2019 (COVID-19) pandemic till the beginning of December 2023, thus justifying the many efforts made for the continuous development of novel treatments against this disease [175].

SARS-CoV-2, its mechanism of action, and many computational tools available for the development of small molecule ligands against viral proteins have already been described in several previous reviews [176–182]. Briefly, SARS-CoV-2 belongs to the enveloped viruses, including a positive-sense single-stranded RNA, and its genome encodes different viral proteins supporting its infective action [176]. For instance, the SARS-CoV-2 spike protein binds to the human receptor angiotensin-converting enzyme 2 (ACE2) through a receptor-binding domain (RBD), thus allowing viral entry into host cells [176,183–190]. In addition, the 3C-like protease (3CL_{pro}), also known as main protease (M_{pro}), is a viral protein with a key role in the processing of the polyprotein precursors pp1a and pp1b into the functional proteins necessary for viral propagation [176,191–194].

The search for the peptide inhibitors of SARS-CoV-2 proteins through *in silico* approaches can be achieved following diverse canonical key stages and combining docking studies and molecular dynamics simulations [183].

Several *in silico* studies concerning the inhibition of viral entry into host cells by blocking the interaction of the RBD domain of the SARS-CoV-2 spike protein (spike RBD) with the human ACE2 receptor have been reported in the literature (Figure 2a) [183–185,187,195]. In one such study, the authors retrieved from the PDB several structures of the RBD from the SARS-CoV-2 spike glycoprotein/ACE2 complex (i.e., PDB ID: 6M0J [196], 7C8D [197], and 7A95 [198]) and deeply analyzed them also through the MM-GBSA approach. MM-GBSA was used to evaluate the per-residue contribution to the binding free energy and predict the crucial amino acids at the binding interface, also accounting for the reproducibility of results in all the considered structures [195]. This investigation revealed that the interacting residues were mainly positioned within two ACE2 regions (i.e., the α 1 helix and the β 4- β 5 sheets, Figure 2a). Next, the gained information was employed to design a 49-mer peptide (SARS-CoV-2 PEP 49) encompassing the amino acids of both the α 1 and β 4- β 5 regions of ACE2, the 3D structure of which was predicted with PEP-FOLD3 [140] and implemented in docking studies against the SARS-CoV-2 spike protein. The docking results pointed out that the designed peptide could target the RBD domain of the spike protein with a better interaction energy than ACE2 itself. In the end, the *in silico* data showed that PEP49 could potentially work as a good inhibitor of viral entry and consequently be implemented to design original antiviral therapeutics [195].

Nevertheless, a strategy called EvoDesign was employed to generate *in silico* several peptides targeting the spike RBD to block viral entry into host cells [199]. Starting from the analyses of a crystal structure of the SARS-CoV-2 spike RBD/ACE2 complex (PDB ID: 6M0J [196]), a peptide scaffold was first built by combining, through a glycine linker, two segments extracted from the ACE2 receptor interface. This initial scaffold was used as a starting point in bioinformatic tools for *de novo* protein design (i.e., EvoEF2 [200] and EvoDesign [201]) to generate novel peptide sequences optimized for interaction affinity towards the SARS-CoV-2 spike RBD. Twelve of the best peptides, according to the *in silico* binding studies, were also subjected to another computational screening. In this secondary screening, MD simulations of the spike RBD/peptide complexes were carried out to analyze the binding affinities and stability, looking at diverse features like the hydrogen bonding aptitude and RMSD (Root Mean Square Deviation) values. The results, combined with the predictions of the peptide secondary structures and stabilities in aqueous solution, led in the end to the choice of the best peptide candidates (e.g., Peptide 6 and Peptide 7 in Figure 2a) that could be conjugated to a graphene sheet or a carbon nanotube to generate a bio-sensor for SARS-CoV-2 detection [184].

In a similar computational approach, the binding of the SARS-CoV-2 spike RBD to ACE2 (PDB ID: 6LZG [202]) was investigated through MD and MM-PBSA free energy calculations to establish the key residues at the binding interface of the two proteins (Figure 2b). Based on the hot spot residues identified on ACE2 (i.e., D30, E37, D38, and Y41), a library of 6-mer peptides (*consensus* sequence: “DX1X2EDY” where X1 and X2 represent any among the 20 natural amino acids) was designed. Through docking-based virtual screening against the spike RBD, six peptides were selected and further validated

by MD. Finally, three peptides (DDFEDY, DEYEDY, and DFVEDY in Figure 2b) resulted as the best candidate inhibitors of viral entry and were further analyzed by bioinformatic tools to estimate the allergenicity, toxicity, and solubility [185].

An additional *in silico* approach to attack viral infectivity by blocking the entry stage started from a diverse virtual library of peptides, but again exploited the concept that a peptide somehow similar to the ACE2 sequence responsible for interaction with the spike protein might work as an inhibitor of the spike RBD/ACE2 complex and thus block viral infection. In this context, a peptide library was assembled by mutating ACE2 residues not involved in binding to spike RBD according to the analyses of experimental structures. In detail, starting from a peptide fragment encompassing the α 1 helix in ACE2, mutations were inserted at 12 non-interacting amino acid sites; the introduction of point and multiple mutations led to the generation of 136 peptides. Docking-based virtual screening against the SARS-CoV-2 spike RBD (Figure 2a) was conducted, and seven peptides with potential high binding affinities could be predicted. The MD simulations further pointed out one potential peptide inhibitor (peptide 13 in Figure 2a) of viral entry with a ~3-fold increased interaction affinity for spike RBD with respect to a reference ACE2 peptide encompassing the α 1 helix [187].

Although these studies [185,187,195] provide useful insights for the design of potential peptide inhibitors of viral infectivity modulating the viral entry stage, the lack of experimental validation remains a major drawback. On the contrary, an elegant study by Hu and collaborators reported on the successful *in silico* identification and experimental validation of a peptide able to reduce SARS-CoV-2 entry into host cells by targeting simultaneously the RBD of the SARS-CoV-2 spike protein and human neuropilin-1 (NRP1) [186]. The NRP1 receptor represents another important player in SARS-CoV-2 entry as it has been proposed as a co-receptor that might adjuvate ACE2 to favor virus attachment to olfactory and respiratory epithelial cells, thus enhancing infectivity [186,203]. The NRP1 extracellular b1 domain is generally used as the target to find NRP1 inhibitors; furin is a host protease that cleaves the SARS-CoV-2 precursor spike S protein into the S1 and S2 fragments and consequently generates a basic stretch at the C-terminal end of S1 (i.e., the C-end rule (CendR) motif) responsible for interacting with the NRP1 receptor [204]. To identify peptides able to bind the spike RBD and NRP1 b1 domain simultaneously, virtual screening was conducted by employing both pharmacophore-based docking and structure-based docking. The software Molecular Operating Environment (MOE [205], Chemical Computing Group Inc, Montreal, Quebec, Canada) was used for the pharmacophore modeling. A library composed of 24,000 peptides was built by the QuaSAR-CombiGen module of MOE through the random creation of connections between fragments of different lengths (i.e., 4-mer, 7-mer, 9-mer, 12-mer) [186]. The resulting 2D peptide models were converted into 3D structures; meanwhile, the crystal structure of ACE2 in complex with the SARS-CoV-2 spike RBD (PDB ID: 6MOJ [196]) and the crystal structure of the human NRP1 b1 domain in complex with the SARS-CoV-2 S1 CendR peptide (PDB ID: 7JJC [204]) (Figure 2a,c) were downloaded from PDB, hydrogens were added, and the protonation state optimized. Next, the MOE program was employed to deeply analyze the interactions between the SARS-CoV-2 spike RBD and ACE2 from the crystal structure [186]. The resulting data were exploited by the Pharmacophore Editor module of MOE to provide the pharmacophore profile for spike RBD (i.e., an ensemble of characteristics that peptides should have to match the features of spike RBD binding pocket). A pharmacophore-based docking simulation was conducted to identify potential peptide ligands of spike RBD through the docking module of MOE. In detail, to screen the peptide database through pharmacophore-based docking, the pharmacophore model was used as a 3D query, while the binding site included hot spot residues on the interaction surface of spike RBD. The binding free energies for the spike RBD/peptide complexes were evaluated based on the docking scores. The peptides provided with docking scores lower than a chosen cut-off (i.e., -13.5 kcal/mol) were further assayed through structure-based docking into the NRP1 b1 active site, and the top five peptides were selected for *in vitro* biological testing. This clever approach was

conducted for the identification of one peptide (i.e., RN-4, Figure 2c) able to simultaneously bind both the spike RBD and NRP1 b1. In *in vitro* tests, RN-4 showed nanomolar affinities against both proteins, and in the pseudovirus infection assay, the capacity to significantly lower the SARS-CoV-2 entry into cells without provoking substantial adverse effects was also demonstrated [186].

To discover peptide ligands of the spike RBD that could reduce SARS-CoV-2 infectivity by blocking viral entry, a work by Ramirez-Acosta and collaborators reported on a virtual screening strategy employing antimicrobial peptides [190]. A library of 104 peptides was assembled starting from the Antimicrobial Peptide Database (APD) of the University of Nebraska Medical Center [170,206] by focusing on those peptides with known antiviral activities [190]. Since lysozyme is an innate immune system component with antimicrobial activity, three lysozyme fragments were inserted into the library [190]. In addition, for comparison purposes, an ACE2-derived peptide (i.e., "IEEQAKTFLDKFNHEAEDLFYQSS") was included in the screening dataset. The structures of the preselected peptides that could not be retrieved from the PDB were predicted by different sequence-based *in silico* tools (I-TASSER (Iterative Threading ASSEMBLY Refinement) [207,208] and PEP-FOLD 3.5 [140]) and then used as input for a docking-based virtual screening against the spike RBD [190]. The docking analysis predicted for the twelve antimicrobial peptides and the three lysozyme-derived peptides better binding affinities for the spike RBD than that predicted for the ACE2 reference peptide. Thus, after analysis of the most recurrent amino acid types interacting with the RBD and based on H-bond pattern, extra optimized peptides were designed and employed in docking studies against the spike RBD. The ability of the peptides to escape the immune system was also investigated computationally through the TepiTool, and for a few of them, a low immunogenic response was predicted [209]. In the end, a few optimized peptides with enhanced predicted binding affinity towards diverse RBD variants (including the Delta one and a theoretic form derived by combining the Alpha, Beta, and Gamma variants) were identified [190]. This represents a promising protocol for discovering antiviral agents.

Thakkar and colleagues reported instead on the development, through the combination of protein design programs and molecular dynamic simulations, of a 17-mer stapled peptide (pep39) targeting the spike RBD and blocking its interaction with human ACE2 [210]. The initial design started from a template peptide encompassing residues 26–42 of the ACE2 α 1-helix and from the structure of the human ACE2/spike RBD complex (PDB ID: 6LZG [202]) (Figure 2b). The structure-based FlexPepDock protocol [157] was implemented to build a peptide that is able to interact with the spike RBD. Following this strategy, a wide search of the conformational space accessible to the peptide backbone was achieved using FlexPepDock. Next, docking poses provided with both low energies and a certain degree of structural similarity with the template peptide were selected. To enhance the affinity of the chosen poses for the spike RBD, the protocol proceeded by substitutions with the common 20 amino acids and the optimization of the sidechains and rotamers. Several optimization cycles led to many diverse possible sequences, but certain amino acid residues appeared to be preferred in specific peptide positions: proline and glycine at positions 8 and 10, respectively, and hydrophobic amino acids at positions 2, 12, 13, and 14. The first 41 optimized peptide structures in complex with the spike RBD were subjected to molecular dynamics simulations. The MM-GBSA binding free energy calculations allowed for the further filtering of 9/41 optimized peptides. Moreover, to favor structure resemblance with the ACE2 α 1 helix, a propene staple and an amide bond between the side chains of a negatively charged residue and a positively charged one were introduced into *ad hoc* chosen positions to enhance the peptide helical structures. The MD simulations coupled to the MM-GBSA binding free energy calculations demonstrated the largest stabilities of the stapled peptide conformations, as well as their higher ability to remain in the bound state. The analysis of the MD trajectories demonstrated that the so-called "pep39" peptide outperformed and was able to interact well with the spike RBD in the ACE2 binding site. This most promising "pep39" peptide was experimentally vali-

dated. The interaction studies through bio-layer interferometry demonstrated the binding of pep39 to the spike RBD (dissociation constant $K_D = 570 \pm 50$ nM) and its delta variant ($K_D = 4.1 \pm 1.4$ μ M). In addition, the cell-based assays suggested that this stapled peptide was able to block SARS-CoV-2 replication and could function as a potential anti-COVID-19 therapeutic agent [210].

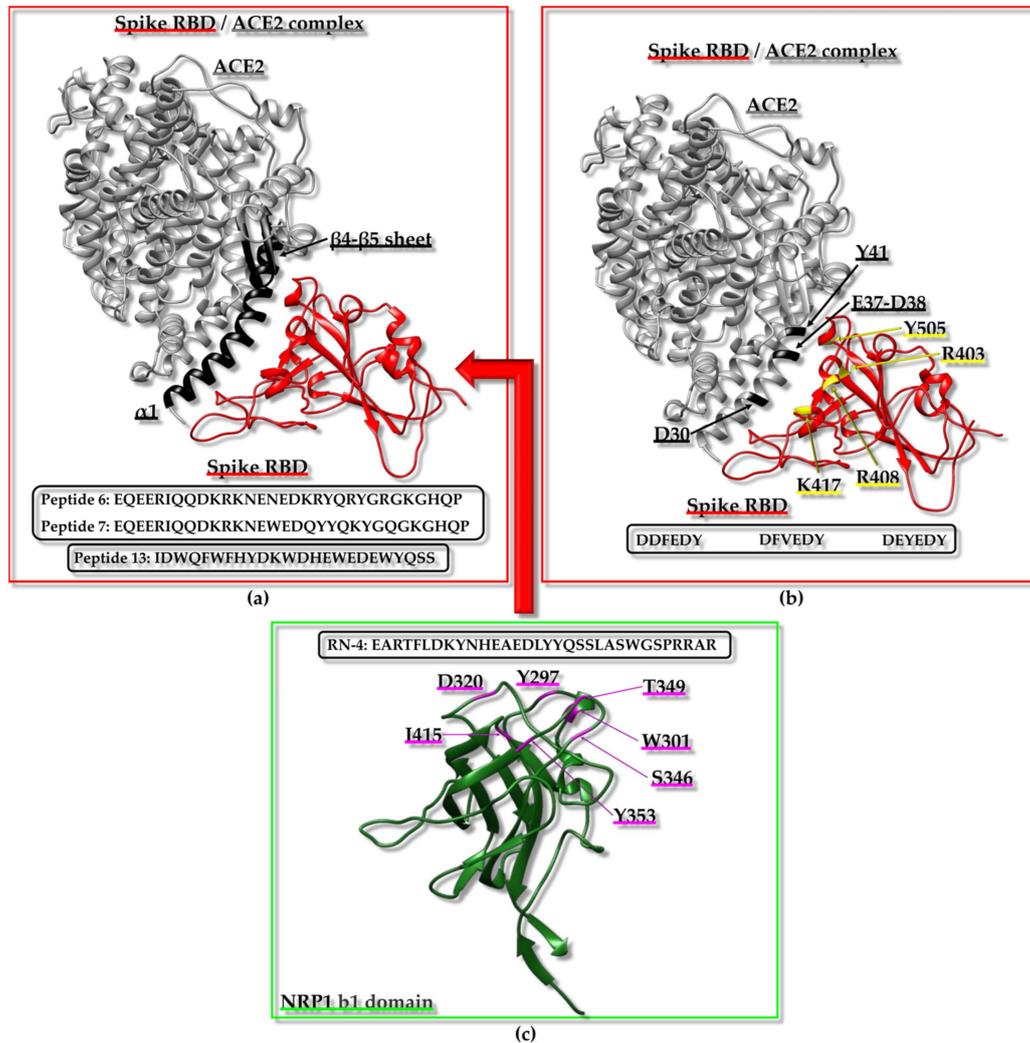


Figure 2. In silico approaches to discover antiviral peptides targeting the receptor-binding domain of SARS-CoV-2 spike protein (spike RBD) (a,b) and neuropilin 1 (NRP1) receptor (c). (a) The crystal structure of spike RBD in complex with ACE2 receptor peptidase domain (PD) (PDB ID: 6M0J [196]) is shown in a ribbon representation where the spike RBD is colored in red and ACE2 in grey with regions important for the interaction (i.e., $\alpha 1$ helix and $\beta 4$ - $\beta 5$ sheets) highlighted in black. Peptides 6, 7, and 13, the sequences of which are reported in the one-letter amino acid code, were identified through in silico binding studies of the spike RBD structure extracted from the shown PDB entry [184,187]. (b) X-ray structure of spike RBD (red)/ACE2 (grey) complex (PDB ID: 6LZG [202]). Hot spot residues on ACE and spike RBD are colored in black and yellow, respectively. The sequences of a few RBD spike-targeting peptides discovered by docking-based virtual screening are shown [185]. (c) Crystal structure of NRP1 b1 domain (green) (PDB ID: 7JJC, chain A) extracted from the complex with the CendR peptide [204]. Residues of the NRP1 peptide binding pocket are labeled and colored in magenta. The RNF4 peptide is an in silico-identified dual ligand of spike RBD and NRP1 b1 domain [186].

A diverse computational strategy relying on the combination of docking analysis and MD simulations provided four dual peptide inhibitors targeting both the spike RBD and the Mpro protease (Figure 3a). The screening of a peptide library was achieved through the docking module in the MOE [205] software. The crystal structures of both the viral protease Mpro complexed with a cyclic peptide (PDB ID: 7RNV [211]) and spike RBD in complex with ACE2 (PDB ID: 6M0J [196]) were implemented in this study. A virtual peptide library was combinatorially built using the QuaSAR-CombiGen module of MOE starting from diverse peptide fragments (i.e., 15-mer cyclic peptides, heptapeptides, and 16- and 18-mer linear peptides) and included 27,000 cyclic peptides made up of an α -helix and a cyclic segment [188]. The α -helix was inserted considering that this structure element in ACE is important for the interaction with the spike RBD (Figure 2). All the peptides were first screened against Mpro, and the best hits (i.e., docking scores lower than -13.7 kcal/mol) were docked towards the spike RBD. The best four dual ligands were further validated using MD simulations to analyze the stabilities of the protein-peptide complexes. Finally, the four *in silico* hits were experimentally validated. First, the MST (Microscale thermophoresis) interaction assays provided evidence of peptide binding to both the Mpro and spike RBD with dissociation constants in the nanomolar range and an increase in the binding affinity compared to two other peptides used as positive controls [188]. In addition, a pseudovirus infection assay showed that more than half of the SARS-CoV-2 pseudovirus was inhibited by these peptides without considerable cytotoxicity to the host cells [188].

As mentioned before, the protease Mpro represents another appealing target in antiviral drug discovery to counteract SARS-CoV-2 infection. To target Mpro, a virtual peptide library was designed considering that D-amino acids provide some advantages when incorporated into a peptide sequence (i.e., a higher resistance to proteases, improved intestinal absorption upon oral administration, and low or absent immunogenicity), and that small peptides (i.e., 3-mer and 4-mer) perform better in docking algorithms since they possess less freely-rotatable bonds [189]. A virtual screening was conducted against the crystallographic structure of free Mpro (PDB ID: 6Y2E [212]) (Figure 3a) [189]. A library of D-tri- and tetra-peptides was built through specific commands of Amber20 [213]; the peptides were assembled by considering combinations of the common 21 residues (i.e., the 20 natural amino acids, and considering both neutral tautomers of histidine), and including the N-terminal acetyl and C-terminal N-methyl amide capping groups. Starting from the natural amino acids, the L-configuration was inverted to the D-configuration for each residue, but for Ile and Thr, modeling was performed to obtain D-allo-isoleucine and D-allo-threonine diastereomers. The 3D structures of each peptide were generated, and were energy minimized [189]. Then, structure-based virtual screening was first conducted by docking the D-peptide database in the Mpro active site; the resulting predicted Mpro/peptide complexes were subjected to rescoring through the MM-GBSA approach and the evaluation of binding free energies, followed by MD simulations and the analysis of different parameters [i.e., RMSD, RMSF (root mean square fluctuation), and the number of H-bonds] [189]. The proposed protocol provided four D-tetrapeptides (i.e., 4P1, 4P2, 4P3, and 4P4 in Figure 3a) that were evaluated in *in vitro* enzymatic assays and demonstrated the capacity to block from 50 to 85% of the Mpro activity, thus proving that the implemented computational strategy could successfully be conducted to identify novel anti SARS-CoV-2 agents functioning as inhibitors of the Mpro protease [189].

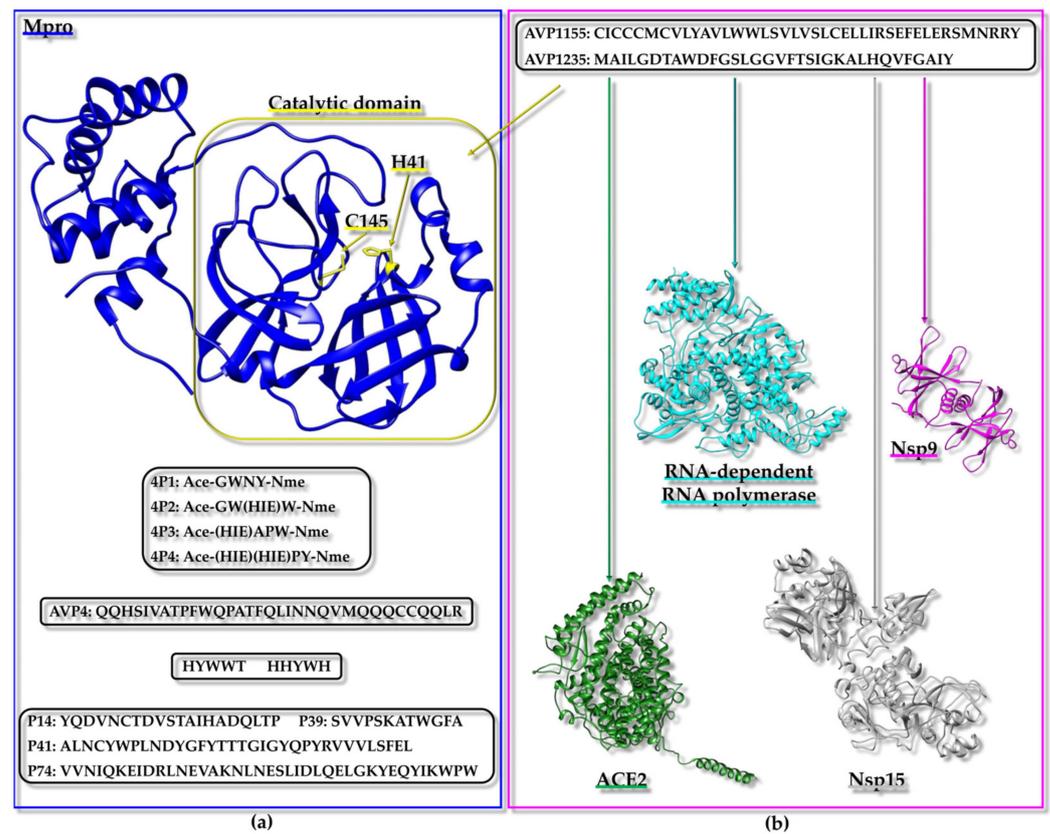


Figure 3. (a) The ribbon representation of free Mpro (crystal structure with PDB ID: 6Y2E [212]) is reported in blue, the catalytic domain is highlighted with a yellow box, and the side chains of the residues (Cys145 and His41), forming the catalytic dyad, are shown. The amino acid sequences of several in silico identified potential peptide inhibitors of Mpro are indicated [189,191,192,214]. The “4P” peptide series contains only amino acids in D-configuration; Ace stands for the N-terminal Acetyl group, Nme stands for N-methyl amidation at C-termini, and HIE represents the E tautomer of histidine [189]. (b) Electron microscopy (EM) structure of full-length ACE2 receptor (green) (extracted from the RBD/ACE2-B0AT1 complex PDB ID: 6M17 [215], chain D) and diverse SARS-CoV-2 proteins: RNA-dependent RNA polymerase (cyan) (extracted from the EM structure of SARS-CoV-2 RNA-dependent RNA polymerase in complex with cofactors with PDB ID: 6M71 [216], chain A); Nsp15 (grey) (crystal structure—dimeric form—PDB ID: 6VWW [217], chains A and B); Nsp9 (magenta) (crystal structure -dimeric form- PDB ID: 6W4B [218], chains A and B). Antiviral peptides AVP1155 and AVP1235 [219] were identified through a multi-target in silico screening against all proteins in panel (b) as well as Mpro (a).

In order to find Mpro inhibitors, another study aimed to generate a peptide library by the in silico digestion of the four more abundant proteins in rice bran (i.e., albumin, glutelin, globulin, and prolamin) by the proteases pepsin, trypsin, and chymotrypsin [191]. Then, the results were analyzed by diverse bioinformatic tools to predict potential antiviral/antimicrobial peptides (several programs were implemented: AVPpred [220], MetaAVP [221], AMPfun [222], and ENNAVIA (Employing Neural Networks for Antiviral Activity) [223]). The peptide amino acid composition and cytotoxicity were also estimated by the COPid (COMposition based Protein Identification) [224] and ToxinPred [225] tools. Through all these analyses, 10 antiviral peptides were chosen, and their 3D structures were predicted by PEP-FOLD 3 [140] and used as the input for docking analyses conducted through the GalaxyPepDock [70] program against the crystal structure of Mpro (PDB ID: 7C2Q [226]) [191]. This stage indicated that all peptides target a region of Mpro near its active site, and the best ligand, according to the docking scores, was the AVP4 peptide (Figure 3a). The Mpro/AVP4 peptide complex was further analyzed by MD simulations

that pointed out a prolonged and robust interaction, thus confirming the validity of this virtual screening approach [191].

As mentioned in Section 2.1, genetic algorithms can be employed to generate peptide libraries and lower the high computing power usually requested to handle libraries made up of tens of thousands of components [192]. In a computational strategy to find Mpro inhibitors, the Mpro structure (PDB ID: 6LZE [227]) was employed to discover high-affinity peptide ligands by means of a genetic algorithm. A group of peptide sequences representing the populations was evolved through the genetic algorithm by using docking scores so that several iterations were carried out, and each time, a new population was generated from the previous one by sequence pairing, thus leading to better-predicted peptide binding affinities according to the docking scores. In other words, the implemented algorithm applied a sort of “Darwinian selection” in which the accomplishment of lower docking scores against SARS-CoV-2 Mpro was used as a criterion for evolutive steps [192]. The approach started from an initial population including 19 5-mer peptides made up by the repetition of a single residue (consensus sequence “XXXXX”, where X is the same single amino acid, chosen from the 20 natural ones, excluding proline) [192]. To achieve greater levels of sequence diversity by ensuring random pairing, the same score was attributed to each member of the starting population [192]. Starting from the second iteration, the resulting couples were subjected to a crossing-over process, during which the 50 sequence pairs with the best docking scores were saved and selected for a roulette wheel pairing step in which the new population of sequence couples is built for the next evolutive step [192]. In addition, each cycle included the removal of the worst ten sequences [192]. The protocol led to two peptides (i.e., HHYWH and HYWWT, Figure 3a) that could interact with SARS-CoV-2 Mpro with an affinity greater than those against human proteases, thus indicating the proposed strategy as a valid tool for the development of a starting lead peptide for the design of optimized potential antiviral agents [192].

As previously mentioned, an intriguing source of potential inhibitors of SARS-CoV-2 viral proteins can be supplied by the database AVPdb, which includes peptides the antiviral roles of which have been experimentally proven [194]. In addition, for the peptides included within this database, useful features are provided (like sequence, source, target virus, virus family, efficacy from the qualitative and/or quantitative point of view, assays used to establish the efficacy, the physicochemical characteristics, and PubMed references) along with a predicted structure [194]. A list of 88 antiviral peptides was extracted from AVPdb based on anti-SARS-CoV-2 activities [214], and their 3D structures were modeled by PEP-FOLD 3 [140]. These peptides were virtually screened against the X-ray structure of Mpro (PDB ID: 6LU7 [228]). The CASTp webserver [151] was employed to clearly identify the binding pocket on the surface of Mpro. Next, docking was carried out first with PatchDock [229], and then further docking runs were performed with FireDock (Fast interaction refinement in molecular docking) [230] starting from the best 10 conformations of each Mpro/peptide complex; to additionally validate the results, the docking was also conducted with ClusPro [231]. Four peptides (i.e., P14, P39, P41, and P74 in Figure 3a) scored as the best ligands, according to several docking tools, and their complexes with the Mpro active site were subjected to MD simulations to study the conformational stability [214]. Different MD-related parameters (i.e., the RMSD, solvent accessible surface area, radius of gyration, RMSF, and number of H-bonds) indicated the higher stability of the protein–peptide complexes even with respect to Mpro in the apo form [214].

In a second work [219], a set of 434 AVPs were extracted from AVPdb [194] following three main criteria: 1—the exclusion of peptides with proven activity against viruses of the coronaviridae class, as the goal was to discover original anti-SARS-CoV-2 agents, 2—a lack of cytotoxicity, and 3—a length of at least 27 residues to allow for the efficient modeling of 3D structures by the Robetta webserver [232]. The selected antiviral peptides were virtually screened by docking against both the SARS-CoV-2 Mpro (PDB ID: 6M03 [233]) and the non-structural protein 9 (Nsp9) (PDB ID: 6W4B [218]) with the ClusPro webserver [231] (Figure 3a,b). From this first screening, eleven peptides with the most promising docking

scores (based also on a comparison with reference peptides) towards both targets were selected. The eleven peptides were next employed in a second docking-based virtual screening approach against other SARS-CoV-2-related targets (Figure 3b) [i.e., the ACE2 receptor-binding domain (PDB ID: 6M17 [215]); SARS-CoV-2 RNA-dependent RNA polymerase (PDB ID: 6M71 [216]), and Nsp15 endoribonuclease (PDB ID: 6VWW [217])] [219]. From the second screening, two peptides (1155 and 1235 in Figure 3b) were selected with good scores against all the proteins. In the end, the complexes of all five of the target proteins and the two best peptides were further analyzed by MD simulations to confirm the structural stability of the protein–peptide complexes [219].

In summary, during the COVID-19 pandemic, and especially during the initial outbreak, *in silico* approaches to fight SARS-CoV-2 flourished due to the urgent need to find therapeutic agents. Many putative peptide inhibitors of SARS-CoV-2 proteins have been identified *in silico* but not experimentally validated. The confirmation of the antiviral activities of these peptides could open the road to the establishment of novel therapeutics, avoiding undesired side effects.

3.4. Inhibiting Fibril Formation: $A\beta_{42}$ and hLAPP as Targets

Virtual screening strategies have also been designed to identify peptides able to inhibit protein self-aggregation, which is often the cause of pathological events [234–236]. AD (Alzheimer’s disease) represents a neurodegenerative disorder due to accumulation in the extracellular space of plaques. Neuritic plaques found in the brains of AD patients are mostly composed of $A\beta$ peptides resulting from the proteolytic cleavage of amyloid precursor proteins that form neurotoxic fibrillary β -sheet structures [237]. $A\beta$ -(1–42) (= $A\beta_{42}$) represents the major form found in plaques. It has been reported that monomeric $A\beta_{42}$ might initially assume an α -helical or disordered state before switching into a β -sheet conformation. This conformational change represents the crucial stage in the $A\beta$ fibrillogenesis process (Figure 4a,b) [238].

Starting from a previously identified inhibitor of $A\beta$ aggregation and toxicity (i.e., the “RIIGL” peptide), a virtual peptide library was designed. In detail, this library was built by substituting diverse amino acid positions of the starting peptide with residues provided with similar side chains. For example, the first “R” in the “RIIGL” peptide was mutated in a few polar and basic amino acids (K, H) while the “I” was substituted with some non-polar aliphatic residues. The final library included 912 5-mer peptides that were virtually screened, through docking techniques, against $A\beta_{42}$ in the monomeric form (PDB ID: 1IYT [237]) [234]. The best 10 peptide ligands, according to the docking scores, were next assessed through molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) to evaluate the free energy of the interaction with $A\beta_{42}$. Based on this analysis, the three best peptides (“RLAPV”, “RVVPI”, and “RIAPA”) (Figure 4a) were further investigated by MD simulations to study the conformational variations of $A\beta_{42}$ in the presence and the absence of each peptide. Interestingly, the MD revealed that in the presence of the “RVVPI” and “RIAPA” peptides, there was a higher stabilization of $A\beta_{42}$ helical states (Figure 4a) with respect to the β -sheet conformations characterizing instead the fibrillar species (Figure 4b). Thus, this clever computational strategy let us speculate that “RVVPI” and “RIAPA” could prevent the conformational switch of the $A\beta_{42}$ monomer towards aggregation-prone structures [234].

The same *in silico* protocol to find inhibitors of $A\beta_{42}$ self-association was tested starting from the β -breaker peptide “LPFFD” and the structure of $A\beta_{42}$ in the aggregated fibrillar state (Figure 4b) [235]. “LPFFD” is a known inhibitor of $A\beta$ aggregation that can dissociate fibrils *in vitro* but also decrease fibrillogenesis and $A\beta$ deposition in models of rat brains. A virtual library of 867 pentapeptides was assembled by incorporating mutations in the starting β -breaker peptide. Virtual screening by docking was first performed using as a target the structure of $A\beta_{42}$ in the fibrillar form (PDB ID: 2BEG [239] Figure 4b), and, afterwards, MM-PBSA analyses and MD simulations were carried out. This approach selected the “PPFFE” peptide as the most promising β -breaker peptide able to interact with

the core region of the $A\beta_{42}$ protofibrils, favoring energy minima characterized by a lower β -sheet content and a decrease in the H-bonds [235].

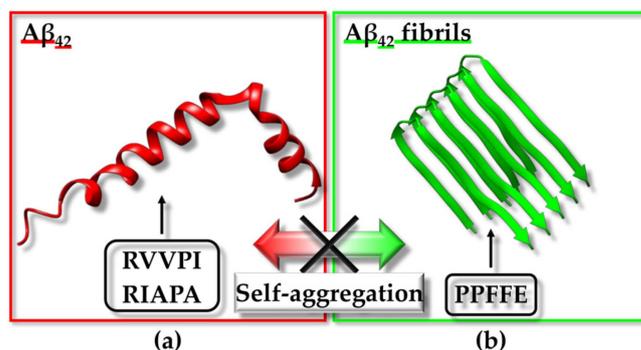


Figure 4. (a) NMR structure of $A\beta_{42}$ monomer (red) calculated in HFIP (Hexafluoroisopropanol/water) 80/20 *v/v*, (PDB ID: 1IYT [237], first conformer). RVVPI and RIAPA peptides [234] that were identified by vs. against the monomeric form of $A\beta_{42}$ could stabilize the $A\beta_{42}$ helical conformation. (b) NMR structure of $A\beta_{42}$ fibrils (green) in aqueous buffer (PDB ID: 2BEG [239], first conformer). The PPFEE peptide should work as a β -breaker and was selected by initial vs. against $A\beta_{42}$ fibrils [235].

A similar computational strategy was also employed to discover peptide inhibitors of hIAPP (Human islet amyloid polypeptide) self-aggregation that leads to cytotoxic fibers connected to the pathogenicity of type 2 diabetes (Figure 5a,b). Small peptide fragments encompassing the amyloidogenic region of hIAPP work as blockers of self-assembly. In particular, the fragment HSSNN_{18–22} was identified as an amyloidogenic sequence demonstrating elevated antiproliferative properties towards RIN-5F cells. Thus, the “HSSNN” peptide was used as a model to generate a virtual library of mutated peptides. This library was virtually screened against the structure of hIAPP in the monomeric state (PDB entry: 2L86 [240]) (Figure 5a). The subsequent MM-PBSA investigation of the best docking hits pointed out two peptides (HSSQN and HSSNQ) that were able to target monomeric hIAPP with a high affinity. Further MD simulations highlighted that when in a complex with each peptide, monomeric hIAPP underwent the enhanced sampling of helical conformation with a consequent decrease in the aggregation tendency [236].

In conclusion, the *in silico* approach combining several computational tools (i.e., docking-based VS, MM-PBSA, and MD) represents a suitable instrument to select the most promising peptides, potentially able to block the self-aggregation mechanisms related to AD and type 2 diabetes to be experimentally validated.

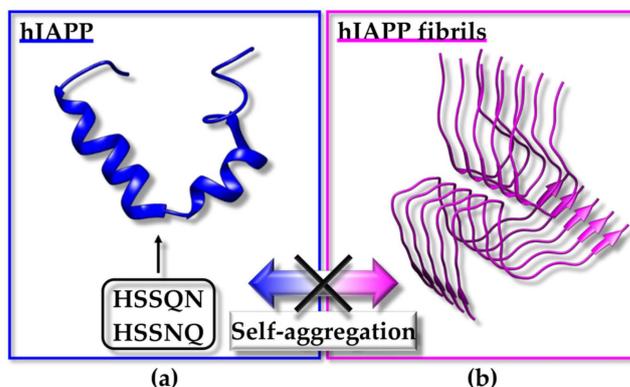


Figure 5. (a) NMR structure of hIAPP monomer (blue) calculated in SDS (Sodium Dodecyl Sulphate) micelles, (PDB ID: 2L86 [240], first conformer). The *in silico* identified HSSQN and HSSNQ peptides could favor the helical state in monomeric hIAPP [236]. (b) Cryo-EM structure of hIAPP fibrils (magenta) (PDB ID: 7M61 [241]).

4. Conclusions

This review can be considered a sort of toolkit for researchers who intend to start a drug discovery study using docking-based virtual screening approaches and peptide libraries to target proteins involved in different diseases.

Most functions within cells are mediated by networks of PPIs, the dysregulation of which is usually related to pathological conditions; thus, there is growing interest in the discovery of compounds able to modulate such interactions positively or negatively. PPIs usually occur through vast surfaces characterized by smoothness and are often unable to be attacked by small molecules. In this framework, peptides, that can also be easily produced by well-established synthetic routes, represent a valid alternative [2,3]. Nevertheless, in comparison to small molecules, peptides can be provided with a higher efficacy and specificity for their protein targets due to their chemical variety as well as conformational properties (i.e., the ability to assume a multiplicity of secondary structure elements along with the capacity to undergo conformational variations when binding to a protein surface) [2,3,242]. During the last few years, tremendous improvements have been made in the medicinal chemistry field related to the production of peptides with better drug-like features, and a variety of synthetic routes to prepare peptides provided with specific secondary structure elements have been set up [35,41]. However, the costs connected to the production of synthetic peptide libraries are relevant, especially if certain non-natural amino acids or cyclic organizations need to be included to increase stability and affinity for the target [242]. It also has to be considered that, when looking for a potential therapeutic peptide acting on a specific protein or PPI, the screening of very large peptide libraries, provided with the proper chemical diversity grade, needs to be carried out to improve the chance of finding a good hit. To perform the initial experimental screenings for evaluating peptide–protein interactions, additional costs, related to the chemicals and materials needed to run the binding assays, must be faced, as well as the costs connected to recombinant protein production (for example, when choosing to perform screening by NMR, ^{15}N and/or ^{13}C -labeled proteins are often required, but their production is more expensive). Such experimental screenings are also time-consuming.

However, when the target protein or protein–protein interaction has been well characterized from a structural point of view and high-resolution atomic coordinates can be downloaded from PDB, *in silico* approaches can be very useful in preselecting putative peptide ligands to be submitted to experimental validation, thus reducing the time and costs connected to a drug discovery campaign. Indeed, as described in the previous paragraphs, large libraries of virtual peptides can be generated through a variety of *in silico* tools, and diverse strategies can be adopted to design the library that best suits the system under study. These libraries can next be implemented in structure-based virtual screening [81,110,112,186,243]. To predict binding poses, peptides can be docked inside the protein binding pockets extracted directly from the complex structure, if available, or alternatively, diverse bioinformatic tools can be implemented to evidence pockets and grooves and establish a putative binding site on the protein surface [151]; even blind docking can be performed, but, in this case, less accurate results might be obtained and longer computational times, connected to the exploration of the whole protein surface, are required [244]. Another aspect to be considered is that, when performing docking-based virtual screening, flexibility is important, and most docking tools allow for a certain treatment of flexibility not only concerning the ligand but also the protein receptor [76,80,82,245]. In the end, docking scores can be employed to preselect a certain number of compounds, but, as selection criteria may not be straightforward enough to be established [246], before starting experimental validations, it can be helpful to further perform MD simulations to obtain some insights into the stability of the predicted docking poses [92,184,214].

All these topics and the major drawbacks have been treated within this review simply to allow the less experienced researchers to gain inspiration for their studies.

Through this review, we describe as well diverse examples of virtual screening approaches to identify peptides targeting a variety of proteins and PPIs involved in cancer, Alzheimer's disease, diabetes, and COVID-19 [153,187,192,234].

By reviewing the literature, we noticed that many putative peptide ligands have been identified in silico, but often the proper experimental validation of computational results is lacking, and in the end, it is important to consider that a large slice of virtually identified hits might not be able to reach the desired biological effects [81,247]. As much research effort is being spent in the field, and rapid improvements are being achieved, an even larger increase in the success rate of SBVS must be expected in the near future. Meanwhile, it could be interesting to create databases of in silico discovered bioactive compounds and validate those without experimentally proved functions. These data could open additional avenues for generating therapeutic agents and/or setting up novel therapeutic routes.

Author Contributions: Conceptualization and supervision, M.L.; writing and reviewing, M.L., F.A.M. and M.V.; figure generation, M.V. and F.A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created within this review paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gasbarri, C.; Rosignoli, S.; Janson, G.; Boi, D.; Paiardini, A. Prediction and Modeling of Protein-Protein Interactions Using "Spotted" Peptides with a Template-Based Approach. *Biomolecules* **2022**, *12*, 201. [[CrossRef](#)] [[PubMed](#)]
2. Bruzzoni-Giovanelli, H.; Alezra, V.; Wolff, N.; Dong, C.Z.; Tuffery, P.; Rebollo, A. Interfering peptides targeting protein-protein interactions: The next generation of drugs? *Drug Discov. Today* **2018**, *23*, 272–285. [[CrossRef](#)]
3. Hashemi, Z.S.; Zarei, M.; Fath, M.K.; Ganji, M.; Farahani, M.S.; Afsharnouri, F.; Pourzardosht, N.; Khalesi, B.; Jahangiri, A.; Rahbar, M.R.; et al. In silico Approaches for the Design and Optimization of Interfering Peptides Against Protein-Protein Interactions. *Front. Mol. Biosci.* **2021**, *8*, 669431. [[CrossRef](#)] [[PubMed](#)]
4. Anand, U.; Bandyopadhyay, A.; Jha, N.K.; Perez de la Lastra, J.M.; Dey, A. Translational aspect in peptide drug discovery and development: An emerging therapeutic candidate. *Biofactors* **2023**, *49*, 251–269. [[CrossRef](#)] [[PubMed](#)]
5. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)] [[PubMed](#)]
6. Skolnick, J.; Gao, M.; Zhou, H.; Singh, S. AlphaFold 2: Why It Works and Its Implications for Understanding the Relationships of Protein Sequence, Structure, and Function. *J. Chem. Inf. Model.* **2021**, *61*, 4827–4831. [[CrossRef](#)] [[PubMed](#)]
7. Bertoline, L.M.F.; Lima, A.N.; Krieger, J.E.; Teixeira, S.K. Before and after AlphaFold2: An overview of protein structure prediction. *Front. Bioinform.* **2023**, *3*, 1120370. [[CrossRef](#)] [[PubMed](#)]
8. Watson, J.L.; Juergens, D.; Bennett, N.R.; Trippe, B.L.; Yim, J.; Eisenach, H.E.; Ahern, W.; Borst, A.J.; Ragotte, R.J.; Milles, L.F.; et al. De novo design of protein structure and function with RFdiffusion. *Nature* **2023**, *620*, 1089–1100. [[CrossRef](#)]
9. Swinney, D.C. Phenotypic vs. target-based drug discovery for first-in-class medicines. *Clin. Pharmacol. Ther.* **2013**, *93*, 299–301. [[CrossRef](#)]
10. Diller, D.J.; Jarosinski, M.; Sawyer, T.K.; Audie, J. Peptide drug discovery: Innovative technologies and transformational medicines. In *Advances in the Discovery and Development of Peptide Therapeutics*; Kruger, G., Albericio, F., Eds.; Future Medicine: London, UK, 2015; pp. 8–27. [[CrossRef](#)]
11. Winkler, D.F.; Campbell, W.D. The spot technique: Synthesis and screening of peptide macroarrays on cellulose membranes. *Methods Mol. Biol.* **2008**, *494*, 47–70. [[CrossRef](#)]
12. De Vries, S.J.; Rey, J.; Schindler, C.E.M.; Zacharias, M.; Tuffery, P. The pepATTRACT web server for blind, large-scale peptide-protein docking. *Nucleic Acids Res.* **2017**, *45*, W361–W364. [[CrossRef](#)] [[PubMed](#)]
13. Ramirez, J.; Nomine, Y. High-Quality Data of Protein/Peptide Interaction by Isothermal Titration Calorimetry. *Methods Mol. Biol.* **2019**, *1964*, 99–117. [[CrossRef](#)] [[PubMed](#)]
14. Drescher, D.G.; Selvakumar, D.; Drescher, M.J. Analysis of Protein Interactions by Surface Plasmon Resonance. *Adv. Protein Chem. Struct. Biol.* **2018**, *110*, 1–30. [[CrossRef](#)]
15. Ren, B.; Sayed, A.M.M.; Tan, H.L.; Mok, Y.K.; Chen, E.S. Identifying Protein Interactions with Histone Peptides Using Bio-layer Interferometry. *Bio Protoc.* **2018**, *8*, e3012. [[CrossRef](#)]

16. Plach, M.G.; Grasser, K.; Schubert, T. MicroScale Thermophoresis as a Tool to Study Protein-peptide Interactions in the Context of Large Eukaryotic Protein Complexes. *Bio Protoc.* **2017**, *7*, e2632. [[CrossRef](#)] [[PubMed](#)]
17. Pellecchia, M.; Bertini, I.; Cowburn, D.; Dalvit, C.; Giralt, E.; Jahnke, W.; James, T.L.; Homans, S.W.; Kessler, H.; Luchinat, C.; et al. Perspectives on NMR in drug discovery: A technique comes of age. *Nat. Rev. Drug Discov.* **2008**, *7*, 738–745. [[CrossRef](#)]
18. Pellecchia, M.; Sem, D.S.; Wuthrich, K. NMR in drug discovery. *Nat. Rev. Drug Discov.* **2002**, *1*, 211–219. [[CrossRef](#)]
19. Leone, M.; Barile, E.; Dahl, R.; Pellecchia, M. Design and NMR studies of cyclic peptides targeting the N-terminal domain of the protein tyrosine phosphatase YopH. *Chem. Biol. Drug Des.* **2011**, *77*, 12–19. [[CrossRef](#)]
20. Furukawa, A.; Konuma, T.; Yanaka, S.; Sugase, K. Quantitative analysis of protein-ligand interactions by NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *96*, 47–57. [[CrossRef](#)]
21. Hobbs, B.; Drant, J.; Williamson, M.P. The measurement of binding affinities by NMR chemical shift perturbation. *J. Biomol. NMR* **2022**, *76*, 153–163. [[CrossRef](#)]
22. Vincenzi, M.; Mercurio, F.A.; Leone, M. About TFE: Old and New Findings. *Curr. Protein Pept. Sci.* **2019**, *20*, 425–451. [[CrossRef](#)] [[PubMed](#)]
23. Mercurio, F.A.; Marasco, D.; Di Natale, C.; Pirone, L.; Costantini, S.; Pedone, E.M.; Leone, M. Targeting EphA2-Sam and Its Interactome: Design and Evaluation of Helical Peptides Enriched in Charged Residues. *Chembiochem* **2016**, *17*, 2179–2188. [[CrossRef](#)] [[PubMed](#)]
24. Kruger, D.M.; Glas, A.; Bier, D.; Pospiech, N.; Wallraven, K.; Dietrich, L.; Ottmann, C.; Koch, O.; Hennig, S.; Grossmann, T.N. Structure-Based Design of Non-natural Macrocyclic Peptides That Inhibit Protein-Protein Interactions. *J. Med. Chem.* **2017**, *60*, 8982–8988. [[CrossRef](#)] [[PubMed](#)]
25. Mercurio, F.A.; Pirone, L.; Di Natale, C.; Marasco, D.; Pedone, E.M.; Leone, M. Sam domain-based stapled peptides: Structural analysis and interaction studies with the Sam domains from the EphA2 receptor and the lipid phosphatase Ship2. *Bioorg. Chem.* **2018**, *80*, 602–610. [[CrossRef](#)] [[PubMed](#)]
26. Masuya, K. New trends in drug discovery and development by constrained peptides. *Nihon Yakurigaku Zasshi* **2016**, *148*, 322–328. [[CrossRef](#)] [[PubMed](#)]
27. Szabo, I.; Yousef, M.; Soltesz, D.; Bato, C.; Mezo, G.; Banoczi, Z. Redesigning of Cell-Penetrating Peptides to Improve Their Efficacy as a Drug Delivery System. *Pharmaceutics* **2022**, *14*, 907. [[CrossRef](#)]
28. Komin, A.; Russell, L.M.; Hristova, K.A.; Searson, P.C. Peptide-based strategies for enhanced cell uptake, transcellular transport, and circulation: Mechanisms and challenges. *Adv. Drug Deliv. Rev.* **2017**, *110–111*, 52–64. [[CrossRef](#)]
29. Al Musaimi, O.; Lombardi, L.; Williams, D.R.; Albericio, F. Strategies for Improving Peptide Stability and Delivery. *Pharmaceutics* **2022**, *15*, 1283. [[CrossRef](#)]
30. Fominaya, J.; Bravo, J.; Rebollo, A. Strategies to stabilize cell penetrating peptides for in vivo applications. *Ther. Deliv.* **2015**, *6*, 1171–1194. [[CrossRef](#)]
31. Kalafatovic, D.; Giralt, E. Cell-Penetrating Peptides: Design Strategies beyond Primary Structure and Amphipathicity. *Molecules* **2017**, *22*, 1929. [[CrossRef](#)] [[PubMed](#)]
32. Chen, S.; Gfeller, D.; Buth, S.A.; Michielin, O.; Leiman, P.G.; Heinis, C. Improving binding affinity and stability of peptide ligands by substituting glycines with D-amino acids. *Chembiochem* **2013**, *14*, 1316–1322. [[CrossRef](#)] [[PubMed](#)]
33. Wang, L.; Wang, N.; Zhang, W.; Cheng, X.; Yan, Z.; Shao, G.; Wang, X.; Wang, R.; Fu, C. Therapeutic peptides: Current applications and future directions. *Signal Transduct. Target. Ther.* **2022**, *7*, 48. [[CrossRef](#)] [[PubMed](#)]
34. Lau, J.L.; Dunn, M.K. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorg. Med. Chem.* **2018**, *26*, 2700–2707. [[CrossRef](#)]
35. Lamers, C. Overcoming the shortcomings of peptide-based therapeutics. *Future Drug Discov.* **2022**, *4*, FDD75. [[CrossRef](#)]
36. Muttenthaler, M.; King, G.F.; Adams, D.J.; Alewood, P.F. Trends in peptide drug discovery. *Nat. Rev. Drug Discov.* **2021**, *20*, 309–325. [[CrossRef](#)] [[PubMed](#)]
37. Hummel, G.; Reineke, U.; Reimer, U. Translating peptides into small molecules. *Mol. Biosyst.* **2006**, *2*, 499–508. [[CrossRef](#)]
38. Zane, D.; Feldman, P.L.; Sawyer, T.; Sobol, Z.; Hawes, J. Development and Regulatory Challenges for Peptide Therapeutics. *Int. J. Toxicol.* **2021**, *40*, 108–124. [[CrossRef](#)] [[PubMed](#)]
39. Otvos, L., Jr.; Wade, J.D. Current challenges in peptide-based drug discovery. *Front. Chem.* **2014**, *2*, 62. [[CrossRef](#)]
40. Cary, D.R.; Ohuchi, M.; Reid, P.C.; Masuya, K. Constrained Peptides in Drug Discovery and Development. *J. Synth. Org. Chem. Jpn.* **2017**, *75*, 1171–1178. [[CrossRef](#)]
41. Walensky, L.D.; Bird, G.H. Hydrocarbon-stapled peptides: Principles, practice, and progress. *J. Med. Chem.* **2014**, *57*, 6275–6288. [[CrossRef](#)]
42. Sun, Z.; Huang, J.; Fishelson, Z.; Wang, C.; Zhang, S. Cell-Penetrating Peptide-Based Delivery of Macromolecular Drugs: Development, Strategies, and Progress. *Biomedicines* **2023**, *11*, 1971. [[CrossRef](#)]
43. Torchilin, V. Intracellular delivery of protein and peptide therapeutics. *Drug Discov. Today Technol.* **2008**, *5*, e95–e103. [[CrossRef](#)] [[PubMed](#)]
44. Yoshikawa, T.; Sugita, T.; Mukai, Y.; Abe, Y.; Nakagawa, S.; Kamada, H.; Tsunoda, S.; Tsutsumi, Y. The augmentation of intracellular delivery of peptide therapeutics by artificial protein transduction domains. *Biomaterials* **2009**, *30*, 3318–3323. [[CrossRef](#)] [[PubMed](#)]
45. Lv, S.; Sylvestre, M.; Prossnitz, A.N.; Yang, L.F.; Pun, S.H. Design of Polymeric Carriers for Intracellular Peptide Delivery in Oncology Applications. *Chem. Rev.* **2021**, *121*, 11653–11698. [[CrossRef](#)] [[PubMed](#)]

46. Matsson, P.; Doak, B.C.; Over, B.; Kihlberg, J. Cell permeability beyond the rule of 5. *Adv. Drug Deliv. Rev.* **2016**, *101*, 42–61. [[CrossRef](#)]
47. Howell, L.A.; Beekman, A.M. In silico peptide-directed ligand design complements experimental peptide-directed binding for protein-protein interaction modulator discovery. *RSC Chem. Biol.* **2021**, *2*, 215–219. [[CrossRef](#)] [[PubMed](#)]
48. Beekman, A.M.; Cominetti, M.M.D.; Walpole, S.J.; Prabhu, S.; O'Connell, M.A.; Angulo, J.; Searcey, M. Identification of selective protein-protein interaction inhibitors using efficient in silico peptide-directed ligand design. *Chem. Sci.* **2019**, *10*, 4502–4508. [[CrossRef](#)]
49. Coscueta, E.R.; Batista, P.; Gomes, J.E.G.; da Silva, R.; Pintado, M.M. Screening of Novel Bioactive Peptides from Goat Casein: In Silico to In Vitro Validation. *Int. J. Mol. Sci.* **2022**, *23*, 2439. [[CrossRef](#)]
50. Mahmoodi-Reihani, M.; Abbasitabar, F.; Zare-Shahabadi, V. In Silico Rational Design and Virtual Screening of Bioactive Peptides Based on QSAR Modeling. *ACS Omega* **2020**, *5*, 5951–5958. [[CrossRef](#)]
51. Prasasty, V.D.; Istyastono, E.P. Data of small peptides in SMILES and three-dimensional formats for virtual screening campaigns. *Data Brief* **2019**, *27*, 104607. [[CrossRef](#)]
52. Joshi, J.; Blankenberg, D. PDAUG: A Galaxy based toolset for peptide library analysis, visualization, and machine learning modeling. *BMC Bioinform.* **2022**, *23*, 197. [[CrossRef](#)]
53. Crooks, R.O.; Baxter, D.; Panek, A.S.; Lubben, A.T.; Mason, J.M. Deriving Heterospecific Self-Assembling Protein-Protein Interactions Using a Computational Interactome Screen. *J. Mol. Biol.* **2016**, *428*, 385–398. [[CrossRef](#)] [[PubMed](#)]
54. Engel, H.; Guischart, F.; Krause, F.; Nandy, J.; Kaas, P.; Hofflin, N.; Kohn, M.; Kilb, N.; Voigt, K.; Wolf, S.; et al. finDr: A web server for in silico D-peptide ligand identification. *Synth. Syst. Biotechnol.* **2021**, *6*, 402–413. [[CrossRef](#)]
55. Frederix, P.W.; Ulijn, R.V.; Hunt, N.T.; Tuttle, T. Virtual Screening for Dipeptide Aggregation: Toward Predictive Tools for Peptide Self-Assembly. *J. Phys. Chem. Lett.* **2011**, *2*, 2380–2384. [[CrossRef](#)] [[PubMed](#)]
56. Amarasinghe, K.N.; De Maria, L.; Tyrchan, C.; Eriksson, L.A.; Sadowski, J.; Petrovic, D. Virtual Screening Expands the Non-Natural Amino Acid Palette for Peptide Optimization. *J. Chem. Inf. Model.* **2022**, *62*, 2999–3007. [[CrossRef](#)] [[PubMed](#)]
57. Foight, G.W.; Chen, T.S.; Richman, D.; Keating, A.E. Enriching Peptide Libraries for Binding Affinity and Specificity Through Computationally Directed Library Design. *Methods Mol. Biol.* **2017**, *1561*, 213–232. [[CrossRef](#)] [[PubMed](#)]
58. Golosov, A.A.; Flyer, A.N.; Amin, J.; Babu, C.; Gampe, C.; Li, J.; Liu, E.; Nakajima, K.; Nettleton, D.; Patel, T.J.; et al. Design of Thioether Cyclic Peptide Scaffolds with Passive Permeability and Oral Exposure. *J. Med. Chem.* **2021**, *64*, 2622–2633. [[CrossRef](#)] [[PubMed](#)]
59. Saha, I.; Dang, E.K.; Svatunek, D.; Houk, K.N.; Harran, P.G. Computational generation of an annotated gigalibrary of synthesizable, composite peptidic macrocycles. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 24679–24690. [[CrossRef](#)] [[PubMed](#)]
60. Gautam, A.; Chaudhary, K.; Kumar, R.; Sharma, A.; Kapoor, P.; Tyagi, A.; Open source drug discovery, c.; Raghava, G.P. In silico approaches for designing highly effective cell penetrating peptides. *J. Transl. Med.* **2013**, *11*, 74. [[CrossRef](#)]
61. Park, H.; Park, J.H.; Kim, M.S.; Cho, K.; Shin, J.M. In Silico Screening and Optimization of Cell-Penetrating Peptides Using Deep Learning Methods. *Biomolecules* **2023**, *13*, 522. [[CrossRef](#)]
62. Viart, B.; Dias-Lopes, C.; Kozlova, E.; Oliveira, C.F.; Nguyen, C.; Neshich, G.; Chavez-Olortegui, C.; Molina, F.; Felicori, L.F. EPI-peptide designer: A tool for designing peptide ligand libraries based on epitope-paratope interactions. *Bioinformatics* **2016**, *32*, 1462–1470. [[CrossRef](#)]
63. Jukic, M.; Kralj, S.; Kolaric, A.; Bren, U. Design of Tetra-Peptide Ligands of Antibody Fc Regions Using In Silico Combinatorial Library Screening. *Pharmaceuticals* **2023**, *16*, 1170. [[CrossRef](#)] [[PubMed](#)]
64. Yagi, Y.; Terada, K.; Noma, T.; Ikebukuro, K.; Sode, K. In silico panning for a non-competitive peptide inhibitor. *BMC Bioinform.* **2007**, *8*, 11. [[CrossRef](#)] [[PubMed](#)]
65. Kalafatovic, D.; Maus, G.; Todorovski, T.; Giral, E. Algorithm-supported, mass and sequence diversity-oriented random peptide library design. *J. Cheminform.* **2019**, *11*, 25. [[CrossRef](#)] [[PubMed](#)]
66. Chang, L.; Mondal, A.; Perez, A. Towards rational computational peptide design. *Front. Bioinform.* **2022**, *2*, 1046493. [[CrossRef](#)] [[PubMed](#)]
67. Tripathi, N.M.; Bandyopadhyay, A. High throughput virtual screening (HTVS) of peptide library: Technological advancement in ligand discovery. *Eur. J. Med. Chem.* **2022**, *243*, 114766. [[CrossRef](#)] [[PubMed](#)]
68. Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* **2018**, *9*, 923. [[CrossRef](#)] [[PubMed](#)]
69. Weng, G.; Gao, J.; Wang, Z.; Wang, E.; Hu, X.; Yao, X.; Cao, D.; Hou, T. Comprehensive Evaluation of Fourteen Docking Programs on Protein-Peptide Complexes. *J. Chem. Theory Comput.* **2020**, *16*, 3959–3969. [[CrossRef](#)] [[PubMed](#)]
70. Lee, H.; Heo, L.; Lee, M.S.; Seok, C. GalaxyPepDock: A protein-peptide docking tool based on interaction similarity and energy optimization. *Nucleic Acids Res.* **2015**, *43*, W431–W435. [[CrossRef](#)]
71. Zhou, P.; Jin, B.; Li, H.; Huang, S.Y. HPEPDOCK: A web server for blind peptide-protein docking based on a hierarchical algorithm. *Nucleic Acids Res.* **2018**, *46*, W443–W450. [[CrossRef](#)]
72. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [[CrossRef](#)]
73. Rentzsch, R.; Renard, B.Y. Docking small peptides remains a great challenge: An assessment using AutoDock Vina. *Brief Bioinform.* **2015**, *16*, 1045–1056. [[CrossRef](#)] [[PubMed](#)]

74. Verdonk, M.L.; Cole, J.C.; Hartshorn, M.J.; Murray, C.W.; Taylor, R.D. Improved protein-ligand docking using GOLD. *Proteins* **2003**, *52*, 609–623. [[CrossRef](#)] [[PubMed](#)]
75. Jones, G.; Willett, P.; Glen, R.C.; Leach, A.R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748. [[CrossRef](#)]
76. Geng, C.; Narasimhan, S.; Rodrigues, J.P.; Bonvin, A.M. Information-Driven, Ensemble Flexible Peptide Docking Using HADDOCK. *Methods Mol. Biol.* **2017**, *1561*, 109–138. [[CrossRef](#)] [[PubMed](#)]
77. Van Zundert, G.C.; Bonvin, A.M. Modeling protein-protein complexes using the HADDOCK webserver “modeling protein complexes with HADDOCK”. *Methods Mol. Biol.* **2014**, *1137*, 163–179. [[CrossRef](#)] [[PubMed](#)]
78. Antunes, D.A.; Moll, M.; Devaurs, D.; Jackson, K.R.; Lizee, G.; Kavraki, L.E. DINC 2.0: A New Protein-Peptide Docking Webserver Using an Incremental Approach. *Cancer Res.* **2017**, *77*, e55–e57. [[CrossRef](#)] [[PubMed](#)]
79. Paul, D.S.; Karthe, P. Improved docking of peptides and small molecules in iMOLSDOCK. *J. Mol. Model.* **2022**, *29*, 12. [[CrossRef](#)] [[PubMed](#)]
80. Kurcinski, M.; Badaczewska-Dawid, A.; Kolinski, M.; Kolinski, A.; Kmiecik, S. Flexible docking of peptides to proteins using CABS-dock. *Protein Sci.* **2020**, *29*, 211–222. [[CrossRef](#)]
81. Almeida Filho, J.L.d.; Fernandez, J.H. HTP SurfexDock: A web tool for Structure-Based Virtual Screening analysis based on the Ensemble Docking protocol. In Proceedings of the Brazilian e-Science Workshop (BRESKI), Porto Alegre, Brazil, 31 July 2022; Sociedade Brasileira de Computação: Porto Alegre, Brazil, 2022; Volume 16, pp. 81–88.
82. Ansar, S.; Vetrivel, U. PepVis: An integrated peptide virtual screening pipeline for ensemble and flexible docking protocols. *Chem. Biol. Drug Des.* **2019**, *94*, 2041–2050. [[CrossRef](#)]
83. Pierce, B.G.; Wiehe, K.; Hwang, H.; Kim, B.H.; Vreven, T.; Weng, Z. ZDOCK server: Interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* **2014**, *30*, 1771–1773. [[CrossRef](#)] [[PubMed](#)]
84. Zhang, Y.; Sanner, M.F. AutoDock CrankPep: Combining folding and docking to predict protein-peptide complexes. *Bioinformatics* **2019**, *35*, 5121–5127. [[CrossRef](#)] [[PubMed](#)]
85. Yan, Y.; Zhang, D.; Huang, S.Y. Efficient conformational ensemble generation of protein-bound peptides. *J. Cheminform.* **2017**, *9*, 59. [[CrossRef](#)] [[PubMed](#)]
86. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718. [[CrossRef](#)] [[PubMed](#)]
87. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [[CrossRef](#)]
88. Pierce, B.; Weng, Z. A combination of rescoring and refinement significantly improves protein docking performance. *Proteins* **2008**, *72*, 270–279. [[CrossRef](#)]
89. Alam, N.; Goldstein, O.; Xia, B.; Porter, K.A.; Kozakov, D.; Schueler-Furman, O. High-resolution global peptide-protein docking using fragments-based PIPER-FlexPepDock. *PLoS Comput. Biol.* **2017**, *13*, e1005905. [[CrossRef](#)]
90. Perez, G.M.; Salomon, L.A.; Montero-Cabrera, L.A.; de la Vega, J.M.; Mascini, M. Integrating sampling techniques and inverse virtual screening: Toward the discovery of artificial peptide-based receptors for ligands. *Mol. Divers.* **2016**, *20*, 421–438. [[CrossRef](#)]
91. Duffy, F.J.; O'Donovan, D.; Devocelle, M.; Moran, N.; O'Connell, D.J.; Shields, D.C. Virtual screening using combinatorial cyclic peptide libraries reveals protein interfaces readily targetable by cyclic peptides. *J. Chem. Inf. Model.* **2015**, *55*, 600–613. [[CrossRef](#)]
92. Santini, B.L.; Zacharias, M. Rapid in silico Design of Potential Cyclic Peptide Binders Targeting Protein-Protein Interfaces. *Front. Chem.* **2020**, *8*, 573259. [[CrossRef](#)]
93. Bender, B.J.; Gahbauer, S.; Lutten, A.; Lyu, J.; Webb, C.M.; Stein, R.M.; Fink, E.A.; Balius, T.E.; Carlsson, J.; Irwin, J.J.; et al. A practical guide to large-scale docking. *Nat. Protoc.* **2021**, *16*, 4799–4832. [[CrossRef](#)]
94. Paul Daniel Phillips, P.D.; Andersen, T.; McDougal, O.M. Assessing the utility and limitations of high throughput virtual screening. *AIMS Mol. Sci.* **2016**, *3*, 238–245. [[CrossRef](#)]
95. Waszkowycz, B.; Clark, D.E.; Gancia, E. Outstanding challenges in protein-ligand docking and structure-based virtual screening. *WIREs Comput. Mol. Sci.* **2011**, *1*, 229–259. [[CrossRef](#)]
96. Berry, M.; Fielding, B.; Gamielien, J. Practical Considerations in Virtual Screening and Molecular Docking. In *Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology*, 1st ed.; Tran, Q.N., Arabnia, H.R., Eds.; Wiley Inc.: Hoboken, NJ, USA, 2015; pp. 487–502. [[CrossRef](#)]
97. Budipramana, K.; Sangande, F. Molecular docking-based virtual screening: Challenges in hits identification for Anti-SARS-CoV-2 activity. *Pharmacia* **2022**, *69*, 1047–1056. [[CrossRef](#)]
98. Gimeno, A.; Ojeda-Montes, M.J.; Tomas-Hernandez, S.; Cereto-Massague, A.; Beltran-Debon, R.; Mulero, M.; Pujadas, G.; Garcia-Vallve, S. The Light and Dark Sides of Virtual Screening: What Is There to Know? *Int. J. Mol. Sci.* **2019**, *20*, 1375. [[CrossRef](#)]
99. Macip, G.; Garcia-Segura, P.; Mestres-Truyol, J.; Saldivar-Espinoza, B.; Ojeda-Montes, M.J.; Gimeno, A.; Cereto-Massague, A.; Garcia-Vallve, S.; Pujadas, G. Haste makes waste: A critical review of docking-based virtual screening in drug repurposing for SARS-CoV-2 main protease (M-pro) inhibition. *Med. Res. Rev.* **2022**, *42*, 744–769. [[CrossRef](#)] [[PubMed](#)]
100. Agu, P.C.; Afiukwa, C.A.; Orji, O.U.; Ezech, E.M.; Ofoke, I.H.; Ogbu, C.O.; Ugwuja, E.I.; Aja, P.M. Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. *Sci. Rep.* **2023**, *13*, 13398. [[CrossRef](#)]
101. Xie, D.; Du, L.; Lin, H.; Su, E.; Shen, Y.; Xie, J.; Wei, D. In vitro-in silico screening strategy and mechanism of angiotensin I-converting enzyme inhibitory peptides from α -lactalbumin. *Food Sci. Technol.* **2022**, *156*, 112984. [[CrossRef](#)]

102. Yu, Z.; Wu, S.; Zhao, W.; Ding, L.; Li, J.; Liu, J. Virtual screening and molecular docking for exploring ACE inhibitory peptides in *Larimichthys crocea* nebulin protein. *Int. Food Res. J.* **2019**, *26*, 1417–1426.
103. Panyayai, T.; Sangsawad, P.; Pacharawongsakda, E.; Sawatdichaikul, O.; Tongsimma, S.; Choowongkamon, K. The potential peptides against angiotensin-I converting enzyme through a virtual tripeptide-constructing library. *Comput. Biol. Chem.* **2018**, *77*, 207–213. [[CrossRef](#)] [[PubMed](#)]
104. Kumar, R.; Chaudhary, K.; Singh Chauhan, J.; Nagpal, G.; Kumar, R.; Sharma, M.; Raghava, G.P. An in silico platform for predicting, screening and designing of antihypertensive peptides. *Sci. Rep.* **2015**, *5*, 12512. [[CrossRef](#)]
105. Stefanucci, A.; Iobbi, V.; Della Valle, A.; Scioli, G.; Pieretti, S.; Minosi, P.; Mirzaie, S.; Novellino, E.; Mollica, A. In Silico Identification of Tripeptides as Lead Compounds for the Design of KOR Ligands. *Molecules* **2021**, *26*, 4767. [[CrossRef](#)] [[PubMed](#)]
106. Malyshev, A.V.; Sukhanova, I.A.; Zlobin, A.S.; Gedzun, V.R.; Pavshintsev, V.V.; Vasileva, E.V.; Zalevsky, A.O.; Doronin, I.I.; Mitkin, N.A.; Golovin, A.V.; et al. In silico Screening and Behavioral Validation of a Novel Peptide, LCGA-17, with Anxiolytic-Like Properties. *Front. Neurosci.* **2021**, *15*, 705590. [[CrossRef](#)] [[PubMed](#)]
107. Tan, J.; Grouleff, J.J.; Jitkova, Y.; Diaz, D.B.; Griffith, E.C.; Shao, W.; Bogdanchikova, A.F.; Poda, G.; Schimmer, A.D.; Lee, R.E.; et al. De Novo Design of Boron-Based Peptidomimetics as Potent Inhibitors of Human ClpP in the Presence of Human ClpX. *J. Med. Chem.* **2019**, *62*, 6377–6390. [[CrossRef](#)] [[PubMed](#)]
108. Hussain, A.; Shaw, P.E.; Hirst, J.D. Molecular dynamics simulations and in silico peptide ligand screening of the Elk-1 ETS domain. *J. Cheminform.* **2011**, *3*, 49. [[CrossRef](#)] [[PubMed](#)]
109. Wang, Y.; Guo, H.; Feng, Z.; Wang, S.; Wang, Y.; He, Q.; Li, G.; Lin, W.; Xie, X.Q.; Lin, Z. PD-1-Targeted Discovery of Peptide Inhibitors by Virtual Screening, Molecular Dynamics Simulation, and Surface Plasmon Resonance. *Molecules* **2019**, *24*, 3784. [[CrossRef](#)] [[PubMed](#)]
110. Yan, F.; Liu, G.; Chen, T.; Fu, X.; Niu, M.M. Structure-Based Virtual Screening and Biological Evaluation of Peptide Inhibitors for Polo-Box Domain. *Molecules* **2019**, *25*, 107. [[CrossRef](#)] [[PubMed](#)]
111. Mastouri, M.; Baachaoui, S.; Mosbah, A.; Raouafi, N. In silico screening for oligopeptides useful as capture and reporting probes for interleukin-6 biosensing. *RSC Adv.* **2022**, *12*, 13003–13013. [[CrossRef](#)] [[PubMed](#)]
112. Zhang, L.; Wei, X.; Zhang, R.; Mozdziak, P.E.; Si, D.; Ahmad, B.; Cheng, Q.; Tong, Y. Design and Immunological Evaluation of a Hybrid Peptide as a Potent TLR2 Agonist by Structure-Based Virtual Screening. *Front. Cell Dev. Biol.* **2021**, *9*, 620370. [[CrossRef](#)]
113. Dama, E.; Cornelie, S.; Camara, M.; Somda, M.B.; Poinsignon, A.; Ilboudo, H.; Elanga Ndille, E.; Jamonneau, V.; Solano, P.; Remoue, F.; et al. In silico identification of a candidate synthetic peptide (Tsgf118-43) to monitor human exposure to tsetse flies in West Africa. *PLoS Negl. Trop. Dis.* **2013**, *7*, e2455. [[CrossRef](#)]
114. Aksoydan, B.; Durdagi, S. Virtual drug repurposing study for the CGRPR identifies pentagastrin and leuprolerin as putative candidates. *J. Mol. Graph. Model.* **2022**, *116*, 108254. [[CrossRef](#)] [[PubMed](#)]
115. Mollica, A.; Zengin, G.; Durdagi, S.; Ekhteiari Salmas, R.; Macedonio, G.; Stefanucci, A.; Dimmito, M.P.; Novellino, E. Combinatorial peptide library screening for discovery of diverse alpha-glucosidase inhibitors using molecular dynamics simulations and binary QSAR models. *J. Biomol. Struct. Dyn.* **2019**, *37*, 726–740. [[CrossRef](#)] [[PubMed](#)]
116. Doytchinova, I.A.; Blythe, M.J.; Flower, D.R. Additive method for the prediction of protein-peptide binding affinity. Application to the MHC class I molecule HLA-A*0201. *J. Proteome Res.* **2002**, *1*, 263–272. [[CrossRef](#)] [[PubMed](#)]
117. Doytchinova, I.A.; Walshe, V.A.; Jones, N.A.; Gloster, S.E.; Borrow, P.; Flower, D.R. Coupling in silico and in vitro analysis of peptide-MHC binding: A bioinformatic approach enabling prediction of superbinding peptides and anchorless epitopes. *J. Immunol.* **2004**, *172*, 7495–7502. [[CrossRef](#)] [[PubMed](#)]
118. Imai, K.; Takeuchi, Y.; Shimizu, K.; Honda, H. In Silico Screening of a Bile Acid Micelle Disruption Peptide for Oral Consumptions from Edible Peptide Database. *Foods* **2021**, *10*, 2496. [[CrossRef](#)] [[PubMed](#)]
119. Deryusheva, E.; Machulin, A.; Litus, E. Virtual Screening of Human Serum Albumin Mutants to Optimize the Search for its Forms that Increase Affinity to Amyloid-B Peptide. *BIO Web Conf.* **2023**, *57*. [[CrossRef](#)]
120. Martini, S.; Cattivelli, A.; Conte, A.; Tagliazucchi, D. Application of a Combined Peptidomics and In Silico Approach for the Identification of Novel Dipeptidyl Peptidase-IV-Inhibitory Peptides in In Vitro Digested Pinto Bean Protein Extract. *Curr. Issues Mol. Biol.* **2021**, *44*, 139–151. [[CrossRef](#)] [[PubMed](#)]
121. Zhao, W.; Sub, L.; Huo, S.; Yua, Z.; Lib, J.; Liuca, J. Virtual screening, molecular docking and identification of umami peptides derived from *Oncorhynchus mykiss*. *Food Sci. Hum. Wellness.* **2023**, *12*, 89–93. [[CrossRef](#)]
122. Mercurio, F.A.; Di Natale, C.; Pirone, L.; Marasco, D.; Calce, E.; Vincenzi, M.; Pedone, E.M.; De Luca, S.; Leone, M. Design and analysis of EphA2-SAM peptide ligands: A multi-disciplinary screening approach. *Bioorg. Chem.* **2019**, *84*, 434–443. [[CrossRef](#)]
123. Mercurio, F.A.; Di Natale, C.; Pirone, L.; Vincenzi, M.; Marasco, D.; De Luca, S.; Pedone, E.M.; Leone, M. Exploring the Ability of Cyclic Peptides to Target SAM Domains: A Computational and Experimental Study. *ChemBioChem* **2020**, *21*, 702–711. [[CrossRef](#)]
124. Mercurio, F.A.; Vincenzi, M.; Leone, M. Hunting for Novel Routes in Anticancer Drug Discovery: Peptides against Sam-Sam Interactions. *Int. J. Mol. Sci.* **2022**, *23*, 10397. [[CrossRef](#)] [[PubMed](#)]
125. Leffler, A.E.; Kuryatov, A.; Zebroski, H.A.; Powell, S.R.; Filipenko, P.; Hussein, A.K.; Gorson, J.; Heizmann, A.; Lyskov, S.; Tsien, R.W.; et al. Discovery of peptide ligands through docking and virtual screening at nicotinic acetylcholine receptor homology models. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E8100–E8109. [[CrossRef](#)]
126. Chiangjong, W.; Chutipongtanate, S.; Hongeng, S. Anticancer peptide: Physicochemical property, functional aspect and trend in clinical application (Review). *Int. J. Oncol.* **2020**, *57*, 678–696. [[CrossRef](#)]

127. International Agency for Research on Cancer—World Health Organization. Available online: <https://www.iarc.who.int/> (accessed on 25 November 2023).
128. Gaspar, D.; Veiga, A.S.; Castanho, M.A. From antimicrobial to anticancer peptides. A review. *Front. Microbiol.* **2013**, *4*, 294. [[CrossRef](#)]
129. Speck-Planche, A.; Cordeiro, M.N.D.S. Speeding Up the Virtual Design and Screening of Therapeutic Peptides. In *Multi-Scale Approaches in Drug Discovery: From Empirical Knowledge to In Silico Experiments and Back*, 1st ed.; Speck-Planche, A., Ed.; Elsevier: Amsterdam, The Netherlands, 2017; pp. 127–147. [[CrossRef](#)]
130. Mercurio, F.A.; Leone, M. The Sam Domain of EphA2 Receptor and its Relevance to Cancer: A Novel Challenge for Drug Discovery? *Curr. Med. Chem.* **2016**, *23*, 4718–4734. [[CrossRef](#)] [[PubMed](#)]
131. Tyagi, A.; Tuknait, A.; Anand, P.; Gupta, S.; Sharma, M.; Mathur, D.; Joshi, A.; Singh, S.; Gautam, A.; Raghava, G.P. CancerPPD: A database of anticancer peptides and proteins. *Nucleic Acids Res.* **2015**, *43*, D837–D843. [[CrossRef](#)] [[PubMed](#)]
132. Tyagi, A.; Kapoor, P.; Kumar, R.; Chaudhary, K.; Gautam, A.; Raghava, G.P. In silico models for designing and discovering novel anticancer peptides. *Sci. Rep.* **2013**, *3*, 2984. [[CrossRef](#)] [[PubMed](#)]
133. Chai, T.T.; Koh, J.A.; Wong, C.C.; Sabri, M.Z.; Wong, F.C. Computational Screening for the Anticancer Potential of Seed-Derived Antioxidant Peptides: A Cheminformatic Approach. *Molecules* **2021**, *26*, 7396. [[CrossRef](#)]
134. Das, D.; Jaiswal, M.; Khan, F.N.; Ahamad, S.; Kumar, S. PlantPepDB: A manually curated plant peptide database. *Sci. Rep.* **2020**, *10*, 2194. [[CrossRef](#)]
135. Agrawal, P.; Bhagat, D.; Mahalwal, M.; Sharma, N.; Raghava, G.P.S. AntiCP 2.0: An updated model for predicting anticancer peptides. *Brief Bioinform.* **2021**, *22*, bbaa153. [[CrossRef](#)]
136. Shaw, S.A.; Vokits, B.P.; Dilger, A.K.; Viet, A.; Clark, C.G.; Abell, L.M.; Locke, G.A.; Duke, G.; Kopcho, L.M.; Dongre, A.; et al. Discovery and structure activity relationships of 7-benzyl triazolopyridines as stable, selective, and reversible inhibitors of myeloperoxidase. *Bioorg. Med. Chem.* **2020**, *28*, 115723. [[CrossRef](#)]
137. Cao, H.; Pauff, J.M.; Hille, R. X-ray crystal structure of a xanthine oxidase complex with the flavonoid inhibitor quercetin. *J. Nat. Prod.* **2014**, *77*, 1693–1699. [[CrossRef](#)]
138. Lo, S.C.; Li, X.; Henzl, M.T.; Beamer, L.J.; Hannink, M. Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. *EMBO J.* **2006**, *25*, 3605–3617. [[CrossRef](#)] [[PubMed](#)]
139. Ogura, K.; Nobuhisa, I.; Yuzawa, S.; Takeya, R.; Torikai, S.; Saikawa, K.; Sumimoto, H.; Inagaki, F. NMR solution structure of the tandem Src homology 3 domains of p47phox complexed with a p22phox-derived proline-rich peptide. *J. Biol. Chem.* **2006**, *281*, 3660–3668. [[CrossRef](#)]
140. Lamiable, A.; Thevenet, P.; Rey, J.; Vavrusa, M.; Derreumaux, P.; Tuffery, P. PEP-FOLD3: Faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res.* **2016**, *44*, W449–W454. [[CrossRef](#)]
141. Kochnev, Y.; Hellemann, E.; Cassidy, K.C.; Durrant, J.D. Webina: An open-source library and web app that runs AutoDock Vina entirely in the web browser. *Bioinformatics* **2020**, *36*, 4513–4515. [[CrossRef](#)]
142. Manavalan, B.; Subramaniyam, S.; Shin, T.H.; Kim, M.O.; Lee, G. Machine-learning-based prediction of cell-penetrating peptides and their uptake efficiency with improved accuracy. *J. Proteome Res.* **2018**, *17*, 2715–2726. [[CrossRef](#)] [[PubMed](#)]
143. Kumar, V.; Patiyal, S.; Dhalla, A.; Sharma, N.; Raghava, G.P.S. B3Pred: A Random-Forest-Based Method for Predicting and Designing Blood-Brain Barrier Penetrating Peptides. *Pharmaceutics* **2021**, *13*, 1237. [[CrossRef](#)] [[PubMed](#)]
144. Mathur, D.; Singh, S.; Mehta, A.; Agrawal, P.; Raghava, G.P.S. In silico approaches for predicting the half-life of natural and modified peptides in blood. *PLoS ONE* **2018**, *13*, e0196829. [[CrossRef](#)]
145. Minkiewicz, P.; Iwaniak, A.; Darewicz, M. BIOPEP-UWM Database of Bioactive Peptides: Current Opportunities. *Int. J. Mol. Sci.* **2019**, *20*, 5978. [[CrossRef](#)]
146. Wood, C.W.; Ibarra, A.A.; Bartlett, G.J.; Wilson, A.J.; Woolfson, D.N.; Sessions, R.B. BAlaS: Fast, interactive and accessible computational alanine-scanning using BudeAlaScan. *Bioinformatics* **2020**, *36*, 2917–2919. [[CrossRef](#)]
147. Ferguson, K.M.; Berger, M.B.; Mendrola, J.M.; Cho, H.S.; Leahy, D.J.; Lemmon, M.A. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol. Cell* **2003**, *11*, 507–517. [[CrossRef](#)]
148. Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639–1662. [[CrossRef](#)]
149. Song, S.; Liu, D.; Peng, J.; Deng, H.; Guo, Y.; Xu, L.X.; Miller, A.D.; Xu, Y. Novel peptide ligand directs liposomes toward EGF-R high-expressing cancer cells in vitro and in vivo. *FASEB J.* **2009**, *23*, 1396–1404. [[CrossRef](#)]
150. Sogabe, S.; Kawakita, Y.; Igaki, S.; Iwata, H.; Miki, H.; Cary, D.R.; Takagi, T.; Takagi, S.; Ohta, Y.; Ishikawa, T. Structure-Based Approach for the Discovery of Pyrrolo[3,2-d]pyrimidine-Based EGFR T790M/L858R Mutant Inhibitors. *ACS Med. Chem. Lett.* **2013**, *4*, 201–205. [[CrossRef](#)] [[PubMed](#)]
151. Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: Computed atlas of surface topography of proteins. *Nucleic Acids Res.* **2018**, *46*, W363–W367. [[CrossRef](#)]
152. Li, Z.; Wan, H.; Shi, Y.; Ouyang, P. Personal experience with four kinds of chemical structure drawing software: Review on ChemDraw, ChemWindow, ISIS/Draw, and ChemSketch. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1886–1890. [[CrossRef](#)] [[PubMed](#)]
153. Bhat, S.S.; Jadhav, G.; Pradeep, S.; Dharmashekar, C.; Jain, A.S.; Devegowda, D.; Shivamallu, C.; Kollur, S.P.; Prasad, S.K. In Silico Examination of Peptides Containing Selenium and Ebselen Backbone to Assess Their Tumoricidal Potential. *Int. J. Health Allied Sci.* **2022**, *11*, 9. [[CrossRef](#)]

154. Schwartz, J.C.; Zhang, X.; Fedorov, A.A.; Nathenson, S.G.; Almo, S.C. Structural basis for co-stimulation by the human CTLA-4/B7-2 complex. *Nature* **2001**, *410*, 604–608. [CrossRef]
155. Darnell, S.J.; Page, D.; Mitchell, J.C. An automated decision-tree approach to predicting protein interaction hot spots. *Proteins* **2007**, *68*, 813–823. [CrossRef]
156. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [CrossRef]
157. Raveh, B.; London, N.; Zimmerman, L.; Schueler-Furman, O. Rosetta FlexPepDock ab-initio: Simultaneous folding, docking and refinement of peptides onto their receptors. *PLoS ONE* **2011**, *6*, e18934. [CrossRef]
158. Chaudhury, S.; Lyskov, S.; Gray, J.J. PyRosetta: A script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* **2010**, *26*, 689–691. [CrossRef]
159. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [CrossRef]
160. Onufriev, A.; Bashford, D.; Case, D.A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **2004**, *55*, 383–394. [CrossRef]
161. Fu, H.; Gumbart, J.C.; Chen, H.; Shao, X.; Cai, W.; Chipot, C. BFEE: A User-Friendly Graphical Interface Facilitating Absolute Binding Free-Energy Calculations. *J. Chem. Inf. Model.* **2018**, *58*, 556–560. [CrossRef]
162. Thakkar, R.; Upreti, D.; Ishiguro, S.; Tamura, M.; Comer, J. Computational design of a cyclic peptide that inhibits the CTLA4 immune checkpoint. *RSC Med. Chem.* **2023**, *14*, 658–670. [CrossRef] [PubMed]
163. Phillips, J.C.; Hardy, D.J.; Maia, J.D.C.; Stone, J.E.; Ribeiro, J.V.; Bernardi, R.C.; Buch, R.; Fiorin, G.; Henin, J.; Jiang, W.; et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 044130. [CrossRef] [PubMed]
164. Kunczewicz, K.; Battin, C.; Sieradzan, A.; Karczynska, A.; Orlikowska, M.; Wardowska, A.; Pikula, M.; Steinberger, P.; Rodziewicz-Motowidlo, S.; Spodzieja, M. Fragments of gD Protein as Inhibitors of BTLA/HVEM Complex Formation-Design, Synthesis, and Cellular Studies. *Int. J. Mol. Sci.* **2020**, *21*, 8876. [CrossRef]
165. Lerksuthirat, T.; On-Yam, P.; Chitphuk, S.; Stitchantrakul, W.; Newburg, D.S.; Morrow, A.L.; Hongeng, S.; Chiangjong, W.; Chutipongtanate, S. ALA-A2 Is a Novel Anticancer Peptide Inspired by Alpha-Lactalbumin: A Discovery from a Computational Peptide Library, In Silico Anticancer Peptide Screening and In Vitro Experimental Validation. *Glob. Chall.* **2023**, *7*, 2200213. [CrossRef] [PubMed]
166. Gan, B.H.; Gaynord, J.; Rowe, S.M.; Deingruber, T.; Spring, D.R. The multifaceted nature of antimicrobial peptides: Current synthetic chemistry approaches and future directions. *Chem. Soc. Rev.* **2021**, *50*, 7820–7880. [CrossRef]
167. Lee, Y.J.; Shirkey, J.D.; Park, J.; Bisht, K.; Cowan, A.J. An Overview of Antiviral Peptides and Rational Biodesign Considerations. *Biodes. Res.* **2022**, *2022*, 9898241. [CrossRef] [PubMed]
168. Yang, M.; Liu, S.; Zhang, C. Antimicrobial peptides with antiviral and anticancer properties and their modification and nanodelivery systems. *Curr. Res. Biotechnol.* **2023**, *5*, 100121. [CrossRef]
169. Izadpanah, A.; Gallo, R.L. Antimicrobial peptides. *J. Am. Acad. Dermatol.* **2005**, *52*, 381–390, quiz 391–382. [CrossRef]
170. Wang, G.; Li, X.; Wang, Z. APD2: The updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Res.* **2009**, *37*, D933–D937. [CrossRef] [PubMed]
171. Fjell, C.D.; Hiss, J.A.; Hancock, R.E.; Schneider, G. Designing antimicrobial peptides: Form follows function. *Nat. Rev. Drug Discov.* **2011**, *11*, 37–51. [CrossRef]
172. Huang, J.; Xu, Y.; Xue, Y.; Huang, Y.; Li, X.; Chen, X.; Xu, Y.; Zhang, D.; Zhang, P.; Zhao, J.; et al. Identification of potent antimicrobial peptides via a machine-learning pipeline that mines the entire space of peptide sequences. *Nat. Biomed. Eng.* **2023**, *7*, 797–810. [CrossRef] [PubMed]
173. Veltri, D.; Kamath, U.; Shehu, A. Improving Recognition of Antimicrobial Peptides and Target Selectivity through Machine Learning and Genetic Programming. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2017**, *14*, 300–313. [CrossRef]
174. Kleandrova, V.V.; Ruso, J.M.; Speck-Planche, A.; Dias Soeiro Cordeiro, M.N. Enabling the Discovery and Virtual Screening of Potent and Safe Antimicrobial Peptides. Simultaneous Prediction of Antibacterial Activity and Cytotoxicity. *ACS Comb. Sci.* **2016**, *18*, 490–498. [CrossRef]
175. WHO Coronavirus (COVID-19) Dashboard—World Health Organization. Available online: <https://covid19.who.int/> (accessed on 2 December 2023).
176. Vincenzi, M.; Mercurio, F.A.; Leone, M. Looking for SARS-CoV-2 Therapeutics Through Computational Approaches. *Curr. Med. Chem.* **2023**, *30*, 3158–3214. [CrossRef]
177. Vincenzi, M.; Leone, M. The Fight against Human Viruses: How NMR Can Help? *Curr. Med. Chem.* **2021**, *28*, 4380–4453. [CrossRef] [PubMed]
178. Yang, H.; Rao, Z. Structural biology of SARS-CoV-2 and implications for therapeutic development. *Nat. Rev. Microbiol.* **2021**, *19*, 685–700. [CrossRef] [PubMed]
179. Yan, W.; Zheng, Y.; Zeng, X.; He, B.; Cheng, W. Structural biology of SARS-CoV-2: Open the door for novel therapies. *Signal Transduct. Target. Ther.* **2022**, *7*, 26. [CrossRef]
180. Jamison, D.A., Jr.; Anand Narayanan, S.; Trovao, N.S.; Guarnieri, J.W.; Topper, M.J.; Moraes-Vieira, P.M.; Zaksas, V.; Singh, K.K.; Wurtele, E.S.; Beheshti, A. A comprehensive SARS-CoV-2 and COVID-19 review, Part 1: Intracellular overdrive for SARS-CoV-2 infection. *Eur. J. Hum. Genet.* **2022**, *30*, 889–898. [CrossRef]

181. Narayanan, S.A.; Jamison, D.A., Jr.; Guarnieri, J.W.; Zaksas, V.; Topper, M.; Koutnik, A.P.; Park, J.; Clark, K.B.; Enguita, F.J.; Leitao, A.L.; et al. A comprehensive SARS-CoV-2 and COVID-19 review, Part 2: Host extracellular to systemic effects of SARS-CoV-2 infection. *Eur. J. Hum. Genet.* **2024**, *32*, 10–20. [[CrossRef](#)]
182. Malone, B.; Urakova, N.; Snijder, E.J.; Campbell, E.A. Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design. *Nat. Rev. Mol. Cell Biol.* **2022**, *23*, 21–39. [[CrossRef](#)]
183. Satpathy, R. In silico Tools and Techniques for Screening and Development of Peptide-Based Spike Protein Inhibitors against Novel Coronavirus (SARS-CoV-2). *Jordan J. Biol. Sci.* **2022**, *15*, 729–738. [[CrossRef](#)]
184. Badhe, Y.; Gupta, R.; Rai, B. In silico design of peptides with binding to the receptor binding domain (RBD) of the SARS-CoV-2 and their utility in bio-sensor development for SARS-CoV-2 detection. *RSC Adv.* **2021**, *11*, 3816–3826. [[CrossRef](#)]
185. Liu, X.; Jiang, L.; Li, L.; Lu, F.; Liu, F. Bionics design of affinity peptide inhibitors for SARS-CoV-2 RBD to block SARS-CoV-2 RBD-ACE2 interactions. *Heliyon* **2023**, *9*, e12890. [[CrossRef](#)]
186. Hu, C.; Guo, T.; Zou, Y.; Gao, J.; Gao, Y.; Niu, M.; Xia, Y.; Shen, X.; Li, J. Discovery of dual S-RBD/NRP1-targeting peptides: Structure-based virtual screening, synthesis, biological evaluation, and molecular dynamics simulation studies. *J. Enzyme Inhib. Med. Chem.* **2023**, *38*, 2212327. [[CrossRef](#)]
187. Panda, S.K.; Sen Gupta, P.S.; Biswal, S.; Ray, A.K.; Rana, M.K. ACE-2-Derived Biomimetic Peptides for the Inhibition of Spike Protein of SARS-CoV-2. *J. Proteome Res.* **2021**, *20*, 1296–1303. [[CrossRef](#)]
188. Xu, Z.; Zou, Y.; Gao, X.; Niu, M.M.; Li, J.; Xue, L.; Jiang, S. Dual-targeting cyclic peptides of receptor-binding domain (RBD) and main protease (Mpro) as potential drug leads for the treatment of SARS-CoV-2 infection. *Front. Pharmacol.* **2022**, *13*, 1041331. [[CrossRef](#)] [[PubMed](#)]
189. Hernandez Gonzalez, J.E.; Eberle, R.J.; Willbold, D.; Coronado, M.A. A Computer-Aided Approach for the Discovery of D-Peptides as Inhibitors of SARS-CoV-2 Main Protease. *Front. Mol. Biosci.* **2021**, *8*, 816166. [[CrossRef](#)] [[PubMed](#)]
190. Ramirez-Acosta, K.; Rosales-Fuerte, I.A.; Perez-Sanchez, J.E.; Nunez-Rivera, A.; Juarez, J.; Cadena-Nava, R.D. Design and selection of peptides to block the SARS-CoV-2 receptor binding domain by molecular docking. *Beilstein J. Nanotechnol.* **2022**, *13*, 699–711. [[CrossRef](#)]
191. Harnkit, N.; Khongsonthi, T.; Masuwan, N.; Prasartkul, P.; Noikaew, T.; Chumnanpuen, P. Virtual Screening for SARS-CoV-2 Main Protease Inhibitory Peptides from the Putative Hydrolyzed Peptidome of Rice Bran. *Antibiotics* **2022**, *11*, 1318. [[CrossRef](#)]
192. Porto, W.F. Virtual screening of peptides with high affinity for SARS-CoV-2 main protease. *Comput. Biol. Med.* **2021**, *133*, 104363. [[CrossRef](#)]
193. Waqas, M.; Haider, A.; Rehman, A.; Qasim, M.; Umar, A.; Sufyan, M.; Akram, H.N.; Mir, A.; Razzaq, R.; Rasool, D.; et al. Immunoinformatics and Molecular Docking Studies Predicted Potential Multiepitope-Based Peptide Vaccine and Novel Compounds against Novel SARS-CoV-2 through Virtual Screening. *Biomed Res. Int.* **2021**, *2021*, 1596834. [[CrossRef](#)]
194. Qureshi, A.; Thakur, N.; Tandon, H.; Kumar, M. AVpdb: A database of experimentally validated antiviral peptides targeting medically important viruses. *Nucleic Acids Res.* **2014**, *42*, D1147–D1153. [[CrossRef](#)] [[PubMed](#)]
195. Khater, I.; Nassar, A. Potential antiviral peptides targeting the SARS-CoV-2 spike protein. *BMC Pharmacol. Toxicol.* **2022**, *23*, 91. [[CrossRef](#)]
196. Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581*, 215–220. [[CrossRef](#)]
197. Wu, L.; Chen, Q.; Liu, K.; Wang, J.; Han, P.; Zhang, Y.; Hu, Y.; Meng, Y.; Pan, X.; Qiao, C.; et al. Broad host range of SARS-CoV-2 and the molecular basis for SARS-CoV-2 binding to cat ACE2. *Cell Discov.* **2020**, *6*, 68. [[CrossRef](#)]
198. Benton, D.J.; Wrobel, A.G.; Xu, P.; Roustan, C.; Martin, S.R.; Rosenthal, P.B.; Skehel, J.J.; Gamblin, S.J. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* **2020**, *588*, 327–330. [[CrossRef](#)] [[PubMed](#)]
199. Huang, X.; Pearce, R.; Zhang, Y. De novo design of protein peptides to block association of the SARS-CoV-2 spike protein with human ACE2. *Aging* **2020**, *12*, 11263–11276. [[CrossRef](#)]
200. Huang, X.; Pearce, R.; Zhang, Y. EvoEF2: Accurate and fast energy function for computational protein design. *Bioinformatics* **2020**, *36*, 1135–1142. [[CrossRef](#)]
201. Pearce, R.; Huang, X.; Setiawan, D.; Zhang, Y. EvoDesign: Designing Protein-Protein Binding Interactions Using Evolutionary Interface Profiles in Conjunction with an Optimized Physical Energy Function. *J. Mol. Biol.* **2019**, *431*, 2467–2476. [[CrossRef](#)]
202. Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.Y.; et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, *181*, 894–904.e9. [[CrossRef](#)]
203. Cantuti-Castelvetri, L.; Ojha, R.; Pedro, L.D.; Djannatian, M.; Franz, J.; Kuivanen, S.; van der Meer, F.; Kallio, K.; Kaya, T.; Anastasina, M.; et al. Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science* **2020**, *370*, 856–860. [[CrossRef](#)] [[PubMed](#)]
204. Daly, J.L.; Simonetti, B.; Klein, K.; Chen, K.E.; Williamson, M.K.; Anton-Plagaro, C.; Shoemark, D.K.; Simon-Gracia, L.; Bauer, M.; Hollandi, R.; et al. Neuropilin-1 is a host factor for SARS-CoV-2 infection. *Science* **2020**, *370*, 861–865. [[CrossRef](#)] [[PubMed](#)]
205. *Molecular Operating Environment (MOE)*, 2022.02; Chemical Computing Group ULC: Montreal, QC, Canada, 2023.
206. Wang, Z.; Wang, G. APD: The Antimicrobial Peptide Database. *Nucleic Acids Res.* **2004**, *32*, D590–D592. [[CrossRef](#)]
207. MacCarthy, E.A.; Zhang, C.; Zhang, Y.; Kc, D.B. GPU-I-TASSER: A GPU accelerated I-TASSER protein structure prediction tool. *Bioinformatics* **2022**, *38*, 1754–1755. [[CrossRef](#)]

208. Yang, J.; Zhang, Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* **2015**, *43*, W174–W181. [CrossRef] [PubMed]
209. Paul, S.; Sidney, J.; Sette, A.; Peters, B. TepiTool: A Pipeline for Computational Prediction of T Cell Epitope Candidates. *Curr. Protoc. Immunol.* **2016**, *114*, 18.19.11–18.19.24. [CrossRef] [PubMed]
210. Thakkar, R.; Agarwal, D.K.; Ranaweera, C.B.; Ishiguro, S.; Conda-Sheridan, M.; Gaudreault, N.N.; Richt, J.A.; Tamura, M.; Comer, J. De novo design of a stapled peptide targeting SARS-CoV-2 spike protein receptor-binding domain. *RSC Med. Chem.* **2023**, *14*, 1722–1733. [CrossRef] [PubMed]
211. Johansen-Leete, J.; Ullrich, S.; Fry, S.E.; Frkic, R.; Bedding, M.J.; Aggarwal, A.; Ashhurst, A.S.; Ekanayake, K.B.; Mahawaththa, M.C.; Sasi, V.M.; et al. Antiviral cyclic peptides targeting the main protease of SARS-CoV-2. *Chem. Sci.* **2022**, *13*, 3826–3836. [CrossRef] [PubMed]
212. Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science* **2020**, *368*, 409–412. [CrossRef] [PubMed]
213. Case, D.A.; Belfon, K.; Ben-Shalom, I.Y.; Brozell, S.R.; Cerutti, D.S.; Cheatham, T.E.; Cruzeiro, V.W.D.; Darden, T.A.; Duke, R.E.; Giambasu, G.; et al. *Amber 2020*; University of California: San Francisco, CA, USA, 2020.
214. Mahmud, S.; Paul, G.K.; Biswas, S.; Afrose, S.; Mita, M.A.; Hasan, M.R.; Shimu, M.S.S.; Hossain, A.; Promi, M.M.; Ema, F.K.; et al. Prospective Role of Peptide-Based Antiviral Therapy Against the Main Protease of SARS-CoV-2. *Front. Mol. Biosci.* **2021**, *8*, 628585. [CrossRef] [PubMed]
215. Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367*, 1444–1448. [CrossRef]
216. Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Zhao, Y.; Cao, L.; Wang, T.; Sun, Q.; Ming, Z.; Zhang, L.; et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **2020**, *368*, 779–782. [CrossRef]
217. Kim, Y.; Jedrzejczak, R.; Maltseva, N.I.; Wilamowski, M.; Endres, M.; Godzik, A.; Michalska, K.; Joachimiak, A. Crystal structure of Nsp15 endoribonuclease NendoU from SARS-CoV-2. *Protein Sci.* **2020**, *29*, 1596–1605. [CrossRef]
218. Tan, K.; Kim, Y.; Jedrzejczak, R.; Maltseva, N.; Endres, M.; Michalska, K.; Joachimiak, A. The Crystal Structure of Nsp9 Replicase Protein of COVID-19 (2020)—Center for Structural Genomics of Infectious Diseases (CSGID). Available online: <https://www.rcsb.org/structure/6w4b> (accessed on 12 December 2023).
219. Singh, S.; Chauhan, P.; Sharma, V.; Rao, A.; Kumbhar, B.V.; Prajapati, V.K. Identification of multi-targeting natural antiviral peptides to impede SARS-CoV-2 infection. *Struct. Chem.* **2022**, *34*, 1743–1758. [CrossRef]
220. Thakur, N.; Qureshi, A.; Kumar, M. AVPPred: Collection and prediction of highly effective antiviral peptides. *Nucleic Acids Res.* **2012**, *40*, W199–W204. [CrossRef] [PubMed]
221. Schaduangrat, N.; Nantasenamat, C.; Prachayasittikul, V.; Shoombuatong, W. Meta-iAVP: A Sequence-Based Meta-Predictor for Improving the Prediction of Antiviral Peptides Using Effective Feature Representation. *Int. J. Mol. Sci.* **2019**, *20*, 5743. [CrossRef] [PubMed]
222. Chung, C.R.; Kuo, T.R.; Wu, L.C.; Lee, T.Y.; Horng, J.T. Characterization and identification of antimicrobial peptides with different functional activities. *Brief Bioinform.* **2019**, *21*, 1098–1114. [CrossRef] [PubMed]
223. Timmons, P.B.; Hewage, C.M. ENNAVIA is a novel method which employs neural networks for antiviral and anti-coronavirus activity prediction for therapeutic peptides. *Brief Bioinform.* **2021**, *22*, bbab258. [CrossRef] [PubMed]
224. Kumar, M.; Thakur, V.; Raghava, G.P. COPid: Composition based protein identification. *Silico Biol.* **2008**, *8*, 121–128.
225. Gupta, S.; Kapoor, P.; Chaudhary, K.; Gautam, A.; Kumar, R.; Open Source Drug Discovery, C.; Raghava, G.P. In silico approach for predicting toxicity of peptides and proteins. *PLoS ONE* **2013**, *8*, e73957. [CrossRef] [PubMed]
226. Zhou, X.; Zhong, F.; Lin, C.; Hu, X.; Zhang, Y.; Xiong, B.; Yin, X.; Fu, J.; He, W.; Duan, J.; et al. Structure of SARS-CoV-2 main protease in the apo state. *Sci. China Life Sci.* **2021**, *64*, 656–659. [CrossRef]
227. Dai, W.; Zhang, B.; Jiang, X.M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F.; et al. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **2020**, *368*, 1331–1335. [CrossRef]
228. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; et al. Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, *582*, 289–293. [CrossRef]
229. Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H.J. PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Res.* **2005**, *33*, W363–W367. [CrossRef]
230. Mashiah, E.; Schneidman-Duhovny, D.; Andrusier, N.; Nussinov, R.; Wolfson, H.J. FireDock: A web server for fast interaction refinement in molecular docking. *Nucleic Acids Res.* **2008**, *36*, W229–W232. [CrossRef]
231. Kozakov, D.; Hall, D.R.; Xia, B.; Porter, K.A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S. The ClusPro web server for protein-protein docking. *Nat. Protoc.* **2017**, *12*, 255–278. [CrossRef] [PubMed]
232. Kim, D.E.; Chivian, D.; Baker, D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* **2004**, *32*, W526–W531. [CrossRef] [PubMed]
233. Zhao, Y.; Zhu, Y.; Liu, X.; Jin, Z.; Duan, Y.; Zhang, Q.; Wu, C.; Feng, L.; Du, X.; Zhao, J.; et al. Structural basis for replicase polyprotein cleavage and substrate specificity of main protease from SARS-CoV-2. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2117142119. [CrossRef] [PubMed]

234. Kaur, A.; Goyal, B. Identification of new pentapeptides as potential inhibitors of amyloid-beta(42) aggregation using virtual screening and molecular dynamics simulations. *J. Mol. Graph. Model.* **2023**, *124*, 108558. [[CrossRef](#)] [[PubMed](#)]
235. Shuaib, S.; Narang, S.S.; Goyal, D.; Goyal, B. Computational design and evaluation of beta-sheet breaker peptides for destabilizing Alzheimer's amyloid-beta(42) protofibrils. *J. Cell. Biochem.* **2019**, *120*, 17935–17950. [[CrossRef](#)]
236. Kaur, A.; Goyal, B. In silico design and identification of new peptides for mitigating hIAPP aggregation in type 2 diabetes. *J. Biomol. Struct. Dyn.* **2023**, 1–16. [[CrossRef](#)]
237. Crescenzi, O.; Tomaselli, S.; Guerrini, R.; Salvadori, S.; D'Ursi, A.M.; Temussi, P.A.; Picone, D. Solution structure of the Alzheimer amyloid beta-peptide (1–42) in an apolar microenvironment. Similarity with a virus fusion domain. *Eur. J. Biochem.* **2002**, *269*, 5642–5648. [[CrossRef](#)]
238. Barrow, C.J.; Yasuda, A.; Kenny, P.T.; Zagorski, M.G. Solution conformations and aggregational properties of synthetic amyloid beta-peptides of Alzheimer's disease. Analysis of circular dichroism spectra. *J. Mol. Biol.* **1992**, *225*, 1075–1093. [[CrossRef](#)]
239. Luhrs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Dobeli, H.; Schubert, D.; Riek, R. 3D structure of Alzheimer's amyloid-beta(1–42) fibrils. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17342–17347. [[CrossRef](#)]
240. Nanga, R.P.; Brender, J.R.; Vivekanandan, S.; Ramamoorthy, A. Structure and membrane orientation of IAPP in its natively amidated form at physiological pH in a membrane environment. *Biochim. Biophys. Acta* **2011**, *1808*, 2337–2342. [[CrossRef](#)] [[PubMed](#)]
241. Cao, Q.; Boyer, D.R.; Sawaya, M.R.; Abskharon, R.; Saelices, L.; Nguyen, B.A.; Lu, J.; Murray, K.A.; Kandeel, F.; Eisenberg, D.S. Cryo-EM structures of hIAPP fibrils seeded by patient-extracted fibrils reveal new polymorphs and conserved fibril cores. *Nat. Struct. Mol. Biol.* **2021**, *28*, 724–730. [[CrossRef](#)] [[PubMed](#)]
242. Wang, X.; Ni, D.; Liu, Y.; Lu, S. Rational Design of Peptide-Based Inhibitors Disrupting Protein-Protein Interactions. *Front. Chem.* **2021**, *9*, 682675. [[CrossRef](#)] [[PubMed](#)]
243. Mo, Q.; Xu, Z.; Yan, H.; Chen, P.; Lu, Y. VSTH: A user-friendly web server for structure-based virtual screening on Tianhe-2. *Bioinformatics* **2023**, *39*, btac740. [[CrossRef](#)] [[PubMed](#)]
244. Jofily, P.; Pascutti, P.G.; Torres, P.H.M. Improving Blind Docking in DOCK6 through an Automated Preliminary Fragment Probing Strategy. *Molecules* **2021**, *26*, 1224. [[CrossRef](#)]
245. Andrusier, N.; Mashiach, E.; Nussinov, R.; Wolfson, H.J. Principles of flexible protein-protein docking. *Proteins* **2008**, *73*, 271–289. [[CrossRef](#)]
246. Zhu, T.; Cao, S.; Su, P.C.; Patel, R.; Shah, D.; Chokshi, H.B.; Szukala, R.; Johnson, M.E.; Hevener, K.E. Hit identification and optimization in virtual screening: Practical recommendations based on a critical literature analysis. *J. Med. Chem.* **2013**, *56*, 6560–6572. [[CrossRef](#)]
247. Corbeil, C.R.; Williams, C.I.; Labute, P. Variability in docking success rates due to dataset preparation. *J. Comput. Aided Mol. Des.* **2012**, *26*, 775–786. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.