Current Topics in Pharmacology
Vol. 15, *2*, 2011

Review

In silico-based direct evolution of peptides and peptidomimetics in drug discovery

Gregorio Fernández-Ballester[1,*], Asia Fernández-Carvajal[1], Isabel Devesa[1], Belén Valenzuela[2], María José Duart[3], José Manuel González-Ros[1], and Antonio Ferrer-Montiel[1]

[1]Molecular and Cellular Biology Institute, University Miguel Hernandez, Elche, Spain
[2]Platform of Oncology, USP Hospital San Jaime, Torrevieja, Alicante, Spain
[3]Pharmaceutical Sciences Department, University Miguel Hernandez, Elche, Spain

ABSTRACT

Because drug discovery efforts have experienced a pronounced decline in productivity, novel approaches to the rational design of new drugs are being introduced and developed. An exciting solution is the use of natural or synthetic peptides and peptidomimetics targeting protein-protein interactions essential for signaling networks function. The combination of several bioinformatic approaches (docking, virtual screening, pharmacophore models, etc.) allows for the use of the vast amount of information on protein-protein interactions deposited in structural databases. In this respect, interacting peptides are susceptible to optimization in order to stabilize or disrupt protein-protein interactions, providing a promising use of peptides and derivatives as therapeutics. In this review we illustrate tools and strategies currently used in peptidomic drug discovery as well as trends in the area of molecular pharmacology.

KEYWORDS: drug discovery, peptide, peptidomimetic, computational approaches, homology models, docking, virtual screening, pharmacology, pharmacokinetics, ADME/Tox

*To whom correspondence should be addressed:
Instituto de Biología Molecular y Celular,
Universidad Miguel Hernández. Av.
De la Universidad s/n, 03202 Elche, Spain.
gregorio@umh.es

ABBREVIATIONS

ADME/Tox, absorption, distribution, metabolism, excretion and toxicity; PK, pharmacokinetic; VS, virtual screening; HTS, high throughput screening; SAR, structure-activity relationship

INTRODUCTION

Protein-protein interactions play a central role in many biological processes, such as signaling and regulatory networks. These interactions have been considered as attractive targets for the inhibition of specific pathways, commonly involved in many pathological processes. Targeting drugs at protein-protein interfaces have the advantage of modulating signal activity without interfering with the catalytic binding sites of the targets [1]. Nowadays, the stabilization [2] or disruption [1, 3] of the protein-protein interactions is being pursued. However, designing small molecules to inhibit protein-protein interactions is considered an enormous challenge [4]. Traditional High Throughput Screening (HTS) has had limited success to identify drugs targeting protein-protein interactions [5, 6]. Drug discovery efforts to find blockers for these targets have encountered several challenges: i) large interaction surfaces [7, 8]; ii) the surfaces are generally flat, with no clear grooves and pockets [3, 9, 10]; iii) limited structural information of binding surfaces; iv) large thermodynamic barriers for small molecule-protein binding; and v) absence of appropriate compounds in chemical libraries [11]. In spite of

this, small molecules still dominate the drug discovery for ion channels, receptors and transporters.

Biological drugs (such as peptides, peptidomimetics, antibodies and proteins) are now emerging as powerful tools for the discovery of inhibitors directed to protein-protein interactions. The use of antibodies as protein therapeutics against membrane-bound receptors or secreted proteins has been very successful [12] although limited to extracellular targets because of the difficulty for their intracellular uptake [13]. On the contrary, peptide and peptidomimetic therapies could, in principle, represent an extension of both, the small molecule and protein therapies, due to their unique structural properties and the recent advances in formulation, delivery [13, 14], and chemistry [15]. These molecules, that do not necessarily comply the Lipinski's [16] or Veber's [17] rules, are suitable for targeting large interfacial areas in protein-protein interactions. Conversely, peptides are being used to study protein-protein interactions [18] since they can be easily synthesized, hot spots are precisely identified, and post-translational modifications or non-natural amino acids can be introduced.

The advantage of peptides over protein or antibodies application is clear: Peptides can penetrate deeper into tissues, they are less immunogenic, and they have higher activity per unit mass and lower manufacturing costs [19]. The use of peptides over small molecules has also several benefits: Peptides offer great efficacy, selectivity and specificity [20]; the degradation products are amino acids, thus reducing the systemic toxicity [21]; peptides do not accumulate in tissues because of their short half-life; and small quantities of them are needed to modulate their receptor targets [19].

So, this work aims at reviewing the current methods and the computational techniques employed to derive inhibitory peptides and peptidomimetics from protein-protein interactions as a general mechanism to block the recognition partners. These compounds can be used as promising leads for the rational design of therapeutic drugs. Nowadays, the computational techniques employed in the peptidomic field include docking and virtual screening, virtual amino acid scanning, and pharmacophoric model evolution, along with the prediction of absorption, distribution, metabolism, excretion and toxicity, pharmacokinetic and pharmacodynamic properties of candidate lead compounds.

Nature of protein-protein interactions

Protein-protein contacts have mainly two models of interaction: Between two globular domains, or between a globular and a linear flexible peptide [22]. There is not an easy way to derive peptides from globular interactions, probably due to the fact that the interaction surface between globular domains is composed of residues which are brought together by the tertiary structure. Since important residues are not contiguous in the polymer chain, no linear sequence can be easily extracted. Furthermore, an inhibitory peptide should compete with a large interaction surface. On the other hand, a small number of hot spots are the major contributors of protein binding interaction. These hot spots represent less than half of the contact residues and are usually located at the centre of the interacting surface [4]. Therefore, the design of inhibitory peptides from this kind of surfaces by targeting only the hot spots is plausible [11]. The advantages are clear: First, the specificity, since they are derived from the native interaction; and second, the small size, which increases the competitive binding by raising the effective peptide concentration at the interface [4]. For these reasons, peptide fragments derived from crystal interfaces of protein-protein interactions are the major sources of inhibitory molecules [11]. The peptide conformation with the higher success has been α-helices [23, 24], probably due to the lower entropic cost paid upon peptide binding. Nevertheless, other conformational structures have been successfully assayed, such as hairpin structures and miniproteins [23, 24].

Interactions between globular proteins are often dominated by a single high-affinity linear peptide segment [25]. These are linear amino acid stretches derived from one of the partners in the interaction. These peptides derive from loops within globular domain, disordered linker regions or protein termini [25]. Within many of the abundant disordered regions in proteins, linear peptides are responsible for thousands of cell

interactions. Detailed studies on peptide-protein complexes have shown strong packing interactions and a large number of main chain-main chain and main chain-side chain hydrogen bonds [26], which would indicate that this kind of peptides are less hydrophobic to avoid aggregation [27] when compared to similar peptides in globular-globular interfaces. Small peptides (6 to 11 residues) usually pose up to 2-3 hot spots residues [27], which are crucial for the highly specificity observed in most peptides. The study of peptide-protein interfaces has also revealed that peptides are elongated structures that do not induce significant changes in their binding partners. The majority of these peptides remain unstructured in solution [28], and they only adopt a stable conformation upon binding to their partners. London *et al*. (2010) [25] analyzed the secondary structure of a set of linear segments from Benchmark 3.0 [29] and found that 76% displayed no defined secondary structure, 20% were α-helix, and only 4% were β-sheets. It is not clear whether or not the isolated inhibitory peptides adopt the same binding conformation as in the original domain, although docking experiments with *FlexPepDock* have shown that a large number of peptides presented a near-native conformation, concluding that they will adopt a similar structure when cut out of their protein context [25].

Deriving peptides from protein complexes

Approaches to discover lead compounds

Protein-protein interactions are therapeutic targets since these interactions govern cellular functions. Peptides can modulate protein-protein interactions as agonists or antagonists, making them attractive tools to identify lead compounds. In addition, peptides can also modulate the oligomerization state of proteins [18]. Figure 1 depicts schematically different ways to affect a protein-protein interaction by a peptide.

Targeting protein-protein interactions by large molecules (peptides and peptidomimetics) needs the adaptation of accompanying disciplines such as suitable methods of compound synthesis, development of biochemical assays to monitor protein-ligand modulation, and computational methods to analyze the interactions at atomic level [30]. Diverse studies have analyzed the protein-protein interaction networks [31]. Today, the structural representation of protein complexes is limited in the Protein Data Bank [32], and there is a need to combine experimental methods (X-ray, nuclear magnetic resonance (NMR), electron microscopy (EM)) and computational predictions (docking, virtual screening (VS), bioinformatics, etc.) to expand the structural information of protein assemblies [33]. A great challenge is the creation of high quality models from low-resolution images, homology models of separated fragments of proteins, and any other source of biochemical information on interaction, expression data, etc. [34]. Recent developments in the field of structural bioinformatics applied to the modeling of protein interactions and complexes, from large macromolecular machines to protein-protein or protein-peptide interactions, were reviewed by Stein *et al*. (2011) [35].

Some databases collect the available high resolution three-dimensional (3D)-structures of known protein-protein interactions: *3did* contains known high-resolution 3D structures [36]; *PepX* is a database of peptide-protein complexes [37]; *BriX* has protein building blocks for structural analysis, modeling and design, containing protein fragments from 4 to 14 residues [38]; *3DComplex* offers three dimensional complexes classification [39]. Regarding the structure-based prediction of peptides to inhibit protein-protein interactions, it is interesting to note that the “absence” of structural information is the general rule for a vast amount of proteins, including membrane proteins and channels. There is, however, a key factor that helps us to escape from this and allows the design of blocking molecules: Peptide binding motifs strongly resemble the intramolecular packing motifs [40]. Thus, looking at the inside of crystallized monomeric proteins, it is possible to harvest the interacting motifs in order to model protein-peptide interactions, thus significantly increasing the templates available. Interestingly, by using this approach, the sequence homology is not a pre-requisite [40].

There are mainly two strategies to discover peptide leads: Sequence- and structure-based approaches, which are commonly used for the rational design of peptidic modulators of protein-

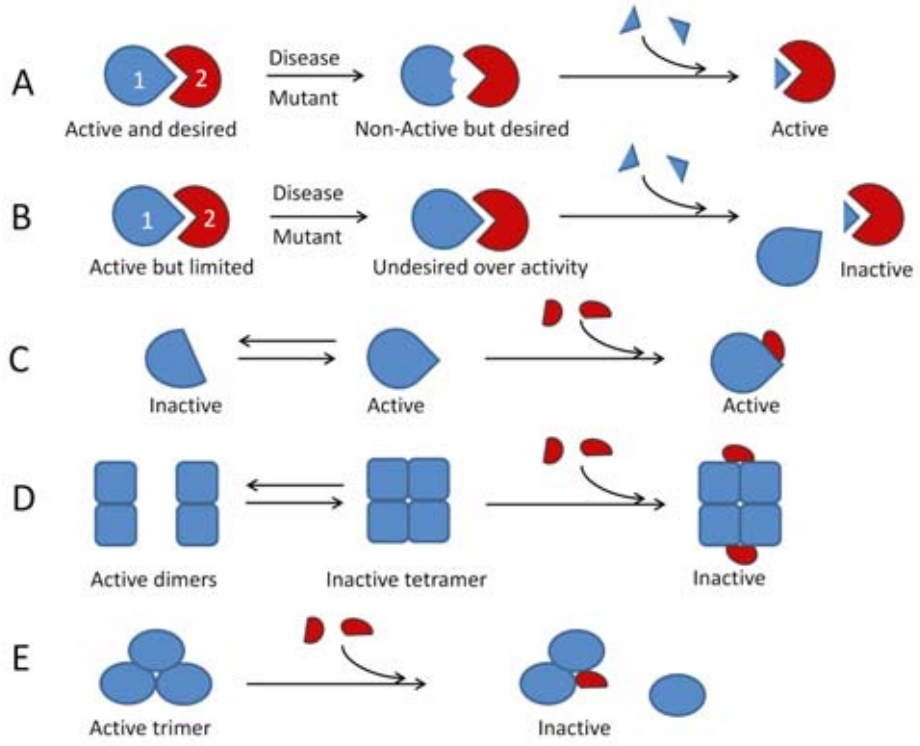

Figure 1

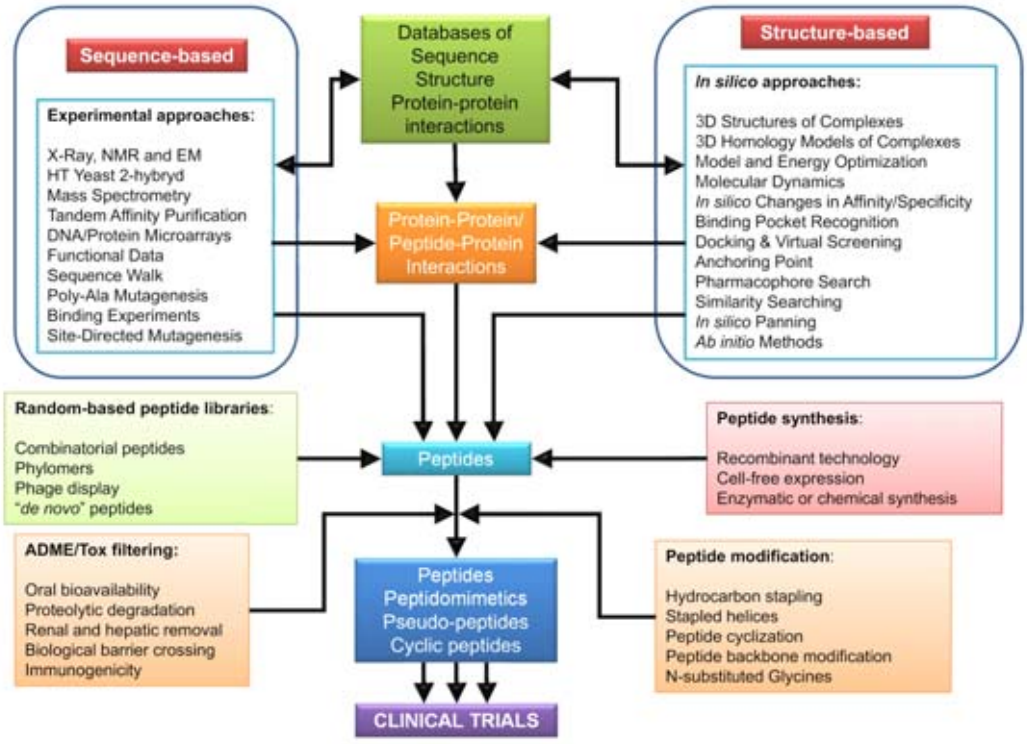

Figure 2

protein interactions. Figure 2 gives a global overview of these approaches.

Sequence-based peptide design

Several experimental methods are suitable for the discovery of protein-protein interactions by screening multiple proteins against a single one. In this sense, HTS and yeast two-hybrid, mass spectrometry combined with tandem affinity purification, DNA or protein microarrays are able to detect and identify direct protein-protein interactions. In addition, databases of already known protein-protein interactions are employed to derive unknown interactions by homology. Once identified a given interaction, the typical strategy to derive peptides able to inhibit the protein-protein interaction is to mimic the sequence of one interacting partner [41], which in turn provides the starting point for the design of high affinity blockers [25].

Inhibition of kinases is an excellent example of targeting protein-protein interactions by the sequence-based approach. Protein kinases regulation is pivotal for cellular function, and its modulation has been pursued for many years. However, actual inhibitors are small molecules directed to the ATP binding site [42] with poor selectivity [43, 44]. Looking at the protein-protein interfaces of kinases, several regions can be selected, such as the catalytic and regulatory domains, the SH2-ligand complex, the substrate binding pocket, and the activation loop. There are many studies dealing with the search of these inhibitors, especially with kinase-substrate interactions, since the substrate binding site of protein kinases is less conserved than the ATP binding site, thus providing an opportunity to design blockers with enhanced specificity [45-47]. Fujiwara *et al*. (2011) [48] have employed a semi-rational strategy, combining phage-displayed libraries and de novo designed peptides. They constructed a peptide library based on the loop region of a helix-loop-helix motif, and screened against Aurora-A, a serine/threonine protein kinase, and found selective peptides with moderate inhibitory activity that in turn could serve for lead optimization.

Other strategy includes phylomers, which represent a new class of peptides derived from natural protein fragments. Libraries contain billions

Legend to Figure 1. Peptides and peptidomimetics can affect protein-protein interactions in different ways. (A) Peptides can act as agonists by replacing the action of protein 1 over protein 2 in cases where protein 1 is less- or non-functional, i.e.: Disease, mutagenesis, or low expression. (B) The over activation of a signaling pathway can be deactivated by antagonists peptides by mimicking the interaction of protein 1 and 2. Panel (C) describes the tertiary equilibrium shift induced by peptides. Protein structural instability is rescued by interacting peptides that stabilize the active state [145]. (D) Quaternary equilibrium shift induced by peptides inactivates the protein by facilitating the oligomerization state [129]. Both cases, C) and D), are reviewed in [18]. (E) Disruption of an active oligomer by intercalation and association, which facilitates de dissociation of a monomer [4].

Legend to Figure 2. Design of peptides and peptidomimetics for therapeutic purposes. This flux diagram represents an overview of the methodologies commonly employed to design peptides and their derivatives to interfere protein-protein interactions. Previous knowledge of protein sequences, structures and/or protein-protein interactions contained in the databases is exploited to extract useful information. Two main approaches are depicted: i) the experimental (sequence-based) approach (left) takes advantage of the information stored in databases (sequences, interaction of homologues, models, etc.), and uses common experimental techniques able to determine new protein-protein interactions. The information obtained contributes to the increase in knowledge of protein-protein interactions and at the same time feeds-back the databases; and ii) the *in silico* (structure-based) approach (right), which exploits all three-dimensional information available for protein-protein or peptide-protein interactions. This information is used to predict and characterize new protein interaction, further increasing knowledge and feeding databases. In addition, random libraries of peptides can be generated by either experimental or *in silico* techniques, and tested using diverse assays. All these approaches allow for the selection of peptides susceptible to become lead compounds. The synthesis of single peptides or peptide libraries is followed by ADME/Tox filtering to evaluate pharmacokinetic parameters. Additional chemical modifications are usually necessary to improve the pharmacokinetic properties of the lead compounds. Those passing all requirements go to clinical trials.

of distinct phylomers and represent a large source of diverse natural secondary and tertiary peptide structures. Watt (2006) has derived libraries of phylomers comprising random and structured peptides encoded by natural genes of bacterial genomes, obtaining hits against different targets with high affinity [11]. These sequences can be refined for affinity and/or specificity with standard mutagenesis and *in vitro* evolution techniques [49]. The approach provides a rich source of peptides that interact specifically and with high affinity with human proteins, allowing the comprehension of discrete interactions within the interactome, and the development of effective drugs targeted to particular protein functions.

A different approach for protein ligand screening was the use of limited proteolysis and MALDI-TOF mass spectrometry [50]. In this work a combinatorial library of 8000 peptides attached to a poly-Pro framework was screened against the SH3 domain of the Abl protein. The results were later contrasted with computer modeling, demonstrating the sensitivity, specificity, speed and low sample consumption of the methodology.

Structure-based peptide design

The computational design of peptides to interfere with protein interactions has been recently reviewed by Vanhee *et al*. (2011) [27]. The most common methodology directly uses a peptide-protein complex structure or employs this structure as template to model a relative homologue. In addition, these complexes can be used to optimize the peptide interaction by fixing the ligand backbone, and by using a sequence space search algorithm to change the ligand affinity and/or specificity [51, 52].

There are many successful examples of peptides interfering with protein interfaces, either as agonists or antagonists that have been derived with computational tools taking advantage of the 3D structure. For instance, the activation of the formyl-peptide receptor-like 1 (FPRL1) has beneficial effects in the therapy of inflammatory diseases. Hecth *et al.* (2009), using a computational platform of prediction, derived a 21-residue peptide that activated the receptor and displayed anti-inflammatory activity *in vivo* [53].

Other strategies, as noted by Vanhee *et al*. (2011) [27], tackle the peptide backbone flexibility by using multiple domain-peptide complexes that allow the superimposition of several ligands on the target. This approach can be only addressed to certain proteins highly represented in the Protein Data Bank. As an example, Fernandez-Ballester *et al*. (2009) [51] modeled most of the SH3 domains of *S. cerevisiae*, and collected 29 ligand conformations from structurally available SH3 domains, to construct all possible ligand-SH3 combinations. These complexes were used firstly to identify key residues determining loop conformation in SH3 domains; and secondly to construct position-specific ligand binding matrices by means of FoldX [54, 55] to determine which sequences were favorable for every SH3 domain. The results, which were validated with available experimental data, opened the way for genomic-wide scale predictions. In fact, the yeast SH3, as well as many other peptide interacting domains (PDZ, WW, SH3, SH2, 14-3-3, etc.) from different organisms was compiled in the database ADAN (http://adan-embl.ibmc.umh.es/) together with its predicted sequences and putative partners [56].

Smith & Kortemme (2010) computationally predicted the sequence space of peptides recognized by PDZ domains, and validated them with a large set of phage display experimental data [57]. The study generated an ensemble of backbone conformations for profile prediction using Monte Carlo simulations and a genetic algorithm-based specificity prediction. The authors were able to predict wild type with 70-80% accuracy, and mutations that increased the binding affinity relative to the starting structure, indicating that the incorporation of backbone sampling improved the accuracy of the predictions. Raveh *et al*. (2010) used a Rosetta *FlexPepDock* protocol for refining peptide-protein coarse models, producing high resolution models of sub-angstrom backbone quality with small deviations from the native structure [58].

Binding pocket recognition, docking and virtual screening

In the absence of detailed information on protein-protein interaction, an anchor residue in the binding pocket and implicit backbone movements are used by certain algorithms to find peptide specificity that is filtered by experimental data

[59]. However, computational costs of this method strongly limit its use, opening the door to other methods, such as binding pocket recognition, docking and virtual screening, as alternative approaches in a wide-scale screen.

The localization of the binding pocket is important because it can be targeted by drug discovery. Several approaches to determine the structural active sites have been reported [60], suggesting many parameters to characterize binding pockets, as pocket compactness, surface roughness and complexity, and total surface area [61]. As an example, *ICMPocketFinder* [62] was used to explore pockets in modeling neuro-transmitter transporters [63]. This strategy defined the determinants of rho1-GABA(C) receptor assembly by detecting the binding surfaces on the ligand-gated ion channel located in the transmembrane region [64]. Unal *et al*. (2010), using known protein-peptide inhibitors complexes, determined the binding site on the protein via a coarse-grained Gaussian Network Model [65]. Interestingly, Tseng & Li (2011) developed an evolutionary approach to predict the binding site residues of proteins from primary sequences [66], which is of particular interest for membrane protein, such as ion channels, since protein databases have scarce structural information on them.

The ability for docking methods to place ligands into a known native structure has been evaluated in an excellent review [67] as well as other reports [68-71], including peptidic modulators [30]. A successful docking approach includes flexible ligand search approximation, which takes into account the ligands degree of freedom, typically higher for peptides than for small molecules. There are three categories of algorithms that include ligand flexibility: Systematic, random or stochastic, and simulation methods [67]. However, few procedures have been developed specifically for peptide docking, with impaired results. A method that uses a heuristic search of the potential energy succeeded in docking tripeptides, but failed with longer peptides [72]. A study using Monte Carlo sampling failed to dock peptides to nuclear receptors [73]. Nevertheless, other authors succeeded in docking peptides to proteins by the anchoring point approximation. This is based on the fact that many linear motifs have a well defined docking region, but a more variable area for the residues flanking the motif [74]. The work by Niv & Weinstein (2005) [75] is based firstly on the knowledge of an anchoring point between the peptide (the α carbon of the peptide C-terminus) and the protein (PDZ domains), and then by the use of simulated annealing molecular dynamic approach. The conformational space is explored by heating the complex and cooling different backbone conformations. After side chains optimization, the complexes are minimized and scored by theoretical binding energy, yielding acceptable results [75]. Other studies accurately predicted docking and optimized peptides against MHC by means of a flexible all-atom model of the entire peptide and an energy minimization method [76]. In general, the limitation of the conformational space using constraints is often crucial for successful ligand docking. These constraints are usually derived from experiments such as NOE data [77] or any other source of biological information, like knowledge-based conserved positions [75].

There are recent examples of successful peptide docking that show the reliability of the method. Unal *et al*. (2010) used a novel de novo peptide design approach to block diverse protein-protein interactions [65]. The method sequentially constructed the peptide by generating all possible peptide pairs and determining the binding energies by means of AutoDock [78]. A Hidden Markov model using the Viterbi algorithm decoding is employed to obtain the best fitting peptide for a given surface [65]. London *et al*. (2010) derived candidate inhibitory peptides by first screening for high-affinity linear segments on the interfaces of globular domains, and then by using peptide docking experiments to assess the peptide interaction on the same region, irrespective of the context of the original peptide domain [25]. The authors employed a protocol for flexible peptide docking, *FlexPepDock* [58], and an energy funnel analysis to present an elaborated framework for *in silico* selection of the inhibitory peptides derived from multiple protein interactions. Hussain *et al*. (2011) combined molecular dynamic simulations with virtual peptide screening to identify a series of potential binders for the Elk-1 transcription factor

dimer interface [79]. A collection of conformations representing the loop dynamics involved in dimerization was exhaustively screened against peptide libraries. The positive tri-peptides made specific interactions with certain residues that are pivotal to the dimeric interface, indicating that the incorporation of dynamic fluctuations in the receptor can help to discover inhibitors. Pang *et al*. (2011) found a short peptide that inhibited cyclophilin A in the same range of binding affinity as the positive control cyclosporine A inhibitor. This short peptide could replace cyclosporine A as an immunosuppressant drug with better oral availability, solubility and less toxicity in clinical applications [80].

Synthesis, modification and drug-like properties of peptides

Peptides are traditionally considered to be poor drug candidates because of their propensity to be metabolized as well as their low oral bioavailability. Recent advances in synthesis and stabilization of peptides have boosted their use to interfere protein-protein interactions [18]. The technologies employed to synthesize these molecules depend primarily on peptide size: Recombinant technology, cell-free expression, enzymatic or chemical synthesis. The solid-phase peptide synthesis [81] revolutionized and made available their chemical synthesis by the standardization and automation of the protocols. In recent years, new synthetic strategies to improve productivity and reduce metabolism have been developed. Advances in peptide chemistry have improved protocols, allowing for the synthesis of peptides longer than 200 residues [82].

The ADME/Tox (absorption, distribution, metabolism, excretion and toxicity) parameters of peptides are crucial to define the disposition of a drug candidate and its potential therapeutic effects [19]. As already mentioned, the main limitations for the use of peptides as therapeutics are low oral bioavailability and rapid degradation by proteolytic enzymes (digestive system and blood plasma), but also rapid renal and hepatic removal from circulation, poor ability to cross physiological barriers, eventual risk of immunogenicity, and high production costs [19]. The bioavailability and distribution is determined by a combination of physicochemical properties: Chemical stability, solubility, hydrophobicity, hydrogen bonds formation, and metabolic stability. Nevertheless, peptides are amenable to acquire drug-like physicochemical and pharmacokinetic (PK) properties such as solubility, lipophilicity, metabolic stability, and bioavailability. After an active peptide is identified, it is usually subjected to chemical modifications to increase its potency by stabilizing or improving specific properties [18]. The chemical optimization is made based on structure-activity (SAR) and/or quantitative structure-activity (QSAR) relationships of modified peptides to assess improvements in bioavailability, resistance to proteases, changes in affinity and/or selectivity.

Peptidomimetics, pseudo-peptides and cyclic peptides

The chemical strategies trying to overcome the above described limitations include the development of peptidomimetics, pseudo-peptides and cyclic peptides. Exhaustive revision of amino acid modifications, backbone modifications, global restrictions by cyclization and synthetic backbone scaffolds can be found in Grauer & Konig (2009) [83] and Vagner *et al*. (2008) [84].

Peptidomimetics are non-protein molecules designed to mimic a peptide. For the development of peptidomimetics it is necessary to understand the forces involved in protein-protein interactions. The strongest interactions between peptides and proteins are based mainly on side chain-side chain interactions, which indicates that peptide backbone itself is not essential for high affinities [83]. This allows for the replacement of the amide backbone total or partially by other chemical structures. They should conserve selectivity or potency while overcoming the susceptibility to proteolysis or poor bioavailability of the peptides. Several examples are reviewed in Zinzalla & Thurston (2009) [85].

A pseudo-peptide includes a chemical modification in the peptide bond. Peptoids are peptide-based backbone and N-substituted glycines for side chain residues, resulting in complete protease resistance [86]. A positional scanning library of N-alkylglycine trimers containing more than 10,000 compounds allowed for the screening

against a variety of biological targets [87-89], including a panel of Gram-positive and Gram-negative bacteria [90] that led to the identification of hits exhibiting antimicrobial activity [91].

However, the most successful way to achieve peptide stabilization is via cyclization [92], which imposes conformational constraints to the peptides. Peptide cyclization offers many advantages, such as an increased stability [19], protection against proteases [93], and an increased affinity [94]. This last effect is probably due to reduced entropy of the unbound state as a cyclization consequence. The cyclization also offers other advantages in terms of bioavaliability and cell permeability [19, 93]. As an example, Tal-Gan *et al*. (2001) performed an extensive SAR study of a potent peptide-based protein kinase B/Akt inhibitor [47]. To overcome the lack of pharmacological properties, the authors synthesized cyclic backbones peptide libraries with varying modes of cyclization, bridge chemistry, and ring size. This work found peptides 10-fold more potent than the corresponding linear peptide, and they became promising lead candidates with enhanced pharmacological properties. The growing data in favor of these macrocyclic compounds was reviewed by Driggers *et al*. (2008) [95].

Other way for peptide stabilization, i.e. α-helices, is by covalent stapling, where one of the main-chain hydrogen bond is replaced by a covalent bond [96], or by formation of a hydrocarbon bridge between two consecutive helix turns [97]. Other peptide conformation, such as bicyclic β-turn analogues of Leu-enkephalin, was successfully designed and synthesized [98], showing significant biological activity. Furthermore, to provide proteolytic stability, miniature proteins are used as scaffolds to graft a protein binding site [85]. Finally, the SPOT synthesis on cellulose sheets [99] allowed for the development of stable peptide arrays for the study of protein-peptide interactions.

Transforming peptides into small molecules

Although this review is focused on the use of peptides to overcome the small molecules, we should keep in mind that sometimes there are powerful reasons to continue using small molecules: Practical (easier synthetic protocols, more efficiency) and economical (lower costs). In fact, there are examples in which small molecules bind to the hot spot in the contact surface of the proteins with drug-like potencies. These examples have been reviewed by Wells *et al*. (2007), including studies on small molecule binders to interleukine 2, B-cell Lymphoma 2 (Bcl-XL), human double minute 2 (HDM2), human papilloma virus (HPV E2), ZipA or tumour necrosis factor (TNF) [4]. In most of the cases, the small molecules were in the mid- to low-nanomolar range, and comparable to the binding affinity observed for the native protein partners. In addition, these molecules bound deep into the interacting surfaces, which is indicative of the adaptability of the contact surfaces when the partners of a interaction are separated, probably due to motions of side chains and small loops perturbations [4]. Based on structure-activity relationships, non-peptide molecules or peptidomimetics can be derived and checked for effects on protein-protein interactions.

A computational tool developed to facilitate the search of molecules mimicking peptides is the *3D Pharmacophore Model*. A protein-peptide structure enables for the construction of a pharmacophore model, which describes the nature and the location of functional groups involved in peptide-target interactions, including hydrogen bond network, hydrophobic and electrostatic interactions. [100]. The pharmacophore model can then be used to design small molecules mimicking the peptide [101, 102]. Briefly, the identification of a 3D structure of a peptide-protein interaction is followed by the construction of a pharmacophore query and a conformational library of small compounds. The pharmacophore query is searched against a library and results are processed to establish a ranking of hits. Positive molecules are later evaluated and tested for correct matching by means of docking [102].

Another standard tool commonly used in computational screening is the ligand-based method *Similarity Searching* [103, 104]. Basically, a 2D similarity fingerprint is applied to peptide-like molecules as an alternative method to facilitate the peptide to peptidomimetic transition. Two-D fingerprint is calculated from the 2D graph representation of the molecules, and is represented as bit strings, where an individual bit denotes a

specific chemical feature or a molecular descriptor. This method is also capable of differentiating between peptide-like molecules having different biological activities [104].

Qvit *et al.* (2008) presented a new approach for the conversion of active sequences of proteins and peptides into small molecules. In this approach they constructed a library of macrocyclic disulfide molecules, and preserved the active pharmacophores of the parent peptide to perform a systematic search for macromolecules in which the pharmacophores are in an appropriate conformation for biological activity [105].

Computational tools

From a computational point of view, there are many tools to accomplish the generation of peptides and derivatives, which in turn can be used for docking and virtual screening, evaluated by binding energy and filtered by ADME/Tox properties. There are long lists of software, free or commercial, developed for these purposes. Table 1 shows several examples of software and servers available. Villoutreix *et al.* (2007) reviewed free resources to assist structure-based experiments [106].

Selected working examples

Modulation of receptors and channels

Cell-penetrating membrane-tethered peptides have been used to target intracellular receptor domains proximal to the plasma membrane. They were derived from the intracellular loops of G protein-coupled receptors and lipidated with a palmitate group [107]. These peptides, coined as pepducins, demonstrated their ability to modulate the activity of G protein-coupled receptors (GPCR) with high efficiency and selectivity [107, 108]. Notably, some of them became leads for drug development. The recent identification of protease-activated receptor 1 (PAR1) as a potential therapeutic target in lung cancer was accomplished by the use of cell-penetrating pepducins, generated from the first and the third intracellular loop of PAR1, which were able to block PAR1-RK1/2 signaling pathways [109].

Recently it was shown that peptides can be successfully used to target well-defined protein-protein interactions involved in TRPV1 channel function [110]. The TRPV1 thermosensory channel has a cytosolic domain, referred to as the TRP region that is involved in subunit oligomerization and functional coupling of stimulus sensing and gate opening. In this sense, interfering with this protein interface could be used to modulate channel function. It was demonstrated that palmitoylated peptides (named as TRPducins) patterned after the membrane proximal TRP domain of TRPV1, behave as moderate and selective channel inhibitors. Figure 3 illustrates the interacting region between TRPducin and TRPV1 channel.

Yin *et al.* (2007) described a computational approach to design peptides that specifically recognize transmembrane helices of natural proteins [111]. This methodology selected the backbone of a pair of helices derived from membrane protein structures, and threaded the sequence of the target onto one of these helices. The selection of the amino acid sequence in the accompanying helix was made by means of a side-chain repacking algorithm. The authors derived peptides that specifically interacted against α-IIb and αV integrins *in situ* in the plasma membrane of platelets.

Exocytosis

Neuronal exocytosis is a cascade of protein-protein interactions that involve several proteins. At the center of this process is the SNARE protein complex, which includes the plasma membrane proteins SNAP-25 and syntaxin, and the vesicle-associated protein synaptobrevin. This complex is responsible for vesicle docking and fusion in synaptic terminals. The identification of SNARE complex modulators inhibiting exocytosis was addressed by Blanes-Mira *et al.* (2003, 2004) [112, 113]. Synthetic peptides patterned after the N-terminus of the SNAP25, a component of the SNARE complex, were revealed as potent inhibitors of the SNARE complex formation [113]. In another study to discover SNARE peptide inhibitors, an α-helix-constrained combinatorial peptide library was synthesized in a positional scanning format, and assayed to prevent the formation of SNARE *in vitro*. The most active 17-mer peptide abrogated the Ca^{2+}-dependent release

Table 1. The table compiles selected examples of 3D databases as well as available software tools and servers to generate, edit, or modify peptides, comprising computational methods such as molecular dynamics, homology modeling, binding site prediction, docking and virtual screening, ADME/Tox prediction, etc., among other tools.

3D protein databases	
Protein Data Bank http://www.pdb.org/pdb/home/home.do	Databank of experimentally-determined structures of proteins and complex assemblies
3did http://www.3did.com/	Database containing known high-resolution 3D structures [36]
PepX http://pepx.switchlab.org/	Database of peptide-protein complexes [37]
BriX http://brix.crg.es/	Protein building blocks for structural analysis, modeling and design [38]
3DComplex http://www.3dcomplex.org/	Three dimensional complexes classification [39]
ADAN http://adan-embl.ibmc.umh.es/	Prediction of sequences and putative partners in protein-peptides complexes [56].
Edition and modeling of molecules	
Pymol http://pymol.org/	Molecular visualization system written in Python.
YASARA http://www.yasara.org/	Molecular visualization program for displaying and building molecules, multiple sequence alignments, molecular dynamics, docking, etc.
Modeller http://salilab.org/modeller/	Software for producing homology models of protein tertiary structures
I-TASSER http://zhanglab.ccmb.med.umich.edu/I-TASSER/	Internet service for protein structure and function predictions
SWISS-MODEL http://swissmodel.expasy.org/	Automated protein structure homology-modeling server, accessible via the ExPASy
Molecular dynamics	
NAMD http://www.ks.uiuc.edu/Research/namd/	Free parallel molecular dynamics designed simulation of large biomolecular systems
GROMACS http://www.gromacs.org/	Chemical Simulations. Free open source molecular dynamics simulation package
Binding site prediction	
ICM-PocketFinder http://www.molsoft.com/technology.html	Binding site predictor provided by Molsoft
PocketPicker http://gecco.org.chemie.uni-frankfurt.de/pocketpicker/download.html	Program for the analysis of ligand binding-sites with shape descriptors
FINDSITE http://cssb.biology.gatech.edu/findsite	Threading-based binding site prediction/protein

Table 1 continued..

Docking and virtual screening	
Autodock http://autodock.scripps.edu/	Docking software with flexible ligand and protein side chains capabilities
MGL Autodock Tools http://mgltools.scripps.edu/	Graphical front-end for setting up and running AutoDock.
GOLD http://www.ccdc.cam.ac.uk/products/life_sciences/gold/	Genetic algorithm-based docking program. Flexible ligand
MOE http://www.chemcomp.com/	Multiple purpose suite of chemistry tools like ligand-receptor docking, ligand optimization in pocket, ligand & structure-based scaffold replacement, pharmacophore search, etc.
FlexPepDock http://flexpepdock.furmanlab.cs.huji.ac.il/	High-resolution peptide docking (refinement) protocol
Binding free energy estimation	
X-score http://sw16.im.med.umich.edu/software/xtool/	Program for computing the binding affinities of ligand-target molecules
FoldX http://foldx.crg.es/	Program for calculating binding energy of proteins, complexes and mutants [54, 55]
DrugScore^{ONLINE} http://pc1664.pharmazie.uni-marburg.de/drugscore/	Web interface for the scoring functions DrugScore^{CSD} and DrugScore^{PDB}
ADME toxicity	
ADMET Predictor http://www.simulations-plus.com/Products.aspx?grpID=1&cID=11&pID=13	Software for advanced predictive modeling of ADMET properties
PrologP/PrologD http://www.compudrug.com/?q=node/42	ToolServer for predicting the logP/logD values using linear and neural network methods
ToxPredict http://apps.ideaconsult.net:8080/ToxPredict	Web to estimate toxicological hazard of a chemical structure
PharmMapper http://59.78.96.61/pharmmapper/	Web-server designed to identify target candidates for small molecules
Free ADME Tools http://www.simcyp.com/ProductServices/FreeADMETools/	ADME Prediction Toolbox of the SimCYP application
Peptides extraction and modification	
PeptideMine http://caps.ncbs.res.in/peptidemine/index.html	Server for design of peptides from protein-protein interactomes [139]
DockoMatic http://dockomatic.sourceforge.net/	Server for generating cyclic peptide analog structure files based on protein database [140]

Table 1 continued..

CycloPs http://bioware.ucd.ie/	Webserver for generating virtual libraries of constrained peptides [141]
Small molecules from peptides	
SuperMimic http://bioinformatics.charite.de/supermimic	Tool for identifying compounds that mimic parts of a protein [142]
SAAMCO http://bioware.ucd.ie/	Screening of motifs with known structures against bioactive compound databases [143]
pepMMsMIMIC http://mms.dsfarm.unipd.it/pepMMsMIMIC	Web-oriented peptidomimetic compound virtual screening tool [144]

of L-[^{3}H]-glutamate in intact hippocampal neurons. Interestingly, this study showed that the discovery of peptide sequences is not restricted to those that mimic domains of SNARE proteins [112].

Other studies showed the involvement of regulated exocytosis with the inflammatory sensitization of TRPV1 ion channel [114]. The small peptide EEQMRR, patterned after the SNAP-25 protein, was able to inhibit TRPV1 translocation and CGRP release in sensory neurons, and displayed analgesic activity. These findings supported the tenet that SNARE complex-mediated exocytosis of TRPV1 may be a valid therapeutic target to treat inflammatory pain.

Cell growth

The interaction between human p53 and MDM2 is a key event in controlling cell growth. Different studies have suggested that a p53 mimic would be sufficient to inhibit MDM2 and reduce cell growth in cancerous tissue. Thus, the inhibition of the MDM2-p53 interaction, and the reactivation of p53 function is a milestone in controlling cell growth [115]. It has been described that a p53 homologue is sufficient to induce p53-dependent cell death in cells overexpressing MDM2 [116]. A peptide as short as 6 residues could bind to MDM2 in the same manner [117]. Furthermore, chemical modifications of that 6-residue peptide can dramatically increase its inhibitory activity [118]. Zhong and Carlson (2005) examined the native binding interface of the MDM2-p53 complex, as well as the effects of mutants, by molecular dynamics simulation and alanine scanning [119]. They designed a mimic of p53 based on a β-proline (isomer of α-proline). The difference in chirality allowed peptides to resist hydrolysis by proteases in the body, giving them different absorption, distribution, metabolism, and excretion (ADME) properties. More recently, Phan *et al*. (2010) used a crystal structure of an inhibitory peptide directed against MDM2 to design and test several peptides with inhibitory properties [120]. This study discovered a peptide with a 5-fold increase in potency, as well as the key molecular features responsible for the enhanced affinity. Finally, Hu *et al*. (2011) described a new class of oligomers of N-acylated-N-aminoethyl amino acids, named AA peptides, that inhibited p53-MDM2 interaction with significant activity and specificity [121]. The efforts to inhibit the mentioned p53-MDM2 complex with peptoids and other molecules of different nature have been reviewed by Murray & Gellman (2007) [115]. All these molecules preserve the structural features of peptides, but are resistant to cell degradation.

Another example is protein kinase CK2 (also known as casein kinase II), an ubiquitous eukaryotic ser/thr protein kinase present in the nucleus and cytoplasm. CK2 is known to phosphorylate more than 100 substrates, many of which are involved in the control of cell division and in signal transduction. Laudet *et al*. (2007) studied the multimeric organization of protein kinase CK2 holoenzyme complex *in vitro* by a combination of site-directed mutagenesis, binding experiments and functional assays [122]. Using the crystal structure of the CK2 holoenzyme, they

designed an 11-mer peptide variant derived from the interface contact of CK2α/CK2β subunits, which was able to antagonize the interaction between the CK2 subunits, and to inhibit the assembly of the CK2 holoenzyme complex, *in vitro* and *in vivo*.

Interleukin-6 (IL-6) signal is transduced through a membrane glycoprotein, gp130, which associates with IL-6 receptor (IL-6-R). This cytokine acts on a wide range of tissues with a variety of biological activities, including cell proliferation. Since viral IL-6 has been shown to mimic human IL-6 functions, Sudarman *et al*. (2008) designed and synthesized peptides based on the crystal structure of extracellular domains of gp130 in complex with viral IL-6 [123]. These peptides were shown to block the interaction of gp130 with viral IL-6, as well as the stimulation of viral IL-6-induced cell proliferation.

Apoptosis

Apoptosis, or programmed cell death, is a critical cell process in normal development and homeostasis of multicellular organisms. Inappropriate regulation of apoptosis has been implicated in many human diseases, including cancer. Thus, targeting critical apoptosis regulators is an attractive approach for the development of new classes of therapies. A successful example is the design of an inhibitor of XIAP, a central apoptosis regulator that inhibits caspase-3/-7 through its BIR2 and BIR3 domains. The natural inhibitor, Smac, antagonizes XIAP by targeting both BIR2 and BIR3 domains. Zobel *et al*. (2006) designed a potent second mitochondrial activator of caspases (Smac) mimetic by using the crystal structure of a peptide displaying high binding affinity, but no measurable biological activity [124]. The authors translated the key components of binding of the peptide onto a non-peptide scaffold with enhanced drug-like properties. Similarly, Sun *et al*. (2007) synthesized and characterized a non-peptide, cell-permeable, small-molecule which mimics Smac protein for targeting XIAP, being 7000 times more potent than the natural Smac peptide [125].

BCL-2 family proteins constitute a critical control point for the regulation of apoptosis. Protein interaction between BCL-2 members is a prominent mechanism of control and is mediated through the amphipathic α-helical BH3 segment, an essential death domain. Walensky *et al*. (2004) induced the apoptosis *in vivo* by developing a hydrocarbon stapled BH3 helix being able to bind to the Bcl-2 family with increased affinity [126].

Another example is the protein complex formed between different cyclins and cyclin-dependent kinases, which are pivotal in cell cycle regulation. Canela *et al*. (2006) employed synthetic combinatorial peptide libraries and discovered a D-amino acid hexapeptide, non-competitive for either ATP or histone H1, that interfered with the formation of the cdk2-cyclin A complex. Furthermore, a cell-permeable derivative of this peptide induced apoptosis and inhibited proliferation of tumor cell lines [127].

Immunodeficiency

The integration of viral DNA into the host chromosome, mediated by the enzyme integrase, is an essential step in the HIV life cycle. The enzyme has no mammalian counterpart, converting it in an attractive target for antiviral drug design. Li *et al*. (2006) employed the “sequence walk” strategy across the entire 288 residues of enzyme [128]. The derived peptides, encompassing conserved amino acids and residues known to be important for catalytic activity, were able to inhibit the activity of the HIV-1 integrase protein. Hayouka *et al*. (2007) presented a strategy for inhibiting proteins by “shiftides”, which are ligand peptides that specifically bind to an inactive oligomeric state of a disease-related protein and modulate its activity by shifting the oligomerization equilibrium of the protein (Figure 1D) [129]. They were able to block the integrase enzyme by using peptides derived from its cellular-binding protein, LEDGF/p75, which specifically inhibited integrase activity by a non-competitive mechanism. Also working with the enzyme integrase, Zawahir & Neamati (2006) derived peptides from the HIV-1 HXB2 Pol gene sequence [130]. They tested them for inhibitory activity against wild-type and mutant integrase, and found that the most potent blocker peptide corresponded to a region of the reverse transcriptase subunit of the Pol protein.

The active site of HIV-1 protease has been targeted for inhibitor design, resulting in potent blockers [131]. However, their use has been

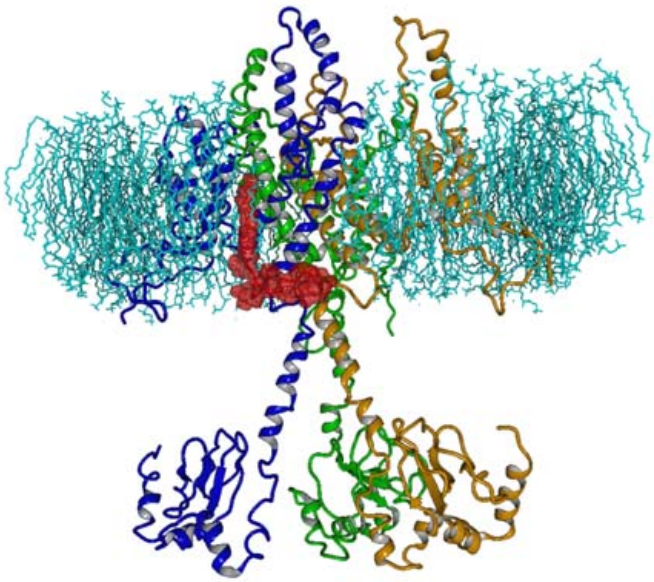

Figure 3. TRPducins modulate the TRPV1 channels. The figure represents the full model of TRPV1 channel inserted in a lipid bilayer [146]. The TRPV1 channel tetramer was represented in ribbons, the subunits were drawn in different colors, and lipid bilayer was colored in cyan. A whole subunit, as well as the big cytoplasmic N-terminus, has been removed for clarity. The transmembrane regions span the membrane, while C-terminus protrudes to the cytosolic space. According to the model and the experimental data, the transmembrane S6 and the contiguous C-terminus of each subunit converge in the cytoplasmic mouth. The most active TRPducin (represented in red surface) was derived from the linker connecting the S6 and proximal C-terminus [110]. The palmytoilation of the TRPducin allowed the peptide to reach the cytosolic space, interfering with tetramerization and normal activity of TRPV1 channel.

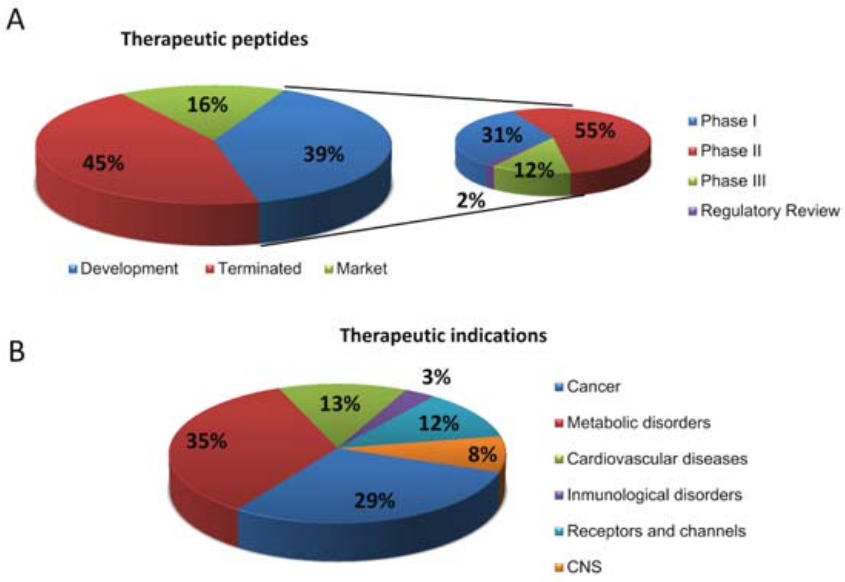

Figure 4. Development status of therapeutic peptides. (A) Distributions of therapeutic peptides by development status; the insert shows the distribution of development peptides by their state in clinical trials. (B) Distribution of development peptides by therapeutic indications. Data from peptide therapeutics foundation: Development trends for peptide foundation (2010), http://clinicaltrials.gov/ and http://www.fda.gov/.

limited due to the high mutation rate of HIV that leads to drug resistance. Shultz *et al*. (2004) targeted the four-stranded β-sheet dimerization interface for inhibition, since this region is relatively free of mutations [132]. The authors designed a focused library based on the interfacial peptides, and found an effective inhibition of the enzyme dimerization. This finding confirmed the oligomerization surface as an alternative region for targeting the active site.

Amyloid formation

Amyloid diseases such as Alzheimer's are associated with the transformation of normally soluble proteins into amyloid fibrils protein aggregates. Inhibitors of pathological amyloid fibril formation may be useful as therapeutics if they are sufficiently specific. Abe *et al*. (2007) screened α-synuclein-binding peptides by *in silico* panning to obtain an effective aggregation inhibitor. Using a genetic algorithm and a docking simulation, the study aimed at discovering peptides able to interact with the region involved in amyloid fibril formation. The selected peptides were able to bind to α-synuclein, thus affecting its aggregation [133]. Sievers *et al*. (2011) employed computer-aided, structure-based design, with known atomic structures of segments of amyloid fibres as templates, to evolve a D-amino-acid peptide that delayed amyloid fibril formation of the tau protein associated with Alzheimer's disease [134].

OUTLOOK

Protein-protein interaction provided by globular domains or short peptides represents the major complex network within the cell. It is estimated that peptide-protein interactions represent near to 15-40% of all interactions [135], which makes peptide and peptidomimetic discovery of great applicability and interest as therapeutic targets. Peptides derived from protein interfaces are readily good candidates for inhibitory purposes because although they are unstructured in solution, they adopt the correct conformation upon binding [25]. Much evidence has been found on the efficacy and selectivity of peptides and peptidomimetics to maximize binding interactions. So, there is a compromise between spatial organization of chemical groups, and enough flexibility to adapt to almost any kind of surface. Peptide drugs offer clear advantages for peptidic therapies: target selectivity, the possibility to act on big surfaces, high flexibility, reduced toxicity, and a large knowledge-based body of experimental and structural data.

There is a clear need to combine experimental methods (X-ray, NMR, EM, etc.) and computational predictions (docking, VS, bioinformatics, etc.) to expand the structural information of protein assemblies. The recent advances in peptide synthesis and the development of computational methods facilitate the rational design of peptidic compounds. The computational tools allow modeling peptide-protein interactions and contribute to the modulation of these interactions. Docking and virtual screening approaches allow searching of optimal interactions at wide scale. The developments in energy potentials improve determination of binding affinities between peptide and proteins. Machine learning approaches help the discrimination between binders and non-binders to build classification models based on interaction energy [136]. All advances in protein design, protein folding, docking and dynamics directly benefit peptide drug discovery [27]. The peptidomimetic field in turn takes advantage of the biochemical studies, synthesis advances and computational techniques [30].

In contrast to the artificially constructed random peptide libraries, many of the used peptides or awaiting approval ones are directly derived from natural sequences. One plausible explanation is that natural subdomains have been selected for stability by evolution [11]. Peptide optimization is also possible with the purpose of increasing stability, affinity and specificity. Either natural or modified peptides can in turn be readily subjected to wet-laboratory experiments and concomitant lead optimization.

Despite the limitations to the use of peptides as drug candidates, there is a considerable and probably unexpected number of peptides currently available as drugs or in clinical trials. This is because peptides and derivatives can be used in multiple pathologies such as arthritis, asthma, diabetes, allergy, infective diseases, inflammation, obesity, immune diseases, oncology, cardiovascular diseases, pain, osteoporosis, etc. [137]. One of the

challenges is the central nervous system, the major therapeutic area with great potential for peptides. Unfortunately, most peptides are unable to cross the blood brain barrier by passive transport because of the high molecular mass and hydrophilicity. In this respect, peptide-based vectors entering the barrier by absorptive-mediated [118] or receptor-mediated transport [138] can contribute to the development of therapeutic peptides targeted to the central nervous system [19].

Although peptides have been studied as drugs for decades, the rate of entry into clinical study was low prior to the 1980s. The average number of new candidates entering study per year has steadily increased; this number was 1.2 per year in the 1970s, 4.6 per year in the 1980s, 9.7 per year in the 1990s, and 16.8 per year so far in the 2000s. In 2003 the global therapeutic peptide market amounted to 1 billion USD and currently more than 54 peptide based products are commercially available with 6 in the registration process. In 2010, six therapeutics peptides have reached global sales over US$1 billion.

Looking forward, the market for peptide drugs will begin to show increased growth as drug candidates in Phase III and Phase II trials gain approval and enter the market (Figure 4). The peptide drug pipeline is robust, with the number of disease indications being investigated growing from 29 to 64. Disease areas with significant peptide therapeutic development activity include cancer, infection, and pain-all billion-dollar markets.

The use of peptides and peptidomimetics, instead of being an alternative to the use of small molecules, may represent an efficient complementary strategy in drug discovery targeted against protein-protein interaction networks.

ACKNOWLEDGMENTS

We are grateful to the members of our laboratory for their continuous support and collaboration. This work was supported by grants from the Ministry of Science and Innovation BFU2009-08346 to A. F. -M., BFU2008-0062 to J. M. G. R., CONSOLIDER-INGENIO 2010 CSD2008-00005 to A. F. -M. and J. M. G. -R. and la Generalitat Valenciana Prometeo/2010/046 to A. F. -M. We are very grateful to Pilar Aguado-Jiménez for the edition of the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Arkin, M. R. and Whitty, A. 2009, Curr. Opin. Chem. Biol., 13, 284.
2. Block, P., Weskamp, N., Wolf, A., and Klebe, G. 2007, Proteins, 68, 170.
3. Fletcher, S. and Hamilton, A. D. 2006, J. R. Soc. Interface, 3, 215.
4. Wells, J. A. and McClendon, C. L. 2007, Nature, 450, 1001.
5. Robin, W. S. 1998, Biotechnol. Bioeng., 61, 61.
6. Cochran, A. G. 2000, Chem. Biol., 7, R85.
7. Jones, S. and Thornton, J. M. 1996, Proc. Natl. Acad. Sci. USA, 93, 13.
8. Lo, C. L., Chothia, C., and Janin, J. 1999, J. Mol. Biol., 285, 2177.
9. Archakov, A. I., Govorun, V. M., Dubanov, A. V., Ivanov, Y. D., Veselovsky, A. V., Lewi, P., and Janssen, P. 2003, Proteomics, 3, 380.
10. Hopkins, A. L. and Groom, C. R. 2002, Nat. Rev. Drug Discov., 1, 727.
11. Watt, P. M. 2006, Nat. Biotechnol., 24, 177.
12. Walsh, G. 2010, Nat. Biotechnol., 28, 917.
13. Patel, L. N., Zaro, J. L., and Shen, W. C. 2007, Pharm. Res., 24, 1977.
14. Tan, M. L., Choong, P. F., and Dass, C. R. 2010, Peptides, 31, 184.
15. Timmerman, P., Beld, J., Puijk, W. C., and Meloen, R. H. 2005, Chembiochem., 6, 821.
16. Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeny, P. J. 1997, Adv. Drug Delivery Rev., 23, 3.
17. Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., and Kopple, K. D. 2002, J. Med. Chem., 45, 2615.
18. Benyamini, H. and Friedler, A. 2010, Future Med. Chem., 2, 989.
19. Vlieghe, P., Lisowski, V., Martinez, J., and Khrestchatisky, M. 2010, Drug Discov. Today, 15, 40.
20. Hummel, G., Reineke, U., and Reimer, U. 2006, Mol. Biosyst., 2, 499.

21. Loffet, A. 2002, J. Pept. Sci., 8, 1.
22. Pawson, T. and Nash, P. 2003, Science, 300, 445.
23. Craik, D. J., Clark, R. J., and Daly, N. L. 2007, Expert. Opin. Investig. Drugs, 16, 595.
24. Gebauer, M. and Skerra, A. 2009, Curr. Opin. Chem. Biol., 13, 245.
25. London, N., Raveh, B., Movshovitz-Attias, D., and Schueler-Furman, O. 2010, Proteins, 78, 3140.
26. London, N., Movshovitz-Attias, D., and Schueler-Furman, O. 2010, Structure, 18, 188.
27. Vanhee, P., van der Sloot, A. M., Verschueren, E., Serrano, L., Rousseau, F., and Schymkowitz, J. 2011, Trends Biotechnol., 29, 231.
28. Dyson, H. J. and Wright, P. E. 2005, Nat. Rev. Mol. Cell Biol., 6, 197.
29. Hwang, H., Pierce, B., Mintseris, J., Janin, J., and Weng, Z. 2008, Proteins, 73, 705.
30. Rubinstein, M. and Niv, M. Y. 2009, Biopol., 91, 505.
31. Aloy, P. and Russell, R. B. 2004, Nat. Biotechnol., 22, 1317.
32. Rose, P. W., Beran, B., Bi, C., Bluhm, W. F., Dimitropoulos, D., Goodsell, D. S., Prlic, A., Quesada, M., Quinn, G. B., Westbrook, J. D., Young, J., Yukich, B., Zardecki, C., Berman, H. M., and Bourne, P. E. 2011, Nucleic Acids Res., 39, D392.
33. Russell, R. B., Alber, F., Aloy, P., Davis, F. P., Korkin, D., Pichaud, M., Topf, M., and Sali, A. 2004, Curr. Opin. Struct. Biol., 14, 313.
34. Aloy, P., Pichaud, M., and Russell, R. B. 2005, Curr. Opin. Struct. Biol., 15, 15.
35. Stein, A., Mosca, R., and Aloy, P. 2011, Curr. Opin. Struct. Biol., 21, 200.
36. Stein, A., Russell, R. B., and Aloy, P. 2005, Nucleic Acids Res., 33, D413.
37. Vanhee, P., Reumers, J., Stricher, F., Baeten, L., Serrano, L., Schymkowitz, J., and Rousseau, F. 2010, Nucleic Acids Res., 38, D545.
38. Vanhee, P., Verschueren, E., Baeten, L., Stricher, F., Serrano, L., Rousseau, F., and Schymkowitz, J. 2011, Nucleic Acids Res., 39, D435.
39. Levy, E. D., Pereira-Leal, J. B., Chothia, C., and Teichmann, S. A. 2006, PLoS. Comput. Biol., 2, e155.
40. Vanhee, P., Stricher, F., Baeten, L., Verschueren, E., Lenaerts, T., Serrano, L., Rousseau, F., and Schymkowitz, J. 2009, Structure, 17, 1128.
41. Peczuh, M. W. and Hamilton, A. D. 2000, Chem. Rev., 100, 2479.
42. Sebolt-Leopold, J. S. and English, J. M. 2006, Nature, 441, 457.
43. Karaman, M. W., Herrgard, S., Treiber, D. K., Gallant, P., Atteridge, C. E., Campbell, B. T., Chan, K. W., Ciceri, P., Davis, M. I., Edeen, P. T., Faraoni, R., Floyd, M., Hunt, J. P., Lockhart, D. J., Milanov, Z. V., Morrison, M. J., Pallares, G., Patel, H. K., Pritchard, S., Wodicka, L. M., and Zarrinkar, P. P. 2008, Nat. Biotechnol., 26, 127.
44. Bogoyevitch, M. A. and Fairlie, D. P. 2007, Drug Discov. Today, 12, 622.
45. Ye, G., Tiwari, R., and Parang, K. 2008, Curr. Opin. Investig. Drugs, 9, 605.
46. Eldar-Finkelman, H. and Eisenstein, M. 2009, Curr. Pharm. Des., 15, 2463.
47. Tal-Gan, Y., Hurevich, M., Klein, S., Ben-Shimon, A., Rosenthal, D., Hazan, C., Shalev, D. E., Niv, M. Y., Levitzki, A., and Gilon, C. 2011, J. Med. Chem., 54, 5154.
48. Fujiwara, D., Ye, Z., Gouda, M., Yokota, K., Tsumuraya, T., and Fujii, I. 2010, Bioorg. Med. Chem. Lett., 20, 1776.
49. Szymkowski, D. E. 2005, Curr. Opin. Drug Discov. Devel, 8, 590.
50. Villanueva, J., Fernandez-Ballester, G., Querol, E., Aviles, F. X., and Serrano, L. 2003, J. Mol. Biol., 330, 1039.
51. Fernandez-Ballester, G., Beltrao, P., Gonzalez, J. M., Song, Y. H., Wilmanns, M., Valencia, A., and Serrano, L. 2009, J. Mol. Biol., 388, 902.
52. Reina, J., Lacroix, E., Hobson, S. D., Fernandez-Ballester, G., Rybin, V., Schwab, M. S., Serrano, L., and Gonzalez, C. 2002, Nat. Struct. Biol., 9, 621.
53. Hecht, I., Rong, J., Sampaio, A. L., Hermesh, C., Rutledge, C., Shemesh, R., Toporik, A., Beiman, M., Dassa, L., Niv, H., Cojocaru, G., Zauberman, A., Rotman, G., Perretti, M., Vinten-Johansen, J., and Cohen, Y. 2009, J. Pharmacol. Exp. Ther., 328, 426.

54. Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. 2005, Nucleic Acids Res., 33, W382.
55. Guerois, R., Nielsen, J. E., and Serrano, L. 2002, J. Mol. Biol., 320, 369.
56. Encinar, J. A., Fernandez-Ballester, G., Sanchez, I. E., Hurtado-Gomez, E., Stricher, F., Beltrao, P., and Serrano, L. 2009, Bioinformatics, 25, 2418.
57. Smith, C. A. and Kortemme, T. 2010, J. Mol. Biol., 402, 460.
58. Raveh, B., London, N., and Schueler-Furman, O. 2010, Proteins, 78, 2029.
59. King, C. A. and Bradley, P. 2010, Proteins, 78, 3437.
60. Ghersi, D. and Sanchez, R. 2011, J. Struct. Funct. Genomics, 12, 109.
61. Hajduk, P. J., Huth, J. R., and Fesik, S. W. 2005, J. Med. Chem., 48, 2518.
62. Cardozo, T., Totrov, M., and Abagyan, R. 1995, Proteins, 23, 403.
63. Ravna, A. W., Sylte, I., and Dahl, S. G. 2009, J. Mol. Model, 15, 1155.
64. Adamian, L., Gussin, H. A., Tseng, Y. Y., Muni, N. J., Feng, F., Qian, H., Pepperberg, D. R., and Liang, J. 2009, Protein Sci., 18, 2371.
65. Unal, E. B., Gursoy, A., and Erman, B. 2010, PLoS. One, 5, e10926.
66. Tseng, Y. Y. and Li, W. H. 2011, Proc. Natl. Acad. Sci. USA, 108, 5313.
67. Sousa, S. F., Fernandes, P. A., and Ramos, M. J. 2006, Proteins, 65, 15.
68. Perola, E., Walters, W. P., and Charifson, P. S. 2004, Proteins, 56, 235.
69. Bursulaya, B. D., Totrov, M., Abagyan, R., and Brooks, C. L., III 2003, J. Comput. Aided Mol. Des., 17, 755.
70. Dominguez, C., Boelens, R., and Bonvin, A. M. 2003, J. Am. Chem. Soc., 125, 1731.
71. Kontoyianni, M., McClellan, L. M., and Sokol, G. S. 2004, J. Med Chem., 47, 558.
72. Arun, P. P. and Gautham, N. 2008, J. Comput. Aided Mol. Des., 22, 815.
73. Kurcinski, M. and Kolinski, A. 2007, J. Steroid Biochem. Mol. Biol., 103, 357.
74. Stein, A. and Aloy, P. 2008, PLoS. One, 3, e2524.
75. Niv, M. Y. and Weinstein, H. 2005, J. Am. Chem. Soc., 127, 14072.
76. Bordner, A. J. and Abagyan, R. 2006, Proteins, 63, 512.
77. Maurer, M. C., Trosset, J. Y., Lester, C. C., DiBella, E. E., and Scheraga, H. A. 1999, Proteins, 34, 29.
78. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. 1998, Journal of Computational Chemistry, 19, 1639.
79. Hussain, A., Shaw, P. E., and Hirst, J. D. 2011, J. Cheminform, 3, 49.
80. Pang, X., Zhang, M., Zhou, L., Xie, F., Lu, H., He, W., Jiang, S., Yu, L., and Zhang, X. 2011, Eur. J. Med. Chem., 46, 1701.
81. Merrifield, R. B. 1969, Adv. Enzymol. Relat Areas Mol. Biol., 32, 221.
82. Torbeev, V. Y. and Kent, S. B. 2007, Angew. Chem. Int. Ed Engl., 46, 1667.
83. Grauer, A. and Konig, B. 2009, European Journal of Organic Chemistry, 2009, 5099.
84. Vagner, J., Qu, H., and Hruby, V. J. 2008, Curr. Opin. Chem Biol., 12, 292.
85. Zinzalla, G. and Thurston, D. E. 2009, Future. Med. Chem., 1, 65.
86. Dohm, M. T., Kapoor, R., and Barron, A. E. 2011, Curr. Pharm. Des., in press.
87. Garcia-Martinez, C., Humet, M., Planells-Cases, R., Gomis, A., Caprini, M., Viana, F., De La, P. E., Sanchez-Baeza, F., Carbonell, T., De, F. C., Perez-Paya, E., Belmonte, C., Messeguer, A., and Ferrer-Montiel, A. 2002, Proc. Natl. Acad. Sci. USA, 99, 2374.
88. Montoliu, C., Humet, M., Canales, J. J., Burda, J., Planells-Cases, R., Sanchez-Baeza, F., Carbonell, T., Perez-Paya, E., Messeguer, A., Ferrer-Montiel, A., and Felipo, V. 2002, J. Pharmacol. Exp. Ther., 301, 29.
89. Planells-Cases, R., Montoliu, C., Humet, M., Fernandez, A. M., Garcia-Martinez, C., Valera, E., Merino, J. M., Perez-Paya, E., Messeguer, A., Felipo, V., and Ferrer-Montiel, A. 2002, J. Pharmacol. Exp. Ther., 302, 163.
90. Humet, M., Carbonell, T., Masip, I., Sanchez-Baeza, F., Mora, P., Canton, E., Gobernado, M., Abad, C., Perez-Paya, E., and Messeguer, A. 2003, J. Comb. Chem., 5, 597.

91. Masip, I., Perez-Paya, E., and Messeguer, A. 2005, Comb. Chem. High Throughput. Screen, 8, 235.
92. Davies, J. S. 2003, J. Pept. Sci., 9, 471.
93. Witt, K. A., Gillespie, T. J., Huber, J. D., Egleton, R. D., and Davis, T. P. 2001, Peptides, 22, 2329.
94. Zhou, H. X. 2004, Acc. Chem. Res., 37, 123.
95. Driggers, E. M., Hale, S. P., Lee, J., and Terrett, N. K. 2008, Nat. Rev. Drug Discov., 7, 608.
96. Liu, J., Wang, D., Zheng, Q., Lu, M., and Arora, P. S. 2008, J. Am. Chem. Soc., 130, 4334.
97. Schafmeister, C. E., Po, J., and Verdine, G. L. 2000, Journal of the American Chemical Society, 122, 5891.
98. Gu, X., Ying, J., Agnes, R. S., Navratilova, E., Davis, P., Stahl, G., Porreca, F., Yamamura, H. I., and Hruby, V. J. 2004, Org. Lett., 6, 3285.
99. Frank, R. and Overwin, H. 1996, Methods Mol. Biol., 66, 149.
100. Seidel, T., Ibis, G., Bendix, F., and Wolber, G. 2010, Drug Discov. Today, 7, e221.
101. Kiran, M. D., Adikesavan, N. V., Cirioni, O., Giacometti, A., Silvestri, C., Scalise, G., Ghiselli, R., Saba, V., Orlando, F., Shoham, M., and Balaban, N. 2008, Mol. Pharmacol., 73, 1578.
102. Parthasarathi, L., Casey, F., Stein, A., Aloy, P., and Shields, D. C. 2008, J. Chem Inf. Model, 48, 1943.
103. Stahura, F. L. and Bajorath, J. 2005, Curr. Pharm. Des., 11, 1189.
104. Eckert, H. and Bajorath, J. 2007, J. Chem. Inf. Model, 47, 1366.
105. Qvit, N., Reuveni, H., Gazal, S., Zundelevich, A., Blum, G., Niv, M. Y., Feldstein, A., Meushar, S., Shalev, D. E., Friedler, A., and Gilon, C. 2008, J. Comb. Chem., 10, 256.
106. Villoutreix, B. O., Renault, N., Lagorce, D., Sperandio, O., Montes, M., and Miteva, M. A. 2007, Curr. Protein Pept. Sci., 8, 381.
107. Covic, L., Gresser, A. L., Talavera, J., Swift, S., and Kuliopulos, A. 2002, Proc. Natl. Acad. Sci. USA, 99, 643.
108. Sevigny, L. M., Zhang, P., Bohm, A., Lazarides, K., Perides, G., Covic, L., and Kuliopulos, A. 2011, Proc. Natl. Acad. Sci. USA, 108, 8491.
109. Cisowski, J., O'Callaghan, K., Kuliopulos, A., Yang, J., Nguyen, N., Deng, Q., Yang, E., Fogel, M., Tressel, S., Foley, C., Agarwal, A., Hunt, S. W., III, McMurry, T., Brinckerhoff, L., and Covic, L. 2011, Am. J. Pathol., 179, 513.
110. Valente, P., Fernandez-Carvajal, A., Camprubi-Robles, M., Gomis, A., Quirce, S., Viana, F., Fernandez-Ballester, G., Gonzalez-Ros, J. M., Belmonte, C., Planells-Cases, R., and Ferrer-Montiel, A. 2011, FASEB J., 25, 1628.
111. Yin, H., Slusky, J. S., Berger, B. W., Walters, R. S., Vilaire, G., Litvinov, R. I., Lear, J. D., Caputo, G. A., Bennett, J. S., and DeGrado, W. F. 2007, Science, 315, 1817.
112. Blanes-Mira, C., Pastor, M. T., Valera, E., Fernandez-Ballester, G., Merino, J. M., Gutierrez, L. M., Perez-Paya, E., and Ferrer-Montiel, A. 2003, Biochem. J., 375, 159.
113. Blanes-Mira, C., Merino, J. M., Valera, E., Fernandez-Ballester, G., Gutierrez, L. M., Viniegra, S., Perez-Paya, E., and Ferrer-Montiel, A. 2004, J. Neurochem., 88, 124.
114. Camprubi-Robles, M., Planells-Cases, R., and Ferrer-Montiel, A. 2009, FASEB J., 23, 3722.
115. Murray, J. K. and Gellman, S. H. 2007, Biopol., 88, 657.
116. Wasylyk, C., Salvi, R., Argentini, M., Dureuil, C., Delumeau, I., Abecassis, J., Debussche, L., and Wasylyk, B. 1999, Oncogene, 18, 1921.
117. Bottger, A., Bottger, V., Garcia-Echeverria, C., Chene, P., Hochkeppel, H. K., Sampson, W., Ang, K., Howard, S. F., Picksley, S. M., and Lane, D. P. 1997, J. Mol. Biol., 269, 744.
118. Garcia-Echeverria, C., Chene, P., Blommers, M. J., and Furet, P. 2000, J. Med. Chem., 43, 3205.
119. Zhong, H. and Carlson, H. A. 2005, Proteins, 58, 222.

120. Phan, J., Li, Z., Kasprzak, A., Li, B., Sebti, S., Guida, W., Schonbrunn, E., and Chen, J. 2010, J. Biol. Chem., 285, 2174.
121. Hu, Y., Li, X., Sebti, S. M., Chen, J., and Cai, J. 2011, Bioorg. Med. Chem. Lett., 21, 1469.
122. Laudet, B., Barette, C., Dulery, V., Renaudet, O., Dumy, P., Metz, A., Prudent, R., Deshiere, A., Dideberg, O., Filhol, O., and Cochet, C. 2007, Biochem. J., 408, 363.
123. Sudarman, E., Bollati-Fogolin, M., Hafner, M., Muller, W., Scheller, J., Rose-John, S., and Eichler, J. 2008, Chem. Biol. Drug Des., 71, 494.
124. Zobel, K., Wang, L., Varfolomeev, E., Franklin, M. C., Elliott, L. O., Wallweber, H. J., Okawa, D. C., Flygare, J. A., Vucic, D., Fairbrother, W. J., and Deshayes, K. 2006, ACS Chem. Biol., 1, 525.
125. Sun, H., Nikolovska-Coleska, Z., Lu, J., Meagher, J. L., Yang, C. Y., Qiu, S., Tomita, Y., Ueda, Y., Jiang, S., Krajewski, K., Roller, P. P., Stuckey, J. A., and Wang, S. 2007, J. Am. Chem. Soc., 129, 15279.
126. Walensky, L. D., Kung, A. L., Escher, I., Malia, T. J., Barbuto, S., Wright, R. D., Wagner, G., Verdine, G. L., and Korsmeyer, S. J. 2004, Science, 305, 1466.
127. Canela, N., Orzaez, M., Fucho, R., Mateo, F., Gutierrez, R., Pineda-Lucena, A., Bachs, O., and Perez-Paya, E. 2006, J. Biol. Chem., 281, 35942.
128. Li, H. Y., Zawahir, Z., Song, L. D., Long, Y. Q., and Neamati, N. 2006, J. Med. Chem., 49, 4477.
129. Hayouka, Z., Rosenbluh, J., Levin, A., Loya, S., Lebendiker, M., Veprintsev, D., Kotler, M., Hizi, A., Loyter, A., and Friedler, A. 2007, Proc. Natl. Acad. Sci. USA, 104, 8316.
130. Zawahir, Z. and Neamati, N. 2006, Bioorg. Med. Chem. Lett., 16, 5199.
131. Ren, S. and Lien, E. J. 2001, Prog. Drug Res., Spec. No, 1.
132. Shultz, M. D., Ham, Y. W., Lee, S. G., Davis, D. A., Brown, C., and Chmielewski, J. 2004, J. Am. Chem. Soc., 126, 9886.
133. Abe, K., Kobayashi, N., Sode, K., and Ikebukuro, K. 2007, BMC. Bioinformatics, 8, 451.
134. Sievers, S. A., Karanicolas, J., Chang, H. W., Zhao, A., Jiang, L., Zirafi, O., Stevens, J. T., Munch, J., Baker, D., and Eisenberg, D. 2011, Nature, 475, 96.
135. Petsalaki, E. and Russell, R. B. 2008, Curr. Opin. Biotechnol., 19, 344.
136. Kinnings, S. L., Liu, N., Tonge, P. J., Jackson, R. M., Xie, L., and Bourne, P. E. 2011, J. Chem. Inf. Model, 51, 408.
137. Stevenson, C. L. 2009, Curr. Pharm. Biotechnol., 10, 122.
138. de Boer, A. G. and Gaillard, P. J. 2007, Clin. Pharmacokinet., 46, 553.
139. Shameer, K., Madan, L. L., Veeranna, S., Gopal, B., and Sowdhamini, R. 2010, BMC. Bioinformatics, 11, 473.
140. Jacob, R. B., Bullock, C. W., Andersen, T., and McDougal, O. M. 2011, J. Comput. Chem., 32, 2936.
141. Duffy, F. J., Verniere, M., Devocelle, M., Bernard, E., Shields, D. C., and Chubb, A. J. 2011, J. Chem. Inf. Model, 51, 829.
142. Goede, A., Michalsky, E., Schmidt, U., and Preissner, R. 2006, BMC. Bioinformatics, 7, 11.
143. Casey, F. P., Davey, N. E., Baran, I., Varekova, R. S., and Shields, D. C. 2008, J. Chem. Inf. Model, 48, 1524.
144. Floris, M., Masciocchi, J., Fanton, M., and Moro, S. 2011, Nucleic Acids Res., 39, W261.
145. Kolter, T. and Wendeler, M. 2003, Chembiochem., 4, 260.
146. Fernandez-Ballester, G. and Ferrer-Montiel, A. 2008, J. Membr. Biol., 223, 161.