University of Szeged

Faculty of Pharmacy

Department of Pharmaceutical Analysis

DESIGN STRATEGIES FOR PROTEIN-PROTEIN INTERACTION INHIBITORS USING NON-NATURAL AMINO ACIDS

Ph. D. Thesis

Attila Tököli

Supervisors:

Prof. Dr. Tamás Martinek, Dr. Gerda Szakonyi

2023

Table of contents

List of publications and lectures

Full papers related to the thesis

I. A. Tököli, B. Mag, É. Bartus, E. Wéber, G. Szakonyi, M. A. Simon, Á. Czibula, É. Monostori, L. Nyitray, T. A. Martinek (2020). Proteomimetis surface fragments distinguish proteins by function. *Chemical Science*, 11, 10390-10398

II. A. Tököli, B. Bodnár, F. Bogár, G. Paragi, A. Hetényi, É. Bartus, E. Wéber, Z. Hegedüs, Z. Szabó, G. Kecskeméti, G. Szakonyi, T. A. Martinek (2023). Structural adaptation of the single-stranded DNA-binding protein C-terminal to DNA metabolizing partners guides inhibitor design. *Pharmaceutics*, 15(4), 1032

Other full papers

I. A. Hetenyi, E. Szabo, N. Imre, K. N. Bhaumik, A. Tököli, T. Füzesi, R. Hollandi, P. Horvath, Á. Czibula, É. Monostori, M. A. Deli, T. A. Martinek (2022). α/β-Peptides as Nanomolar Triggers of Lipid Raft-Mediated Endocytosis through GM1 Ganglioside Recognition. *Pharmaceutics*, *14*(3), 580

II. É. Bartus, A. Tököli, B. Mag, Á. Bajcsi, G. Kecskeméti, E. Wéber, Z. Kele, G. Fenteany, T. A Martinek (2022). Light-fuelled primitive replication and selection in evolvable biomimetic chemical networks. *ChemRxiv*, under major revision, 10.26434/chemrxiv-2021-3dnjt-v4

Scientific lectures related to the thesis

1. A. Tököli, G. Szakonyi, T. Martinek. A bakteriális RecQ helikáz szárnyas doménjének karakterizálása, egy lehetséges antibakteriális célpont.
   MTA Peptidkémiai Munkabizottsági Ülés 2017.
   Balatonszemes – 2017. 05. 29.
2. A. Tököli, É. Bartus, G. Szakonyi, T. Martinek. Characterization of the bacterial winged-helix domain of RecQ helicase. A novel antibacterial target.
   7th BBBB International Conference on Pharmaceutical Sciences –
   Balatonfüred – 2017. 10. 06.
3. A. Tököli, B. Mag, É. Bartus, E. Wéber, G. Szakonyi, M. A. Simon, Á. Czibula, É. Monostori, L. Nyitray, Tamás A. Martinek. Protein affinity patterns of foldameric local surface mimetics: druggability and promiscuity.
   Department of Medicinal Chemistry, Institute Seminar
   Szeged – 2020.11.27.
4. A. Tököli, B. Bodnár, F. Bogár[2], G. Paragi, A. Hetényi, É. Kovács-Bartus, E. Wéber, Z. Hegedüs, G. Szakonyi and T. A. Martinek. Two anchor point-binding of the SSB C-terminal to DNA metabolizing proteins facilitates the development of enhanced inhibitors. Peptide Chemistry and Chemical Biology Working Committees of the Hungarian Academy of Sciences – Symposium 2022.
   Balatonszemes – 2022. 05. 30.

Abbreviations

ACHC: (1S,2S)-2-aminocyclohexanecarboxylic acid

AUC: area under curve

DBU: 1,8-diazabicycloundec-7-ene

DCL: dynamic combinatorial library

DCM: dichloromethane

DIPEA: N,N-diisopropylethylamine

DMF: N,N-dimethylformamide

EC50: half maximal effective concentration

ESI-MS: electrospray ionization mass spectrometry

FB: fraction bound

Fmoc: fluorenylmethyloxycarbonyl

GSH: reduced glutathione

GSSG: oxidized glutathione

HATU: 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]-pyridinium-3-oxid hexafluorophosphate

HBS: hydrogen bond surrogate

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HMW: high molecular weight

HTS: high-throughput screening

IC50: half minimal inhibitory concentration

IDP: intrinsically disordered protein

IDL: intrinsically disordered linker

ITC: isothermal titration calorimetry

MDM2: mouse double minute 2

PPI: protein-protein interaction

SIP: SSB-interacting proteins

SSB: single-stranded DNA-binding protein

SPPS: solid-phase peptide synthesis

TFA: trifluoroacetic acid

TIS: triisopropylsilane

1. Introduction

Proteins are the key elements of life. They govern various cellular functions, serve as molecular skeletons and catalyze chemical reactions. A complicated network of protein-protein interactions is involved in maintaining homeostasis in every living being. Proteins interact with each other through complex interfaces varying in area, size, and chemical characteristics. These interfaces lack deep binding pockets, which could be targeted with small molecules, thus considered „undruggable" or „hard-to-drug" using classical pharmaceutical chemistry approaches. Large molecules, e.g., peptides, peptidomimetics, and proteins are structurally better suited for this purpose.

Skolnick et al. showed that protein-protein interface side chain arrangements could be highly similar in two different proteins with distinct global structural features.[1] Any PPI interface could be sliced up, resulting in surface segments. This phenomenon raises the question: can we design molecules, which are a) short and have secondary structures, b) behave like a segment of an interface exhibiting enough surface area for binding, and c) are able to project side chains in a biomimetic manner? Is it possible to create surface fragments? In principle, a library of these small surface fragments containing a wide variety of amino acid side chains could be reassembled in a bottom-up scenario to yield peptidomimetic molecules with contact surface area sufficient for high affinity and specific binding. This ligand assembly can be carried out using dynamic combinatorial chemistry in the presence of a target protein.[2]

An interesting phenomenon in structural biology is the interaction of proteins where one or both interacting partners are disordered; they do not adopt a well-defined structure under native conditions. In some cases, an IDP folds into a binding conformation when forming contact with its partner. Other disordered interfaces can only be described using multiple conformational states, hence the name fuzzy complex. The latter is an incredibly challenging interface for drug development; the binding surface on the receptor is shallow and diffuse, and the fuzzy partner is usually highly hydrophilic, lacking well-defined hot-spot residues. The amino acid composition in intrinsically disordered regions contains mostly hydrophilic residues, rendering the top-down design of IDP mimetics rather difficult. Extending the available side chain space by incorporating non-canonical amino acids opens up many possibilities in IDP mimetics drug design.

2. Aims

One of our goals was to prove that undruggable protein-protein interfaces can be targeted with a fragment-based approach using surface fragments. Short foldameric helices exhibit a large enough surface area to interact with surface features of other foldameric helices or protein hot spots. Helices built purely from β-amino acids have constrained backbones, stabilizing helical conformation. We intended to demonstrate that these helices are compatible with protein surfaces, can act as local surface mimetics, and serve as surface fragments in a bottom-up design approach. Lastly, we have planned to analyze the interaction of an intrinsically disordered segment, SSB-Ct, with its interacting partners, ExoI and RecO. Our objective was to investigate if the low binding affinity of a conserved and disordered peptide can be improved using non-natural amino acids. We intended to stabilize the fuzzy complex in one binding conformation by introducing modifications that favor enthalpically driven binding on the expanse of conformational entropy.

3. Literature background

3.1. Protein-protein interactions and interfaces

Protein-protein interactions are physical interactions between proteins having highly specific binding through molecular recognition. The recognition event is determined by the free energy of binding, affected by the size, surface complementarity, intermolecular interactions, local secondary structure, and structural changes upon complex formation.[3] PPIs can be categorized in several ways.

A broad classification is based on binding affinity and stability; interacting proteins having high K_D values (micromolar to millimolar) form transient, very short-lived interactions.[4] On the other hand, tight interactions might have under picomolar affinity forming complexes with high kinetic stability that are stable up to days.

Another aspect of classification is the composition of the protein assembly.[5] There are heteromeric and oligomeric complexes formed by oligomerization of the same protein or other different proteins, respectively. It has been noted previously that in *E. coli,* ca. 75% of the proteins are oligomeric.[6] Many hypotheses were stated on why proteins oligomerize. Larger complexes are more stable and have reduced solvent-accessible surface area compared to individual proteins.[6] Complex formation also gives rise to new functions.[7]

Several types of structural changes are associated with forming protein-protein complexes. When there are negligible structural changes upon complex formation, the interface can be described - similarly to protein-small molecule binding – as the lock, and key phenomenon.[8] Conformational selection mechanism hypothesizes that there is one preferred binding conformation out of the multiple conformational ensembles.[9]

Another molecular recognition mechanism is the induced fit model, where conformational changes are induced on the binding partners upon complex formation (e.g., IDP p300-HIF1α interaction).[10]

Protein-protein interaction interfaces have unique amino acid composition compared to the protein core or the non-binding surfaces of the protein.[3] Yan et al. showed in an extensive computational analysis that amino acids with hydrophilic side chains are more common in solvent-accessible areas (Figure 1). At the same time, hydrophobic residues are more frequent in buried parts of the protein. In PPI interfaces, hydrophobic residues (especially aromatic residues) are more frequent than hydrophilic residues; furthermore, Cys-Cys contacts and salt bridges are also abundant, providing selectivity. Protein-protein complex formation in water is entropy-driven[11]; burying hydrophobic

surface patches yields a significant entropy gain and stabilizes the resulting complex. It has also been shown that interface residues are more conserved than non-interface residues.[3]

Protein complexes are formed through an „encounter complex" transition state governed by electrostatic interactions (on the order of weak micromolar interaction). The interfaces are still mostly solvated but are aligned for docking. The final PPI complex is formed through desolvation and the formation of short-range interactions.[12,13]

Most of the binding energy is attributed to a number of residues called „hot-spot" residues. They are residues that, upon mutation to alanine, lose more than 2 kcal mol^{-1} in the binding energy.[14] Another definition states that hot spots are small regions that comprise the subset of residues that contribute the bulk of the binding free energy.[15] Regardless of the definition, hotspots are key areas of PPIs and are the focus of PPI modulation and inhibition since PPIs lack pockets into which small molecules could fit.

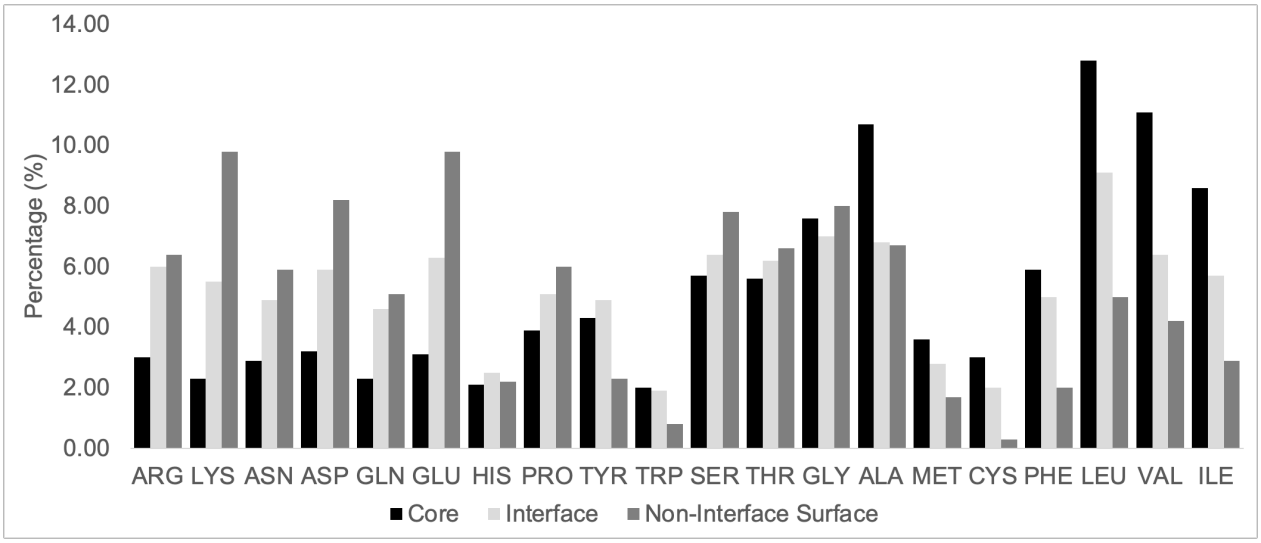

Figure 1. Amino acid composition is shown as the percentage compared to all amino acids in the protein core, interface, and non-interface surfaces, according to Yan et at.[3]

Most of the binding energy is attributed to a number of residues called „hot-spot" residues. They are residues that, upon mutation to alanine, lose more than 2 kcal mol^{-1} in the binding energy.[14] Another definition states that hot spots are small regions that comprise the subset of residues that contribute the bulk of the binding free energy.[15] Regardless of the definition, hotspots are key areas of PPIs and are the focus of PPI modulation and inhibition since PPIs lack pockets into which small molecules could fit.

The amino acid composition of disordered protein sequences differs vastly from the previously presented interfaces. In intrinsically disordered segments, bulky, hydrophobic amino acids are

underrepresented (Ile, Leu, Met, Val, Phe, Tyr, Trp), while there is a high proportion of polar amino acids (Asp, Asn, Glu, Gln, Ser, Lys) and proline. Proline is significantly overrepresented in long disordered sequences (e.g., polyproline helices). Another important descriptor of IPD segments is low sequence complexity, making IDP interfaces carry less information than helical or sheet interfaces.[16]

Different databases help to contain the substantially increasing PPI information publicly available. These powerful tools were assembled by individual scientific publications, enlisting thousands of physical protein-protein interactions across hundreds of different organisms.[17] Data gathered lines up a wide variety of experimental methods such as yeast two-hybrid systems[18] and tandem affinity purification coupled with mass spectrometry[19] as high throughput methods or experimental binding data, representing different levels of PPI evidence. The most extensive databases are the Biological General Repository for Interaction Datasets (BioGRID), the Molecular INTeraction database (MINT), the IntAct molecular interaction database (IntAct), and the Human Protein Reference Database (HPRD). There is a slight difference among databases in reported PPIs. Overlap among databases reaches up to 75 percent.[17] Agile Protein Interaction database (APID) is a colossal meta-database combining datasets from all other databases eliminating shortcomings in standalone databases.

3.2. Targeting protein hot-spots

Numerous methods have been developed to screen PPI inhibitor molecules, such as phenotypic, target-based, and structure-based screening.[20] Phenotypic screening, also known as forward chemical genetics, usually involves screening molecule libraries on a complex biological system to achieve a biological response generated by the modulation of a PPI.[21] Lead molecules can be further optimized. Subsequently, structure-activity relationships can be discovered after identifying the target and/or mechanism of action. Lenalidomide, an anticancer drug, and monastol, an inhibitor of mitotic spindle formation, were found using forward chemical genetics.[22] Target-based screening or reverse chemical genetics also utilizes screening of compound libraries, but the target is known as a part of a biologically validated pathway. Structural information is not necessary for this type of screening. Once a hit is identified, it can be further tested on a complex biological system for a change in phenotype. Several small molecular inhibitors of p53/Mdm2 interaction were discovered this way.[23] Methods involve microarrays, RNA interference, of small-molecule affinity chromatography.[24,25] Chemical genetics has drawbacks; off-target hit-related phenotypic changes may occur, making target identification difficult. The information generated by „omics”

fields facilitates identifying relevant complexes and aids target selection. It helps identify if the protein of interest has any relevant PPIs or is part of any disease pathway.

Structure-based design requires high-resolution structural information on the target of interest. Detailed information required is gathered from X-ray, NMR, or cryoEM studies.[26] For example, the design of venetoclax, a Bcl2 targeting drug used in chronic lymphocytic leukemia, was aided by NMR.[27] An important step in structure-based design is determining binding sites and hot spots involved in the targeted biological process. Alanine scanning is often used to determine amino acids responsible for the majority of binding energy in a protein-ligand or protein-protein interaction. Alanine scanning can be done *in vitro* by mutating putative residues in the target protein or *in silico.* The loss of binding free energy or the reduction of solvent-accessible surface area (ΔSASA) as a result of mutation highlights residues that contribute most to the free energy of complex formation between two proteins.[9,28] Underutilized binding surfaces or residues can also be found using these techniques. An algorithm named AlphaSpace aids the detection of underutilized spaces in the binding surfaces of protein complexes.[29].[30] Computational methods also facilitate the determination of pocket-like structures or shallow binding clefts.[31] When lacking a high-resolution structure, homology modeling may also be used to gain an atomistic model of a target protein. Homology models may be built from homologous proteins with similar biochemical functions and at least 30% sequential homology.[32,33]

3.3. Non-natural amino acids

Protein mimetics with non-natural residues diversify the toolbox available for drug discovery. Monomers such as D-amino acids, β amino acids, or residues with non-natural side chains are all widely used today in designing PPI inhibitors. D-amino acid scanning provides insights into the importance of side chain stereochemistry and provides proteolytic stability, often included as N-terminal caps. Backbone homologation strategies utilizing exclusively β- and γ- amino acids enable tighter control over the peptide's secondary structure, allowing the mimicry of secondary structural elements. α or N-methyl amino acids have their protons on either C_α or amino group substituted with a methyl group. Amino acids with non-natural side chains may uncover underutilized spaces on the binding interface or change the physicochemical characteristics of the peptidomimetic in a desired way. Including any non-canonical amino acid into a peptide sequence increases its proteolytic stability, may enhance affinity and selectivity, and may improve oral absorption and tissue distribution.

3.3.1. β-amino acids

β-amino acids are analogs of natural α amino acids, but the carbon skeleton has been lengthened by one methylene group (C_β) next to the carboxylic acid moiety. There are several substituted β-amino acids; β^2-, β^3-, and $\beta^{2,3}$-amino acids and cyclic β-amino acids (Figure 2). Cyclic β-amino acids are special $\beta^{2,3}$-amino acids containing cycloalkane or heterocyclic rings as side chains. Each substituted β-amino acid requires a different synthetic route to the procedure, many of which are commercially available.[34]

3.3.2. Non-natural amino acids in medicinal chemistry

An increasing number of amino acid, peptide, and peptoid drugs, among other biologically active molecules are reaching clinical trials or the market.[35,36] The drugs and biologically active peptides presented in the following section are an incomplete set of compounds. They are used to showcase unusual amino acids relevant to this thesis; a detailed presentation of all compounds in this area is out of scope.

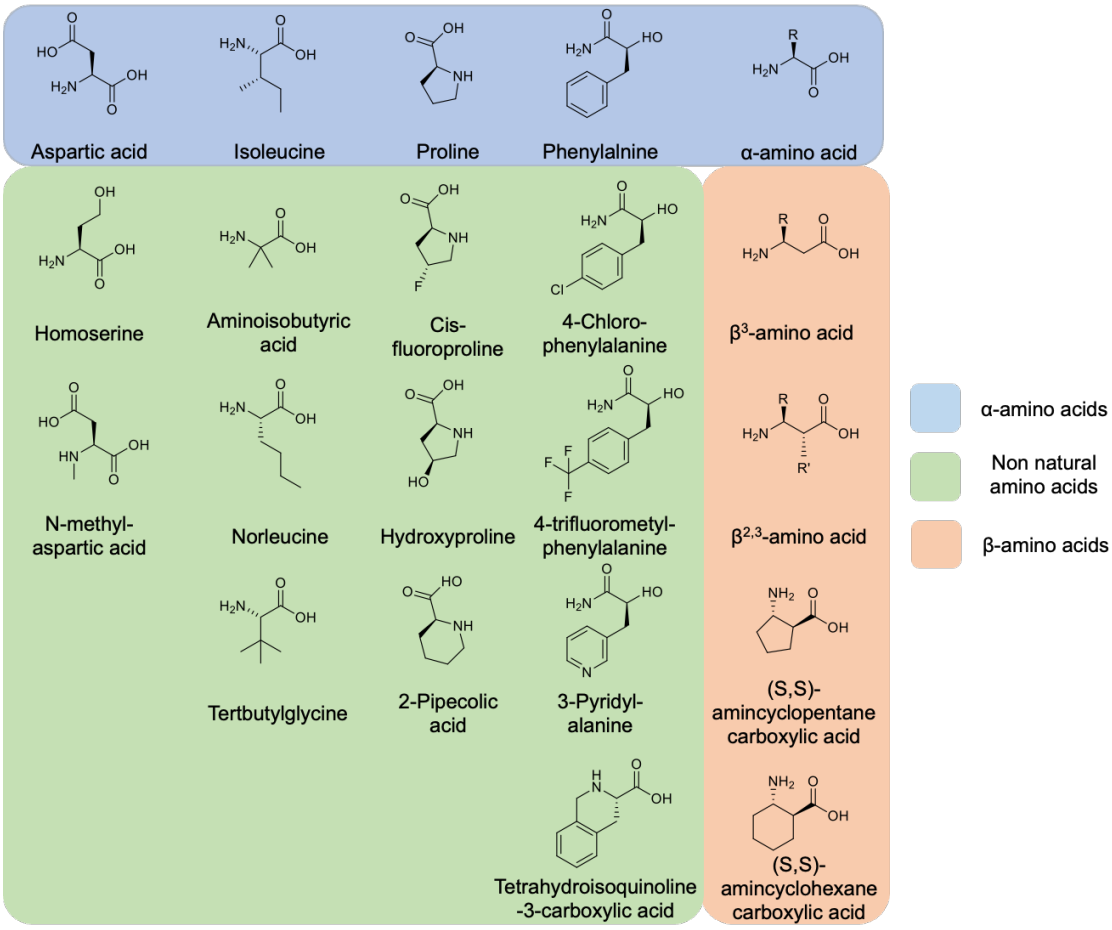

Figure 2. Non-natural amino acids. Showing an α-amino acid, α-amino acids with non-natural side chains, and β-amino acids.

Table 1. A table showing non-natural amino acids and their prevalence in marketed drugs or drug development.

Analog of	Non-natural amino-acid	Compound name	Indication or use
L-aspartic acid and L-glutamic acid derivatives	Several analogs		Modulators of AMPA and NMDA receptors used in schizophrenia, anxiety, obsessive-compulsive disorder, and stroke[37,38]
		Methotrexate analogs	Antitumor agents
Serine derivatives	Homoserine	Faldeprevir	HCV protease inhibitor[39]
		Homoserine lactones	Quorum-sensing inhibition molecules Gram-positive bacteria[40]
		O-allyl serine and O-allyl homoserine	Helix stapling in Ru-catalyzed metathesis reaction[41]
Aliphatic amino acids alternatives	Norleucine		A single mutation can improve the affinity of a hexapeptide to PSD-95 protein[42]
	α-aminoisobutyric acid (Aib)	Taspoglutide	Glucagon-like peptide-1 agonist, diabetes[43]
	tert-butylglycine	Oxytocin	Affinity increase[44]
Proline derivatives	4-hydroxyproline	Collagen	Structural protein in the extracellular matrix
		Pasireotide	Treatment of Cushing disease[45]
		Faldaprevir, asunaprevir	HCV protease inhibitor[46,47]
		Echinocandin B and D	Antifungal cyclic lipopeptides[48]
	Pipecolic acid	Bupivacaine	local anesthetic drug
		Quinupristin, mikamycin B	Cyclopeptide antibiotics
		rapamycin	Macrolide drug used to prevent rejection in organ transplantation
		Simeprevir	HCV protease inhibitor[49-52]
	Fluoroproline		Fluoroproline was used in the development of HIV protease inhibitors[53]
	Cyclic β^3-amino acids		β-peptide foldamers
Phenylalanine derivatives	Thienylalanine	Icatibant	Selective agonist of bradykinin B2 receptor, used to treat acute heart attacks of hereditary angioedema[54]
	Tetrahydro-isoquinoline-3-carboxylic acid (Tic)		
		Oxytocin and somatostatin	Constrain the phenylalanine side chain
	4′-chloro-D-Phe	Cetrorelix	GnRH analog[55]
	Pyridylalanine	cetrorelix	Increases solubility[56]
Tyrosine analogs		L-DOPA	Parkinson's Disease
		Thyroxine	Thyroid hormone deficiency
Tryptophane			Tryptophane is used as an antidepressant and sleep inducer, food intake suppressor, and used in hair loss-promoting agents.[56-58]

3.4. Peptidomimetic PPI inhibitors

PPI inhibitors reported in the literature focus on the mimicry of protein structure. Reported molecules mimic the primary, secondary, tertiary, and even quaternary structures of polypeptides or proteins.[59]

In the case of peptidomimetics, displaying hot-spot residues that match the target protein interface is usually done using secondary or tertiary structure mimetics. Several non-natural polymers tend to fold and are proven to mimic secondary and tertiary protein folds.[60-62]

More than 60% of protein complexes contain helical structures on their interfaces.[63] Hot-spot residues are usually displayed on one or two faces of the helix. Various scaffolds can achieve mimicry of the helix topography to gain an appropriate side chain display.[30,64,65] Ligand flexibility and entropy loss during the binding event are significant issues in peptidomimetic drug discovery. One strategy to mitigate entropic penalty is the stabilization of the helix by side chain stapling[66]. Hydrogen bond surrogates (HBS) are also a viable strategy for stabilization. H-bonds can be substituted with a hydrocarbon chain in the helix.[67] Foldamers are built by using β- and γ- amino acids (exclusively or mixed with α amino acids) and are able to fold into a secondary structure in water. Many PPI inhibitors were designed using foldameric scaffolds.[62,68,69] β-sheets are less prevalent than helices in PPI interfaces and are challenging to mimic due to the increased tendency to aggregate.[70-72]

Cyclic peptides and other macrocycles are also able to mimic secondary structures. Loops and turns are present in over 40% of the PPI interfaces. Cyclotids also display higher in vivo stability, and cyclic peptides have a higher tendency to cross the cell membrane.[73],[74]

Mimicking the tertiary or quaternary structure of proteins might be required when the hot-spot residues are spread out on a larger surface of the target protein. Miniproteins utilizing non-natural scaffolds were also used.[75-78]

The mimicry of pre-existing structural elements mostly dominates peptidomimetic design. Only a few articles touched on the concept of protein surface mimicry[79,80] and fragment-centric modularity in the context of peptidomimetics was investigated by the Arora group.[29] However, no studies concern the bottom-up construction of interfaces for protein surface mimicry.

3.4.1. β-peptides and foldamers

α-peptides are intrinsically unstable, making them less appealing candidates for drug development. Proteolytic instability and the energy cost of finding the optimal binding conformation of a peptide require polymers built out of non-natural monomers where the secondary structure is well controlled. β-amino acids are often used in helical peptidomimetic molecules. β-Substitutions of key residues often increase binding affinity, proteolytic stability, and the helicity of the peptidomimetic molecule. α/β-peptides can be built using the appropriate combinations (ααβ, αααβ, ααβααβ) to retain the helical secondary structure.[81] Peptide sequences constructed from β-

amino acid residues exhibit a strong tendency to fold into secondary structures at short chain lengths. The specific geometry is governed by the backbone stereochemistry pattern and topology of structure-inducing side chains.[82,83] Fully β-peptide foldamers also exhibit interesting secondary structures; 10/12 helix, H12, and H14 helices can be built.[83] β-peptide foldamers have successfully been used for the inhibition of significant drug discovery targets such as Bcl-xL, hDM2-p53, VEGF-VEGFR1 complexes, somatostatin, parathormone, and GLP-1 receptors.[84-89]

β-amino acids exhibit greater conformational flexibility than α-amino acids due to the additional carbon atom in the backbone. The new torsion angle (θ) defined by the C2-C3 bond is forced into a more stable gauche conformation when the amino acid is substituted at position 3, e.g., β^3-amino acids or cyclic amino acids. Here we utilized (1S,2S)-2-aminocyclopentanecarboxylic acid (ACPC) and (1S,2S)-2-aminocyclohexanecarboxylic acid (ACHC) alongside β^3-amino acids.

The alpha helix is the most abundant secondary structure in protein-protein interfaces.[90] The helix is stabilized by intramolecular hydrogen bonds between residues i and i+n. Helices often found in proteins mostly follow the i, i+3 and i, i+4 patterns giving rise to the 3_{10} and 3.6_{13} type helices. The nomenclature shows the number of residues per turn and the number of atoms in the ring formed by the hydrogen bond between the carbonyl group of the amino acid at position i and the amide proton at position i+n in subscript.[90] β-peptides can also fold into helical structures. H14 helix (3_{14}-helix) consists of three residues per turn stabilized by 14-membered rings formed through hydrogen bonds between residues i, i+2. Another helix type is the H12 (2.5_{12}-helix) stabilized by hydrogen bonds between residues i, i+3, containing two and a half amino acids per turn (Figure 3). [83,91,92]

Here in our work, helices H14 and H12 were utilized. These secondary structures can be stabilized by cyclic side chain residues so that they display folding at lengths of hexamers and octamers for H14 ((1S,2S)-2-aminocyclohexanecarboxylic acid, ACHC) and H12 ((1S,2S)-2-aminocyclopentanecarboxylic acid, ACPC), respectively. Such oligomers allow two proteinogenic side chains that point to the identical face of the helix to be incorporated into the sequence. The closely packed side chain distances in α-helices (C^{α}_{i}–$C^{\alpha}_{i+3} \approx 5$ Å and C^{α}_{i}–$C^{\alpha}_{i+4} \approx 6$ Å) and in β-sheets (C^{α}_{i}–$C^{\alpha}_{i+2} \approx 6$ Å and "sideways" C^{α}_{i}–$C^{\alpha}_{i+n} \approx 5$ Å) are comparable with the closest side chain distances in foldameric helices H14 (C^{β}_{i}–$C^{\beta}_{i+3} \approx 5$ Å) and H12 (C^{β}_{i}–$C^{\beta}_{i+2} \approx 6$ Å)(Figure 3).[69] Helix H14 exhibits a strong tendency to fold in aqueous solution. Its side chains are oriented in parallel on one face of the helix. Helix H12 is conformationally less stable in an aqueous medium, demonstrating an elongated geometry, and side chains are juxtaposed with a slight tilt angle.[93]

According to Bartus et al., these helices can recognize hot-spot residues on the surface of calmodulin. They constructed two 256-membered libraries constructed using H12 and H14 scaffolds.[69] These helices are the shortest folded β-peptides and can be considered minimal building blocks among peptidomimetics. They display a surface area of *ca.* 500 Å^2, providing the binding free energy for affinities in the range of 1-500 μM. Foldamer helices used in our experiments will be referred to as building blocks.

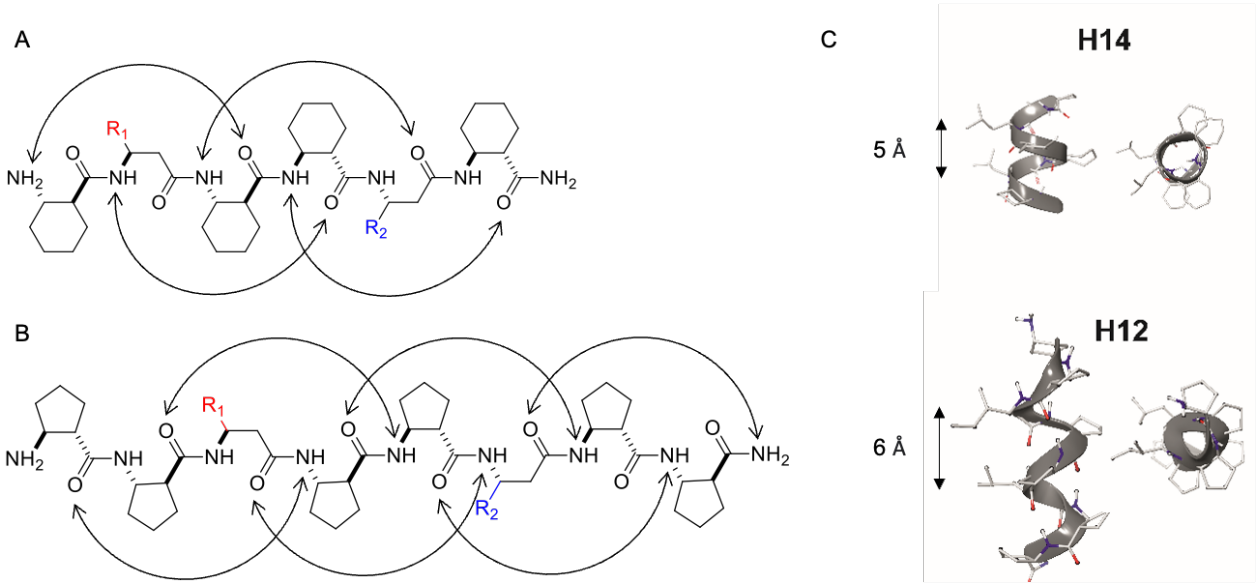

Figure 3. Design of local surface mimetic probe libraries using short foldamers. H14 scaffold (A), H12 scaffold (B), and the structure of H14 and H12 helices (C).

3.5. Design strategies for peptidomimetics

3.5.1. Top-down design of PPI inhibitors

In the search for a peptidomimetic drug candidate, structural data on the target protein and a segment of the interacting partner paves a straightforward way for top-down design. Usually, native peptides are the starting point of structure-based drug design since peptides already have high affinity and/or specificity for their targets.[94-96] The mimicry of binding conformation and critical residues, which has the most significant contribution to binding affinity, is vital in peptidomimetic design. Iterative modification of peptide residues or the rational design of a non-peptidic peptidomimetic is all considered top-down design strategies.[97,98]

Many PPI inhibitors containing β-amino acid residues were reported in the literature.[99-104] Structural compatibility of β-amino acid-containing peptides has been addressed; β-amino acid

substitutions are only tolerated when applied on the non-interface part of the parent molecule.[105,106] β-amino acid substitutions of the interface residues can result in a decrease in binding affinity.[105,106] A top-down discovery scenario requires a large amount of data on the drug target, e.g., high-definition structural information on the target protein. Many of the techniques are explained in section 3.2. can be considered as top-down drug design strategies, e.g., forward chemical genetics or structure-based drug design. Data science-assisted methods such as machine learning and deep learning have already made their way to drug discovery, providing valuable input for top-down design strategies using the large quantities of data available.[107]

3.5.2. Bottom-up approach for the design of PPI inhibitors

The control of peptidomimetics' complex functional and geometrical characteristics is required to develop a PPI inhibitor. Developing peptidomimetic PPI inhibitors is often done by top-down design approaches, for which detailed information on the interacting partners is necessary.[108] Bottom-up design does not necessarily require information on the target protein in atomistic detail. Still, control over the building block assembly is fundamental in this approach. Foldamers have a well-controlled secondary structure and are able to fold at short chain lengths, enabling a foldameric bottom-up design approach.

Bottom-up ligand development has been established with small molecules, which is the fragment-based design. This conceptual framework can be applied to larger, surface mimetic structures.[109] Fragments are weakly binding (0.1 – 10 mM) building blocks that can be linked or further grown after the initial screening to yield a high-affinity compound.[110] Fragment screening may be assisted by NMR, X-ray, fluorescence anisotropy, isothermal titration calorimetry, and surface plasmon resonance techniques. High throughput screening (HTS) is a compatible technique with fragment-based techniques; it enables the screening of enormous libraries, covering large chemical space in a short period of time.[109,110] HTS and fragment-based screens have not been used in foldameric design.

3.5.3. Fragment linking under thermodynamic control using dynamic combinatorial chemistry

Dynamic combinatorial chemistry (DCL) eliminates the labor-intensive processes of conventional drug discovery, enabling the screening and synthesis of drug-like molecules in one step. Dynamic combinatorial libraries enable the reversible assembly of complex molecules from simple building blocks under thermodynamic control. In DCL, disulfide and hydrazone chemistry can be utilized in component exchange reactions, among many other reversible reactions.[111,112] The system can

adapt to external stimuli shifting the equilibrium accordingly. Covalent and noncovalent interactions are involved in the final product interconversion. The final product distribution is dominated by the free energy of a given compound.[113]

Bartus et al. has already shown the feasibility of dynamic covalent assembly of foldamer helices to inhibit protein-protein interaction.[69] The bottom-up assembly of foldameric helices yielded tight binding helix dimers, which could compete with TRPV peptide, one of the natural ligands of calmodulin. However, the thiolate mechanism is rather slow, and the product distribution is under thermodynamic control. A faster disulfide metathesis approach using a radical mechanism and UV light-induced exchange has been reported.[114] This system is far from equilibrium; changing the energy influx enables control over the final product distribution.[115]

4. Experimental methods

4.1. Synthesis and purification of peptides

Foldameric surface probes, competitor peptides for LSM pulldown assay and SSB-Ct peptides were synthesized manually by SPPS, with the utilization of Fmoc/tBu strategy using Rink Amide AM resin (foldamers) and Wang resin (SSB peptides). Loading of the Wang resin was carried out by using 4 equivalents of Fmoc-amino acid, 4 equivalents of hydroxybenzotriazole (HOBt), 4 equivalents of N,N′-diisopropylcarbodiimide (DIC), and 0.1 equivalents of DMAP in dichloromethane (DCM)/*N,N*-dimethylformamide (DMF) (50:50 v/v%) for 4 hours. The resin loading was determined by measuring the absorbance of Fmoc on λ = 290 nm. The deprotection protocol for removing the Fmoc group was done by using 5% piperidine and 2% 1,8-diazabicycloundec-7-ene (DBU) in (DMF) for 3 and 5 minutes, consecutively. The resin was washed using DMF and DCM after every step. Amino acid coupling was carried out by adding 3-fold excess of the following reagents: amino acid (N-Fmoc protected), 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]-pyridinium-3-oxide hexafluorophosphate (HATU) and 6 equivalents of *N,N*-diisopropylethylamine (DIPEA) in DMF for 3 hours. SSB-Ct peptides were capped either with acetylation or 5,6-carboxyfluorescein(F)-Gly-Gly moiety.

Cleavage of rink Amide AM was performed with a mixture of trifluoroacetic acid (TFA)/H_2O/1,4-dithiothreitol (DTT)/triisopropylsilane (TIS) (90:5:2.5:2.5 v/v/m/v%) at room temperature for 3 hours. DTT was omitted in the case of peptides containing Cys residue. Wang resin cleavage was carried out using DCM/TFA/H_2O/TIS (42.5:50:5:2.5 v/v/v/v%) for 1 hour to decrease aspartimide formation. The cleavage cocktail was dried in a vacuum, and the peptide was precipitated in ice-cold, dry diethyl ether. The peptide precipitate was filtered and re-dissolved prior to lyophilization.

4.2. Synthesis of the foldamer fragment libraries

Foldamer libraries were synthesized manually using split and pool synthesis. Rink Amide PS resin was used as solid support. After the initial deprotection of the resin, 3 equivalents of (1*S*,2*S*)-Fmoc-2-aminocyclohexane carboxylic acid was coupled using 3 equivalents of HATU and 6 equivalents of DIPEA. The resin was split into 4, and β^3-homoamino acids in position 5 were coupled. The resin was pooled, and two (1*S*,2*S*)-Fmoc-2-aminocyclohexane carboxylic acid residues were coupled. The resin was split into 16, and β^3-homoamino acids in position 2 were coupled. The resin

was pooled back again, and one (1*S*,2*S*)-Fmoc-2-aminocyclohexane carboxylic acid was coupled. As shown above, the foldamer peptide library was cleaved and precipitated before lyophilization.

4.3. Peptide purification

Peptides and conjugates were purified on a semi-preparative (250 × 10.00 mm) or a preparative (250 × 21.2 mm) RP-HPLC column. Jupiter (particle size: 10 μm, pore size: 300 Å) C18 column and eluents (A) 0.1% TFA in water and (B) 0.1% TFA in ACN/water (80/20) were used. Peptide purity was above 95%, assessed by HPLC-UV and ESI-MS measurements. Foldamer libraries were purified using a gradient (0–90% during 90 minutes with a flow rate of 4 mL min^{-1}). Fractions were analyzed by HPLC-MS. Each fraction containing library members was pooled together prior to lyophilization. Library components were identified by HPLC-MS based on molecular weight and retention time estimated by the hydrophobic properties of the peptides. Purity analysis was based on quantifying the total library members and impurities by integrating the HPLC-MS chromatograms.

4.4. Protein expression and purification

Human S100 proteins and human galectin were gifts from László Nyitrai (Eötvös Lorand University, Faculty of Sciences, Budapest) and Éva Monostori (Biological Research Centre, Szeged), respectively. Every other protein, bovine calmodulin, in *E. coli* RecQ-WH, ExoI, and RecO was produced in our lab. Bovine calmodulin and *E. coli* RecQ-WH were cloned into pET28a+ expression vector. *E. coli* ExoI (in pET14b) was a kind gift from Charles Bell (Addgene plasmid # 104552; http://n2t.net/addgene:104552; RRID:Addgene_104552). Gene coding *E. coli* RecO was synthesized and cloned by Proteogenix (France) into pET28a+ containing a TEV cleavage site between the protein and the His-tag. Every expression cassette was under the control of the T7 promoter and *lac* operator and thus transformed into *Escherichia coli* (BL21 DE3) cells for expression. In the case of *E. coli* RecO, pLysS helper plasmid was used to suppress the basal activity of the T7 promoter. All expressing strains were grown in LB liquid media at 37 °C until OD_{600} = 0.6; then, expression was induced by adding 100 μM IPTG and was carried out overnight at 22 °C. *E. coli* RecO was an exception; due to solid inclusion body formation fermentation was carried out in M9 media and 18°C overnight to slow down expression. Cells were harvested and lysed by sonication. The clear lysate was loaded onto Ni-NTA affinity column. After 30 minutes of incubation, the resin was washed three times, and the protein was eluted in an imidazole-containing buffer. The protein purification process was followed on SDS-PAGE. Proteins RecQ-WH and calmodulin had their affinity tags removed by thrombin cleavage and were subjected to

further gel filtration. Protein ExoI and RecO were treated with thrombin and TEV protease, respectively, and were further purified by Heparin chromatography. The purity and molecular weight of the proteins were assessed by HPLC-ESI-MS. Protein concentration was determined by Thermo Scientific NanoDrop One Microvolume UV-Vis Spectrophotometer after extensive dialysis into a buffer required for the downstream experiment.

4.5. Pulldown assay

Foldameric libraries were screened using a pulldown assay. Cobalt affinity resin (TALON, Takara Bio USA, Inc., Mountain View, CA) was pre-equilibrated with a buffer preferred by proteins. It was incubated with 64 μM protein for 30 minutes in a paper filter spin cup. The resin was washed and incubated with each sublibrary for 30 minutes. Each member of the sublibrary was 1 μM, giving an equimolar setup (64 μM protein and 1 μM helix in a 64-membered library). Competition pulldown assay contained competitor peptides alongside the foldamer library. Non-bound foldamers were collected by centrifugation. Washing steps and elution fraction collection was done by centrifugation in a table-top centrifuge at 100 rpm for 2 minutes. The same pulldown experiment was performed in the absence of protein. It was used as a negative control to assess the non-specific binding between the resin/paper column and the foldamers. Collected samples were measured on HPLC-MS having a Thermo Scientific Dionex UltiMate 3000 HPLC system interfaced to an LTQ ion trap mass spectrometer (Thermo Electron Corp., San Jose, CA, USA) using a Phenomenex Aeris Widepore XB-C18 (250 × 4.6 mm, particle size: 3.6 μm, pore size: 200 Å) column running a gradient elution 5–80% solution B during 25 minutes with a flow rate of 0.7 mL min^{-1}. Eluent compositions were 0.1% formic acid in distilled water (solution A) and 0.1% formic acid in acetonitrile (solution B). Mass spectra were acquired in full scan mode in the 200 to 2000 m/z range. For overlapping peaks, selective reaction monitoring (SRM) was used. Thermo Xcalibur 2.2 software package was used for peak identification and integration. $AUC^{i}_{protein}$ and $AUC^{i}_{control}$ are obtained for compound *i* in the LSM library in the experiment with immobilized protein and the control experiment without protein, respectively. The following formula shows how apparent K_D was calculated. *n* stands for the number of sites, which was 1 for all proteins except for calmodulin. Calmodulin is capable of binding two helices at the same time, according to our ITC experiments (not shown here).

$$K_D = \frac{c^i_{free\ ligand} \times c_{free\ protein}}{c^i_{complex}} = \frac{(1-F^i_B)\times(64\times n - \sum_i^N F^i_B)}{F^i_B} \times 10^{-6} \quad (1)$$

4.6. Fluorescence anisotropy experiments

Triplicate measurements were performed using a Clariostar Plus Microplate Reader in 384-well plates for all fluorescence anisotropy (FA) experiments. The settings were 25 °C, excitation at 485 nm, emission at 510 nm, settling time of 0.2 s, and 200 flashes per well.

Competition fluorescence anisotropy experiments were carried out as follows: protein (0.5 μM either ExoI or RecO) was incubated with 50 nM F-wtSSB peptide and 0–500 μM unlabeled wtSSB (or a variant) peptide for 30 min at room temperature. For direct fluorescence anisotropy measurements, 16 μM ExoI and 7.5 μM RecO were used in the first well. After serial dilution, either 50 nM of F-mSSB or F-sSSB was added to the wells and then incubated for 30 min. Buffers used included 20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1 mM $MgCl_2$, 1 mM 2-mercaptoethanol, 10% (v/v) glycerol, for ExoI, and 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 1 mM 2-mercaptoethanol, 10% (v/v) glycerol, and 0.01% Triton-X 100 for RecO.

Fluorescence anisotropy values were normalized and are shown as bound fractions (BF, %).

$$BF = \frac{r - r_{min}}{\lambda(r_{max} - r) + r - r_{min}} \quad (2)$$

$$I = 2 * PG + S, r = S - \frac{PG}{I} \quad (3)$$

where r: anisotropy, I: total intensity, P: perpendicular intensity, S: parallel intensity, G: an instrument factor set to 1, BF: ligand fraction bound, and $\lambda = \lambda_{bound}/\lambda_{unbound} = 1$.

Competitive fluorescence anisotropy was analyzed in Origin Pro 9.5, and IC_{50} values were determined using the Logistic Nonlinear fit function. EC_{50} values from direct fluorescence anisotropy titrations were plotted in Origin 9.5. Values were determined using the Logistics Nonlinear fit function.

4.7. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were performed using a MicroCal VP-ITC titration microcalorimeter. ExoI was dialyzed extensively against the indicated buffer and cleared through centrifugation at 14,000 rpm for 15 min at 4 °C. RecO was subjected to a quick buffer exchange using an Amicon Ultra ultrafiltration device (Sigma). Protein concentrations were determined after the buffer exchange procedure. Wild-type SSB-Ct or modified SSB-Ct were titrated into ExoI or RecO in 20 mM Tris–HCl (pH 8.0), 100 mM NaCl, 1 mM $MgCl_2$, 4% glycerol, 1 mM β-mercaptoethanol or 50 mM HEPES pH 7.5, 50 mM NaCl, 25% glycerol, and 1 mM TCEP, respectively.

Isothermal titration calorimetry. Isothermal titration calorimetry (ITC) experiments were performed using a MicroCal VP-ITC titration microcalorimeter at 30 °C. 300 rpm stirring was used. Ligand solution was titrated to the protein-containing cell using 20 μL volumes over 300 seconds of injection time.

The raw data were integrated using NITPIC.[116] The binding parameters, association equilibrium constant (K_A), binding enthalpy (ΔH), and binding entropy (ΔS) were obtained by fitting the titration curves to a model of A + B = C in SEDPHAT.[117] Stoichiometry was fixed to 1:1, and incompetent fraction was fitted for protein concentration.

4.8. Modeling

For the structural characterization of the wild-type and the modified SSB-Ct peptides binding to the investigated enzymes, an extended conformational sampling method, the replica-exchange solute tempering molecular dynamics[47,48] was applied as implemented in the Desmond package.[49] Modeling was performed by our colleagues.

5. Results and Discussion

5.1. Local surface mimetics

H14 and H12 helices are the short, folded peptide foldamer building blocks.[118] These helices can display a surface area of ca. 500 Å^2 toward a protein surface, providing the binding free energy for affinities in the range of 1–500 μM. Such affinity is sufficient for probing the surface and is helpful for subsequent fragment-based design. Foldameric fragments will be referred to as local surface mimetics (LSM probes).

5.1.1. Proteins mapped using LSMs

Five proteins (Table 2.) were selected based on the numbers of their interacting protein partners (Figure 4), levels of interprotein structural homology, and types of interactions (protein–protein, protein–carbohydrate). The reference protein in this set was CaM,[119] a highly pleiotropic molecule that plays a vital role in every eukaryotic cell. It displays EF-hand motifs and mediates an extremely large number of signal pathways via the control of enzyme and ion channel activities in a Ca2+-dependent manner. Two additional EF-hand proteins with high structural homology were selected from the S100 family: S100A4 and S100B,[120] both important pharmacological targets for cancer treatment.[121] S100A4 is involved in tumor progression and metastasis and has many interacting partners. S100B is expressed in melanocyte-derived tumors and mature astrocytes,[122] where it plays an important role in neurite extension. The numbers of its known protein partners are significantly less than that for S100A4. Galectin-1 (Gal-1) and the winged helix domain of RecQ helicase (RecQ-WH) were used as target models with a low tendency to form PPIs. Gal-1 has an immunosuppressive effect, promoting cancer progression and metastasis through the recognition of β-galactoside motifs on cell surface glycoproteins.[123] Gal-1 has a β-sandwich structure with a jelly roll topology,[124,125] that may interact with proteins in a carbohydrate-independent manner.[123,126,127] RecQ is a prokaryotic intracellular ATP-dependent DNA helicase that plays an important role in DNA damage response, recombination, replication, and repair.[128,129] RecQ interacts with the disordered C-terminal of the single-stranded DNA-binding protein (SSB) through its WH domain, the single PPI reported for RecQ-WH.[130] All selected targets exhibit solvent-exposed shallow binding sites; however, CaM is able to wrap around α-helices with special side chain motifs, thereby forming a hydrophobic pocket.[131] For competitive binding studies, peptide motifs of PPI partners were selected from the following proteins: transient receptor potential cation

channel subfamily V member 1 (TRPV1), non-muscle myosin IIA (NMIIA), ribosomal S6 kinase 1 (RSK1), and SSB, for CaM, S100A4, S100B, and RecQ-WH, respectively.

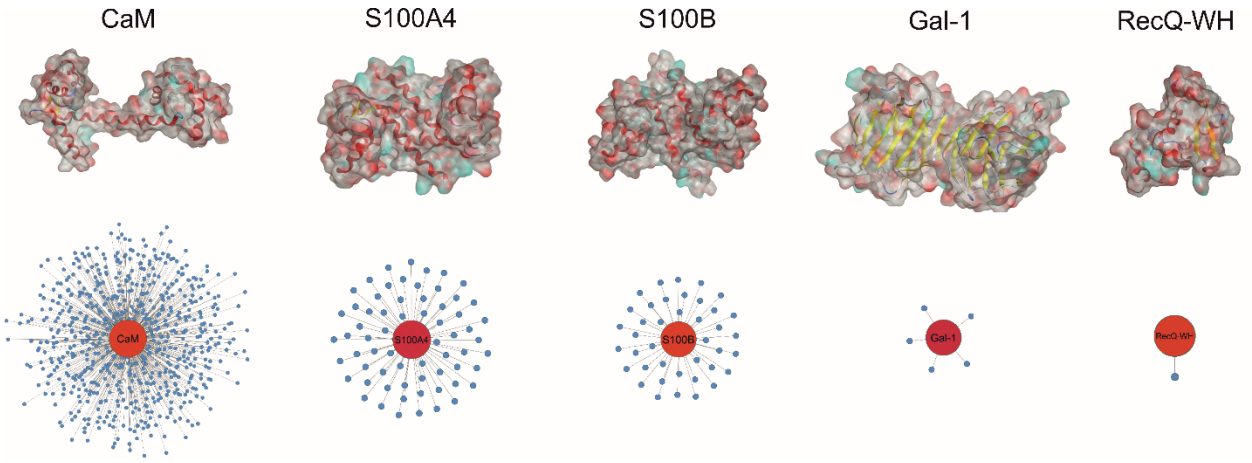

Figure 4. Surface representation and PPI network of selected proteins. The number of interacting proteins (N_{PPI}) is represented as nodes. PPI data were obtained from BioGRID,[132] Wiki-Pi,[133] GPS-Prot,[134] IntAct[135], and APID.[136] For the lectin Gal-1, NPPI is based on the work of Camby et. at.[123]

Table 2. Characteristics of protein screened using LSMs.

	CaM	S100A4	S100B	Gal-1	RecQ-WH
Helix content (%)	62	55	58	0	50
β-sheet content (%)	4	3	2	54	13
Number of PPIs[a]	645	67	41	5	1
Molecular weight (kDa)	16.8	11.7*2[b]	10.7*2[b]	14.6*2[b]	12.9
Isoelectric point[c]	4.09	5.85	4.52	5.30	10.18

[a]Average values from databases, BioGRID, Wiki-Pi, GPS-Prot, IntAct, and APID. For Gal-1, N_{PPI} is given based on the review of Camby et al.[66] [b]Homodimers. [c]Calculated values based on amino acid composition

5.1.1. Design of local surface mimetic library

Peptide sequences constructed from β-amino acid residues exhibit a strong tendency to fold into secondary structures at short chain lengths, and the specific geometry is controlled by the backbone stereochemistry pattern and topology of structure-inducing side chains. The H14 and H12 helices utilized in this study project two proteogenic side chains on the same face of the helix. These scaffolds allow for comparable side chain distances to native PPI interfaces (5-6 Å). Sixteen different β^3-amino acids were substituted at positions R1 and R2, generating 512 local surface mimetics. LSM probes were synthesized and screened as sublibraries (L1-L4) (Figure 5). Four

amino acids were excluded from this pool: glycine, proline, cysteine, and histidine. LSMs are named using the letter codes in the R1 and R2 positions (starting from the N-terminus).

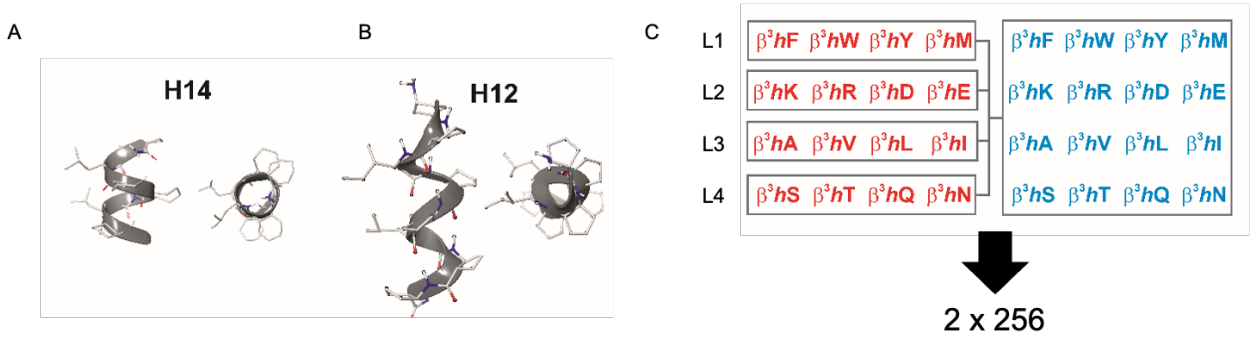

Figure 5. Design of local surface mimetic probe libraries using short foldamers. Structures of H14 (A) and H12 helices (B) and library composition showing sublibrary grouping (C).

5.1.2. LSM affinity patterns indicate secondary structure-dependent structural compatibility

The protein set was probed in pull-down assays. Each protein was equilibrated with probes in an equimolar setup (see section 4. Materials and Methods, 4.5 Pull-down assay) to minimize competition between LSM probes. After 30 minutes of equilibration, the unbound fraction of foldamers was collected and measured using HPLC-MS. Hits were corrected with the background binding of the resin to eliminate unspecific binding. Bound fractions of foldamer hits were used to calculate apparent K_D values (see section 4. Materials and Methods, 4.5 Pull down assay) visualized as heat maps of dissociation constants (Figure 6). Considering the experimental error of fraction bound (FB), the affinity limit was set at 150 µM to filter successful hits (Figure 6, indicated in red), which is 50 µM above the 100 µM guideline normally applied in fragment-based drug design.[137]

A large number of hits were obtained across all proteins, supporting the hypothesis that LSMs can display a sufficient contact area toward their targets. H12 probes were only favored by S100B and yielded fever hits overall, possibly due to less stable helical conformation of H12.

The H14 helical probes have an overall ability to act as a local probe and to adapt to the surface features of different proteins studied here, suggesting structural compatibility between foldamers and proteins. Helices are symmetrical: two helices with the same two proteinogenic side chains only differ in their N and C termini. Interestingly, the orientation of the R1 and R2 proteogenic side chains had a minor impact on binding, as indicated by the diagonal symmetry in the heat maps.

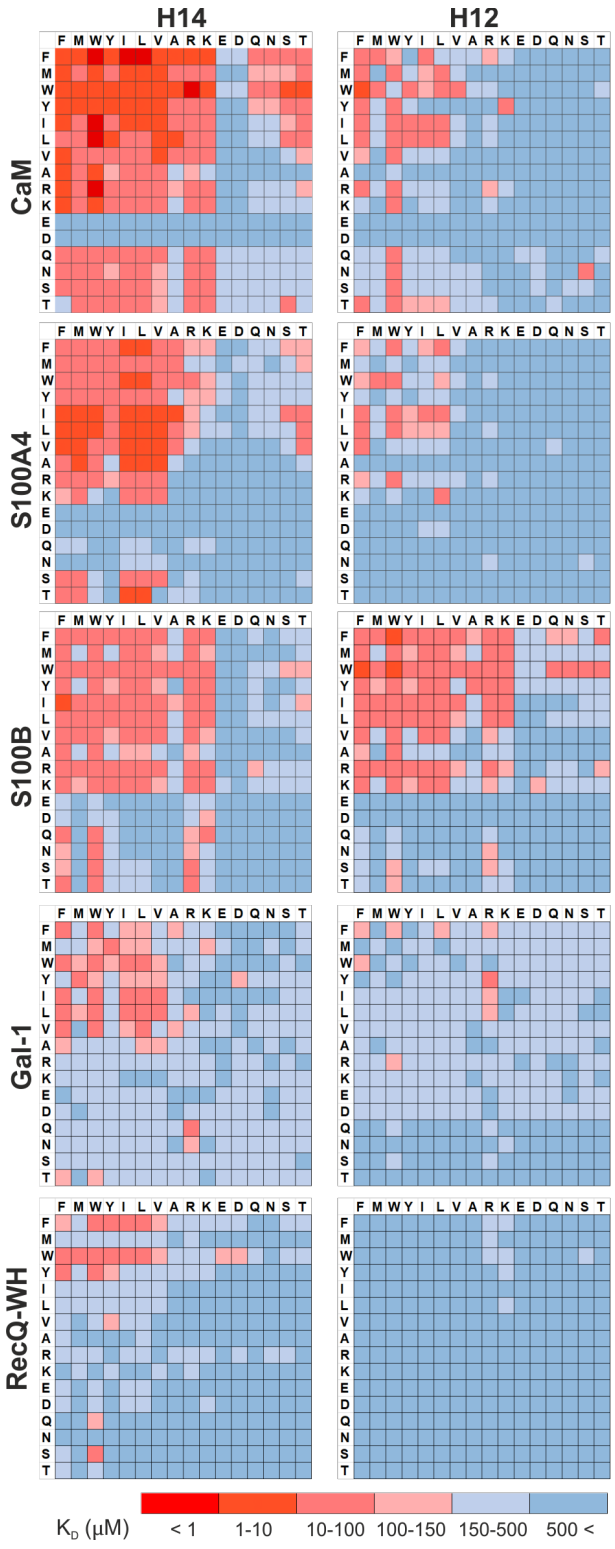

Figure 6. Binding patterns of foldameric LSM probes are represented as heat maps in the K_D dimension (in μM). Apparent dissociation constants are given for H14- and H12-helical LSM libraries. One letter codes displayed at the headings of the rows and columns correspond to the proteinogenic side chains of the β^3-homo-amino acids used in the R1 and R2 positions.

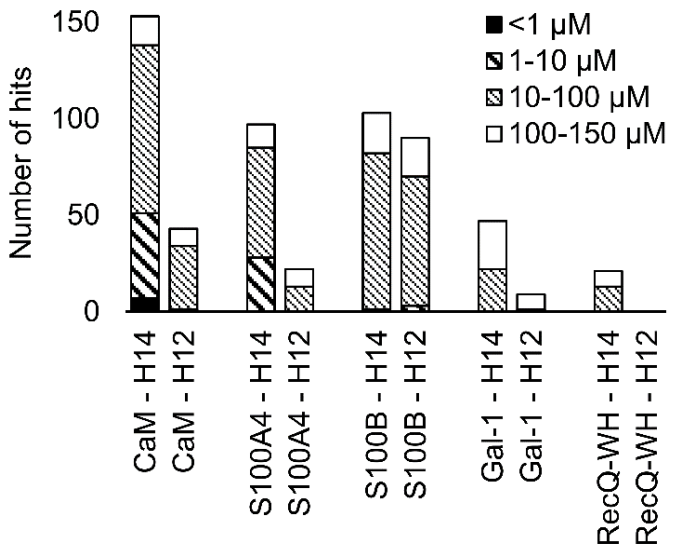

Figure 7. Number of hits obtained in the H14 and H12 LSM pull-down experiments and their affinity distributions.

5.1.3. Foldameric LSMs detect orthosteric and non-orthosteric spots

Interactions between the H14 helical LSMs and the targets were abundant, but safe recognition of the native PPI interfaces on the targets is a criterion of druggability. We have performed pull-down assays in the presence of native ligands of the target proteins (Table 3). For highlighting the effect of the competitor on the binding fingerprints, K_D ratios and the LSM replacement percentages were calculated and are represented as heat maps (Figure 8). We found that native ligands changed the LSM affinity patterns for targets exhibiting direct protein–protein contacts (CaM, S100A4, S100B, and RecQ-WH), and many of the LSM probes were displaced. Since these proteins exhibit geometrically a single surface region to form PPI interfaces, residual binders obviously interacted with the non-orthosteric spots. In contrast, only a few LSM probes displayed replacement for Gal-1. This could be explained by the glycan selective native recognition domain of the protein. We note that the replacement pattern for a single probe can, in theory, be complex. While orthosteric weak-binder probes could be completely replaced from interaction sites, high-affinity foldamer hits might bind to targets even in the presence of competitors, although significantly increased apparent K_D values clearly indicated orthosteric foldamer probes. Moreover, a single probe can bind to both orthosteric and non-orthosteric spots, limiting overall displacement levels. To avoid overlooking interesting orthosteric weak binders, direct replacement levels were assessed.

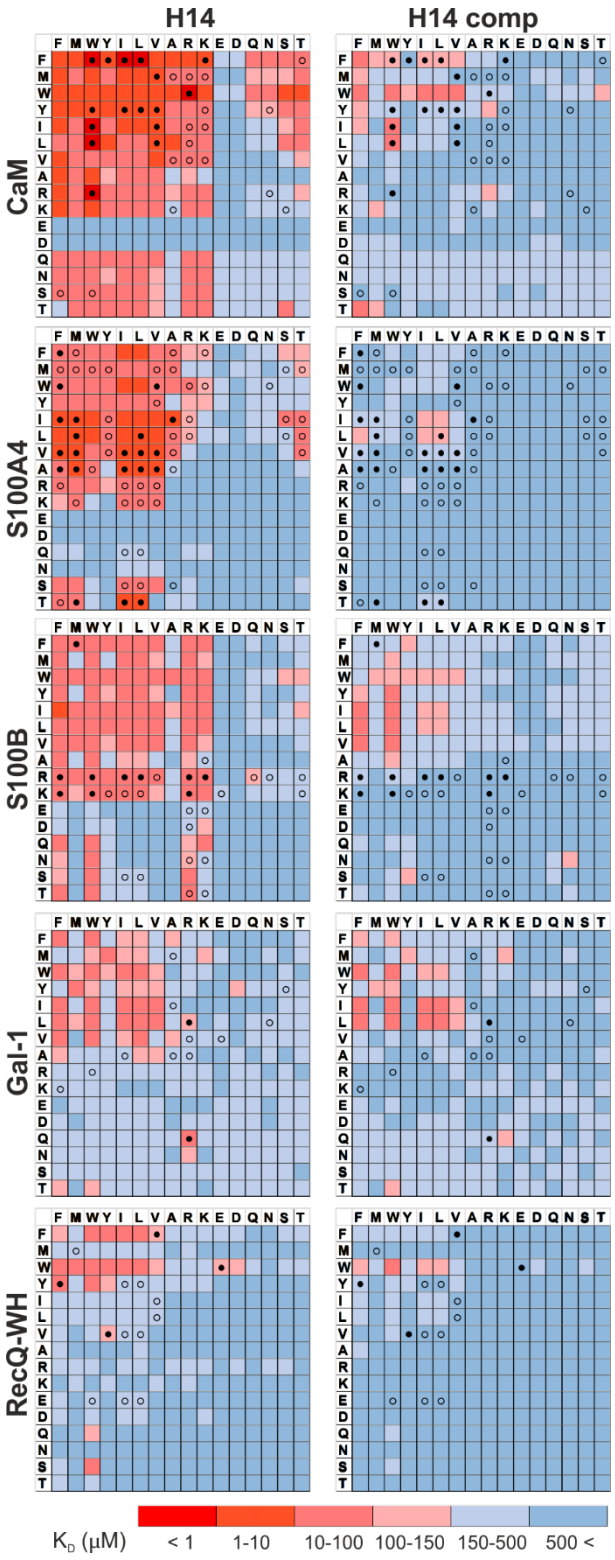

Figure 8. Competition pull-down experiments with the H14 LSM library. K_D heat maps show side chain preferences using H14 helices alone and in the presence of the native ligands as competitors (left and right, respectively). Foldamers with significant replacement percentages (> 80% for RecQ-WH, > 90% for the other proteins) are marked with an open circle, while foldamer hits with significant K_D increase are marked with a filled circle.

Table 3. Sequences and dissociation constants of the native ligands applied as competitors

Protein	Competitor	K_D
CaM	TRPV1-Ct15	30.9 ± 2.1 nM[69]
S100A4	NMIIA (1893–1923)	7 ± 1 nM[138]
S100B	RSK1 (689–735)	1.8 mM[69]
Gal-1	lactose	409 mM[139]
RecQ-WH	SSB-Ct	16.6 ± 0.8 mM[140]

5.1.4. Affinity patterns of foldameric LSMs are characteristic of the target proteins

To investigate if LSM probes are able to distinguish proteins by their PPI interfaces, we have calculated the correlation between the promiscuity of each protein toward LSM probes (bound fraction) and their natural interactome (number of PPI partners). Interactome size was assessed by PPI databases (BioGRID,[132] Wiki-Pi,[133] GPS-Prot,[134] IntAct,[135], and APID[136]). We note that false positive rates may be high in HTS proteomics, although the magnitude of promiscuity is well visible from these data.[141]

There is a correlation between the number of database PPI partners (N_{PPI}) and the average bound fraction values (Figure 9). Logarithmic scaling was used to explain the exponential relationship between the fragment space of foldameric probes and the chemical space of proteins used. Our findings suggest that H14 probes distinguish proteins with different levels of promiscuity, displaying biomimetic behavior, a well-desirable trait for surface mimetic drug design.

We calculated the pairwise covariance values from the bound fraction values for each protein pair to determine the similarities between affinity patterns (Figure 10). These scaled covariance values were obtained from all 256 H14 LSMs.

The pairwise scaled covariances for orthosteric binding were consistently lower, indicating that non-orthosteric binders exhibit promiscuous behavior. Notably, the scaled covariances for orthosteric binders were between 13-20% for proteins with high structural homology (CaM, S100A4, S100B), suggesting that our foldameric probes detected a low level of interface similarity.

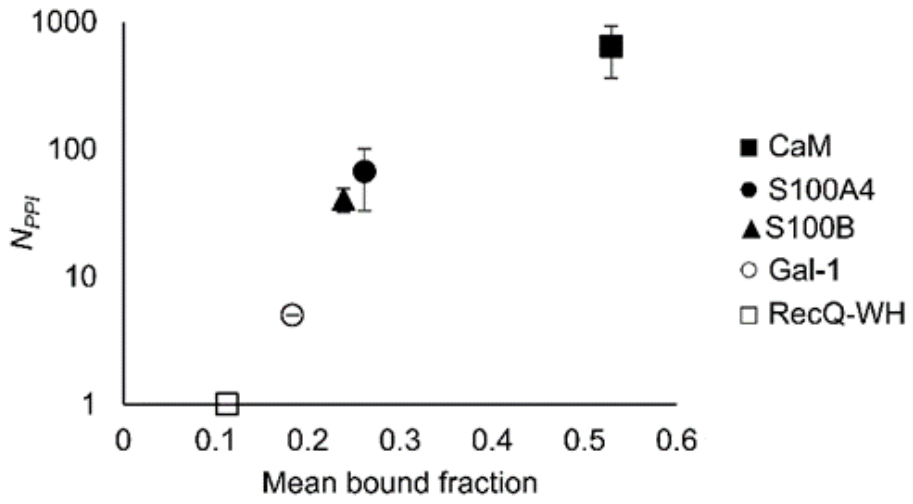

Figure 9. Average bound fractions for the H14 helical LSM library compared with the mean number of PPIs found in databases BioGRID, Wiki-Pi, GPS-Prot, IntAct, and APID. For the lectin, Gal-1, NPPI is based on ref 112.

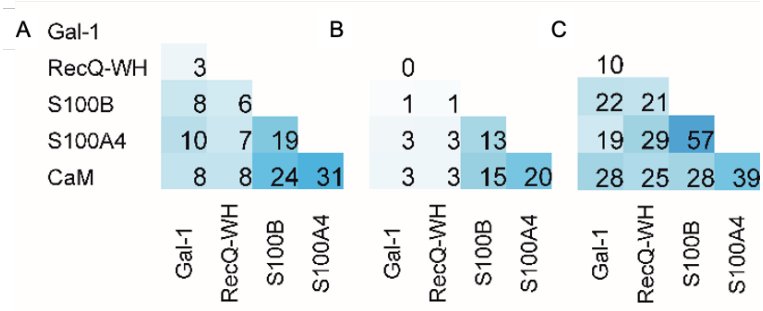

Figure 10. Pairwise scaled covariances (%, maximum similarity: 100%, zero covariance: 0%) of H14 LSM fraction bound (FB) patterns calculated for total (A) and the orthosteric (B) FB values. %. The sequence homologies of proteins are given in panel (C).

5.1.5. Side chain binding propensities are biomimetic

The binding of LSM probes to their targets can be attributed to features at the side chain level. We would have observed uniform baseline binding patterns if the binding was solely driven by the hydrophobic and cyclic β-amino acid residues (ACPC, ACHC). Additionally, many sequences showed no significant binding.

Normalized frequencies (w_j) were calculated for each residue to compare side chain enrichment on the foldamer-protein interface to naturally occurring protein-protein interaction interfaces.

The literature definition of w_j[37] is given in Equation (4):

$$w_j = \frac{f_j}{\sum_m f_m \times v_j} \quad (4)$$

where f_j is the number of interface residues of type j, and normalization factor v_j is the overall prevalence of the amino acid of type j. Index m denotes the residue type. H14 LSM bound fractions were used to approximate residue frequencies, as shown in Equation (5).

$$w_j = \frac{\sum_i F_B^{i,j}}{2 \times \sum_i F_B^i \times v_j} = \frac{\sum_i F_B^{i,j}}{2 \times \sum_i F_B^i \times 0.0625} \quad (5)$$

Index i denotes the LSM sequence, and index j is the residue type. $FB_{i,j}$ is the bound fraction of the LSM probe i, if it contains residue type j. $FB_{i,j}$ is zero if sequence i does not contain residue type j. This formula assumes that a binder LSM probe has both proteinogenic side chains in contact with the protein interface. The LSM library was uniform concerning the 16 side chains, therefore, v_j = 1/16 = 0.0625 was applied as a uniform normalization factor. Normalized frequencies were compared to literature data assessing datasets from the Protein Data Bank: (i) computational Ala scanning[142] and PPI interface analysis[3] (Figure 11).

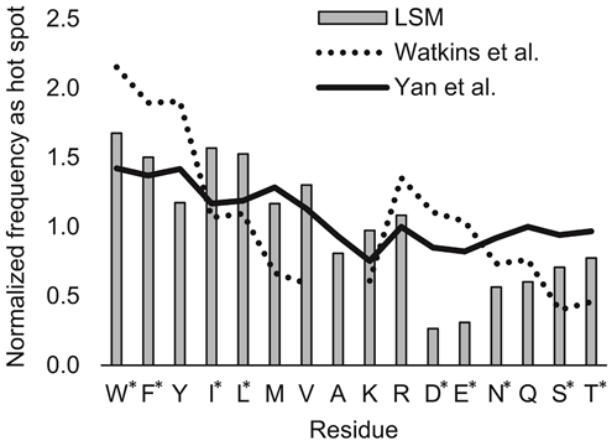

Figure 11. Normalized frequencies were obtained from literature data and calculated for the H14 LSM library. Values from Watkins et al.[142] and Yan et al.[3] are based on computational Ala scanning and PPI interface analysis, respectively.

Our data is in quantitative agreement with previous literature showing an overall enrichment of residues W, F, Y, I, L, M, V, and R and depletion of residues A, D, E, N, Q, S, and T.[3,142,143] For the test protein set, side chain frequency levels varied depending on the protein, highlighting differences in the PPI interfaces. The affinity patterns we observed closely match the natural side chain preferences of each protein tested. Calmodulin prefers cationic and hydrophobic residues as anchor points.[144] S100A4 forms strong interactions with sequences containing Ser, Thr, and Met. S100B interacts with sequences containing Leu, Ile, and cationic side chains, as demonstrated by

the S100B-p53 complexes where Arg, Lys, Leu, and Ile residues are the main drivers of the interaction.[145-147] In contrast to S100A4, the specific enrichment of Arg and Lys was observed in the LSM probe binding fingerprints of S100B. Gal-1 and RecQ-WH were chosen due to their low tendency to form protein-protein complexes. Gal-1 has a small number of PPI partners, reflected in its affinity patterns. RecQ-WH interacts with an intrinsically disordered anionic peptide (DFDDDIPF). Its binding site includes a hydrophobic pocket (binding site of the IPF motif) flanked by polar and positively charged residues.[130] The bulky surface fragments cannot fit into the narrow, slightly bent binding site, leading to a low number of hits. Even negatively charged residues showed no enrichment, possibly due to steric incompatibility. Our findings conclude that the side chains selected from the LSM libraries are similar to the selection of natural interacting partners. We can also conclude that our foldameric probes are biomimetic, meeting critical criteria for drug design, and may serve as minimal surface mimetic motifs.

5.2. Designing bacterial single-stranded DNA-binding protein (SSB) mimetic peptides to inhibit its interactions with DNA metabolizing enzymes

Single-stranded DNA-binding protein is a eubacterial hub protein involved in DNA metabolism, which interacts with more than sixteen proteins in *E. coli*. SSB consists of an N-terminal globular domain and a C-terminal intrinsically disordered protein (IDP) region responsible for interactions with SSB-interacting proteins (SIPs). There are numerous PXXP motifs along the C-terminal IDP region. These PXXP motifs bind to the oligonucleotide-binding (OB) folds (SH3-domains) on either another SSB protein or other SIP proteins, from which several of the latter contain OB-folds. The binding of the SSB IDP region to the OB-fold of another SSB protein plays an important role in the cooperative binding to single-stranded DNA (ssDNA).[148] In a ternary complex of SSB-ssDNA-SIP, there is a competition for OB fold binding between the IDP region and ssDNA, which adds another layer of regulation to the SSB interactome.[149-151] The octapeptidic C-terminus (wtSSB-Ct, Asp1 – Phe2 – Asp3 – Asp4 – Asp5 – Ile6 – Pro7 – Phe8) is an essential recognition segment in the binding.[152] The octapeptide comprises a proximal (DFDD) and distal (DIPF) segment. SSB-Ct is a highly conserved sequence among all eubacteria, where the C-terminal Ile-Pro-Phe (IPF) motif is obligatory for the function. [148-150]

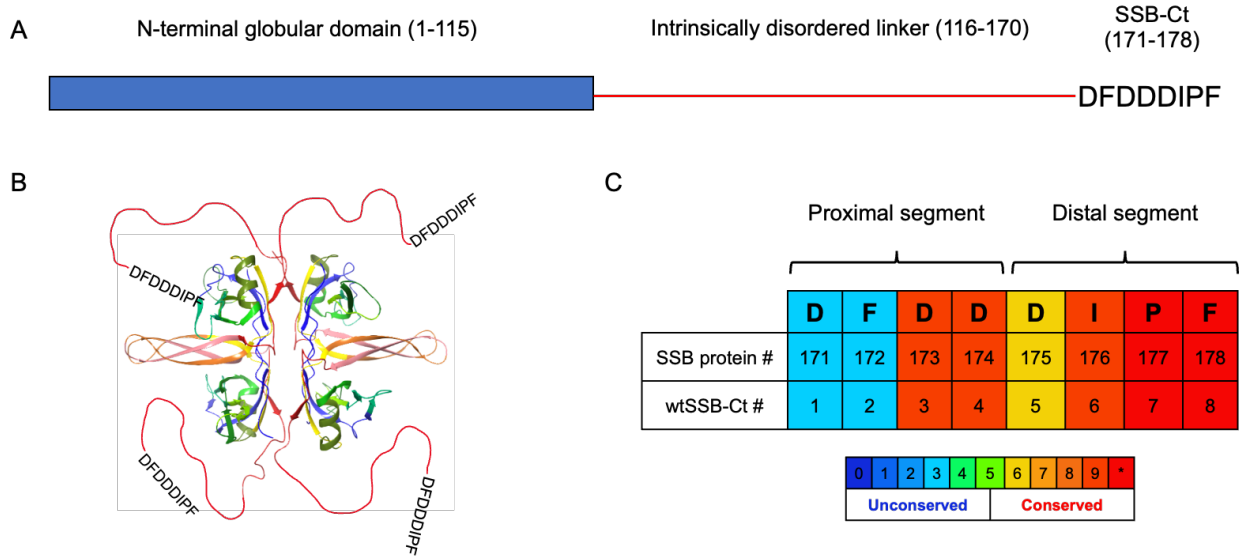

Figure 12. (A) Domain structure of E. coli SSB protein. (B) Structure of E. coli single-stranded-DNA binding protein as a tetramer (PDB structure 4MZ9). The IDP regions are shown as red lines, and acidic tips are shown as amino acid sequences. (C) Residues and numbering of SSB-Ct.

There are numerous crystal structures reported containing SSB-Ct in complex with any SIPs. Interestingly, only the DIPF or IPF motif of SSB-Ct is resolved, where the proximal segment of

the peptide is not visible. Despite its fuzzy nature, the truncation of the proximal segment of the peptide reduces affinity and SSB-stimulated ExoI activation, indicating the importance of the proximal residues in SSB-Ct interactions.[153] Deletion/mutation studies clarified the biological importance of SSB-Ct. Still, the exact role of the proximal segment remains puzzling, especially when we consider the low conservation scores in this segment.[152]

SIPs are involved in a wide variety of pathways of DNA metabolic processes. ExoI shows exonucleolytic activity in the 3' to 5' direction, degrading single-stranded DNA in bacteria.[154] SSB stimulates its activity fourfold.[148] RecO is a recombination mediator protein, part of the RecFOR pathway in bacteria.[155] These proteins interact with SSB physically in vitro through the SSB-Ct.[148,155]

SSB-interacting proteins are potential targets of antibiotic drug discovery.[156-158] Chances of resistance are lower since one point mutation has a lesser effect on the diffuse interface.

Small molecule mimetics of the IPF motif have been reported to inhibit the interactions of SSB with ExoI,[158,159] DnaG,[157] PriA.[156,160] These efforts led to inhibitors with inhibitory concentrations (IC50) at around 1 μM or higher values, which leaves room for improvement. The low affinities accord with the limited and solvent-exposed contact surface available for small molecules on the SIP proteins (Figure 20).[130,161,162] For ExoI, the crystal structures[158] showed a secondary binding site for the SSB-Ct at the SH3 domain, which does not play a primary role in the enzyme activation (Figure 20). However, this site offers a possibility for an additional hot spot to stabilize the binding of a peptidomimetic ligand.

We hypothesized that certain SIPs display surface features near the confirmed IPF binding sites, accessible by the proximal DFDD motif. We set out to test the role of these secondary hot spots for SIPs ExoI and RecO that are highly conserved across eubacterial species. Our data revealed that SSB-Ct could simultaneously land DFDD and DIPF motifs onto SIPs, and the binding involves a separate hot spot beside the confirmed IPF-binding site. Based on this structural model, we aimed to chemically modify the residues, including both terminals, to explore the possibility of increasing affinity by simultaneously enhancing interactions with the two hot spots. Dissecting the interaction of SSB C-terminus peptide with ExoI and RecO using molecular dynamics simulations.

5.2.1. The role of the proximal segment of SSB in binding to its interacting partners

Preliminary X-ray analysis by Lu and Keck identified two hot spots on ExoI, both binding the C-terminal tail of the SSB-Ct peptide.[18] Site A is located at the border of exonuclease and the SH3-like domains, while site B is in the SH3-like domain (Figure 20A). Due to crystal packing effects

and the inherent flexibility of SSB-Ct, only four (site A) and three (site B) C-terminal residues were resolved (Figure 20B) despite the fact that full wtSSB-Ct peptides were used for crystallization studies in both cases. The same phenomenon appeared for RecO; only three residues of SSB-Ct were resolved.

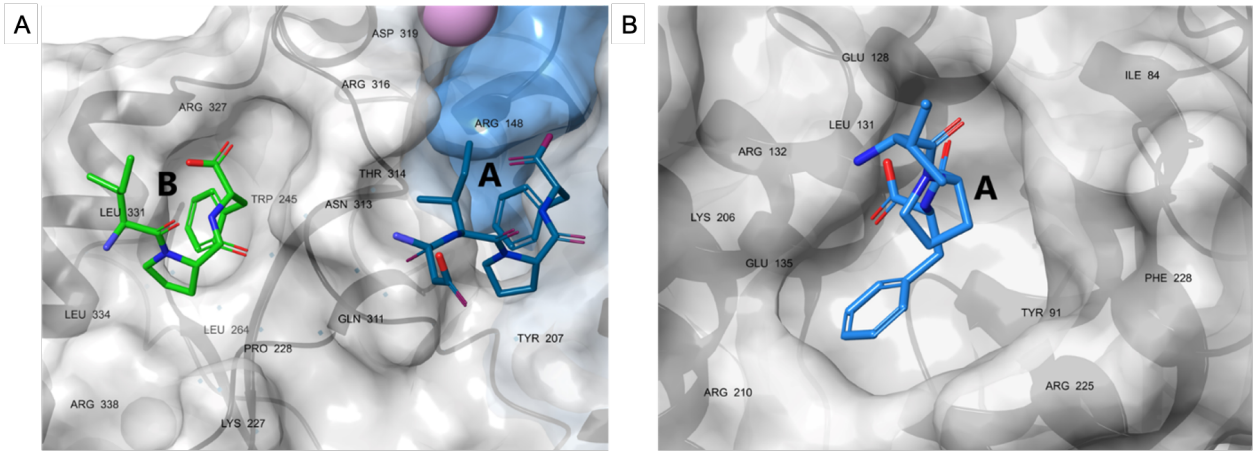

Figure 13. (A) The resolved C-terminal residues of SSB-Ct peptide (sticks) bound to sites A and B of ExoI and their critical interacting residues (PDB structure 3C94). (B) Resolved C-terminal residues of SSB-Ct peptide (sticks) bound to RecO (PDB structure 3Q8D). Interacting residues are labeled black. ExoI and RecO appear in surface representation.

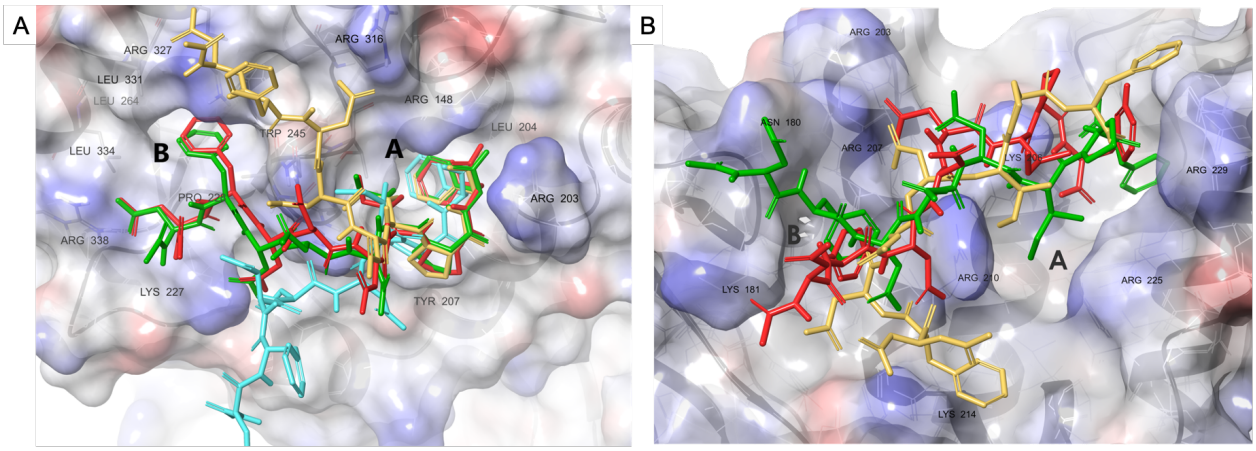

Figure 14. Representative structures of SSB-Ct – ExoI (A) complexes indicating the simultaneous binding of terminals at sites A and B. (B) RecO–SSB-Ct complex and representative structures. Proteins are shown in surface representation, and SSB-Ct peptide conformations appear as sticks.

For the structural characterization of the binding modes of the wild-type and the modified SSB-Ct peptides to the investigated enzymes, an extended conformational sampling method, the replica-exchange solute tempering (REST) molecular dynamics[163,164] was applied as implemented in the Desmond package.[165]

The representative structures of the MD simulation (Figure 21A) illustrate the ability of Phe2 of SSB-Ct to fit into the hydrophobic pocket of site B in ExoI. The modeling results reflect the non-uniform flexibility along the chain in the bound state. The Root Mean Square Fluctuation (RMSF) of the backbone atom coordinated by the C-terminal Phe8 has a value of ~1 Å showing a fixed position. Despite the potential transient interactions, the flexibility grows toward the proximal end, and the RMSF reaches the value of 9 Å (Figure 22).

Modeling results for RecO also revealed a secondary hot spot (site B) in the proximity of the IPF-binding pocket (site A) (Figure 21B). Site B is enclosed by a loop (residues 178-183) and a helix (residues 200-212). The helix structure separates sites A and B. Positively charged amino acids in the helix segment, especially Arg203, Lys206, and Arg210 interact with the negatively charged central segment of SSB-Ct. A computational alanine scan revealed the importance of Arg203 in peptide binding. Accordingly, we observed relatively low RMSF values for the central acidic segment, whereas modeling revealed higher residual flexibility for the terminal IPF and DFD motifs (Figure 22B).

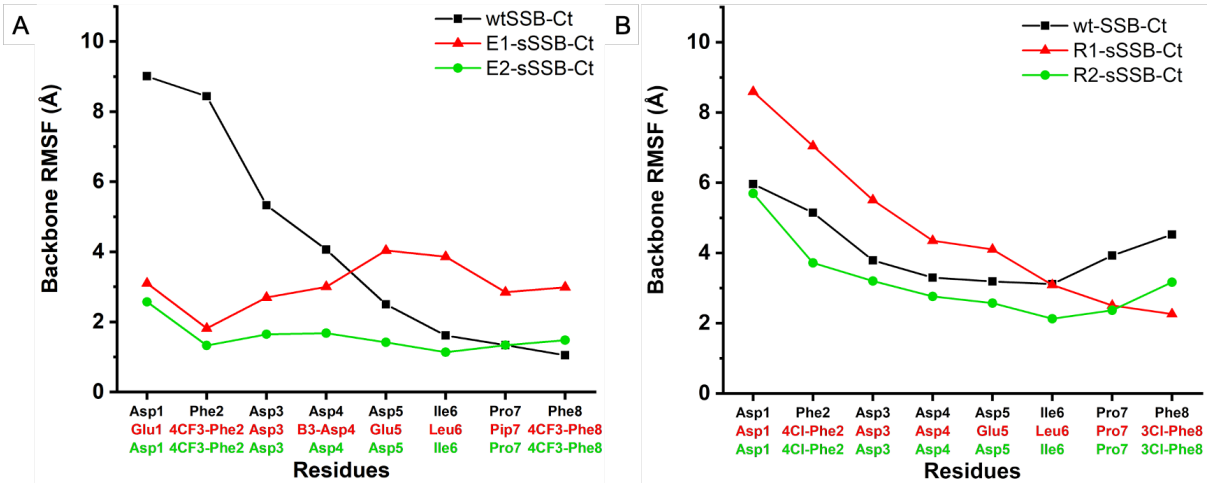

Figure 15. RMSF values of the backbone atom coordinates for wtSSB-Ct (black squares), E1-sSSB-Ct (red triangles), and E2-sSSB-Ct (green circles) in interaction with ExoI (A). RMSF values of the backbone atom coordinates displaying wtSSB-Ct (black squares), R1-sSSB-Ct (red triangles), and R2-sSSB-Ct (green circles) in interaction with RecO (B).

We have performed competitive fluorescence anisotropy assays to assess the importance of the proximal segment of wtSSB-Ct in binding. We synthesized half wtSSB-Ct peptides, namely DFDD and DIPF. The tracer sequence was the fluorescein-labeled wtSSB-Ct. First, we examined the ability of the wtSSB-Ct to compete with the tracer. For ExoI and RecO, the inhibition constants (IC_{50}) of the wtSSB-Ct were 4.54 ± 0.35 μM and 4.66 ± 0.25 μM, respectively. In the case of ExoI,

IC_{50} values were 311.80 ± 16.7 μM and 101.94 ± 6.61 μM for DFDD and DIPF peptides, respectively. RecO competitive titrations revealed that the DFDD peptide does not bind in conditions used here (IC_{50} is higher than 3 mM), while the DIPF peptide has an IC_{50} value of 44.84 ± 4.53 μM.

We concluded from the simulations that the proximal DFD motif can have transient stabilizing contacts with the secondary hot spots on the SIPs studied. Competition studies with the half peptides revealed the importance of the proximal segment in the binding of wtSSB-Ct.

5.2.2. Synthesis and Screening of the single mutant SSB-Ct library.

In the next step, we scanned if chemical modifications of the wild-type SSB-Ct (wtSSB-Ct) can improve affinity and generate a single mutant wtSSB-Ct (mSSB-Ct) peptide library (Figure 23A and B). The importance of the side chain chemistry was investigated using non-natural amino acids having similar physicochemical characteristics and homologous replacements (Figure 23C). Potential effects of sidechain stereochemistry and backbone homology were also addressed using D-enantiomer- and backbone homologation scans. N-methylated amino acids were included in all positions to probe the H-bonding capabilities of the backbone amides.

We synthesized 51 mSSB-Ct sequences and screened them in a fluorescence anisotropy-based competition assay on ExoI and RecO. The library members were evaluated against the IC_{50} values obtained for the wtSSB-Ct (Figures 23A and B).

The best 12 hits were re-synthesized with a fluorescein label and titrated directly to validate the competition results (Table 4). Tolerated or non-significant changes are shown in light pink (IC_{50} = 2.5 – 6.0 μM). Moderate or significant increases in IC_{50} values (6.0 – 8.0 μM and > 8.0 μM) are shown in light and dark blue, respectively. Modifications in grey color were not evaluated due to compound impurity. C) wtSSB-Ct sequence, conservation score, and applied modifications for each occurring amino acid in the wtSSB-Ct sequence.[9]

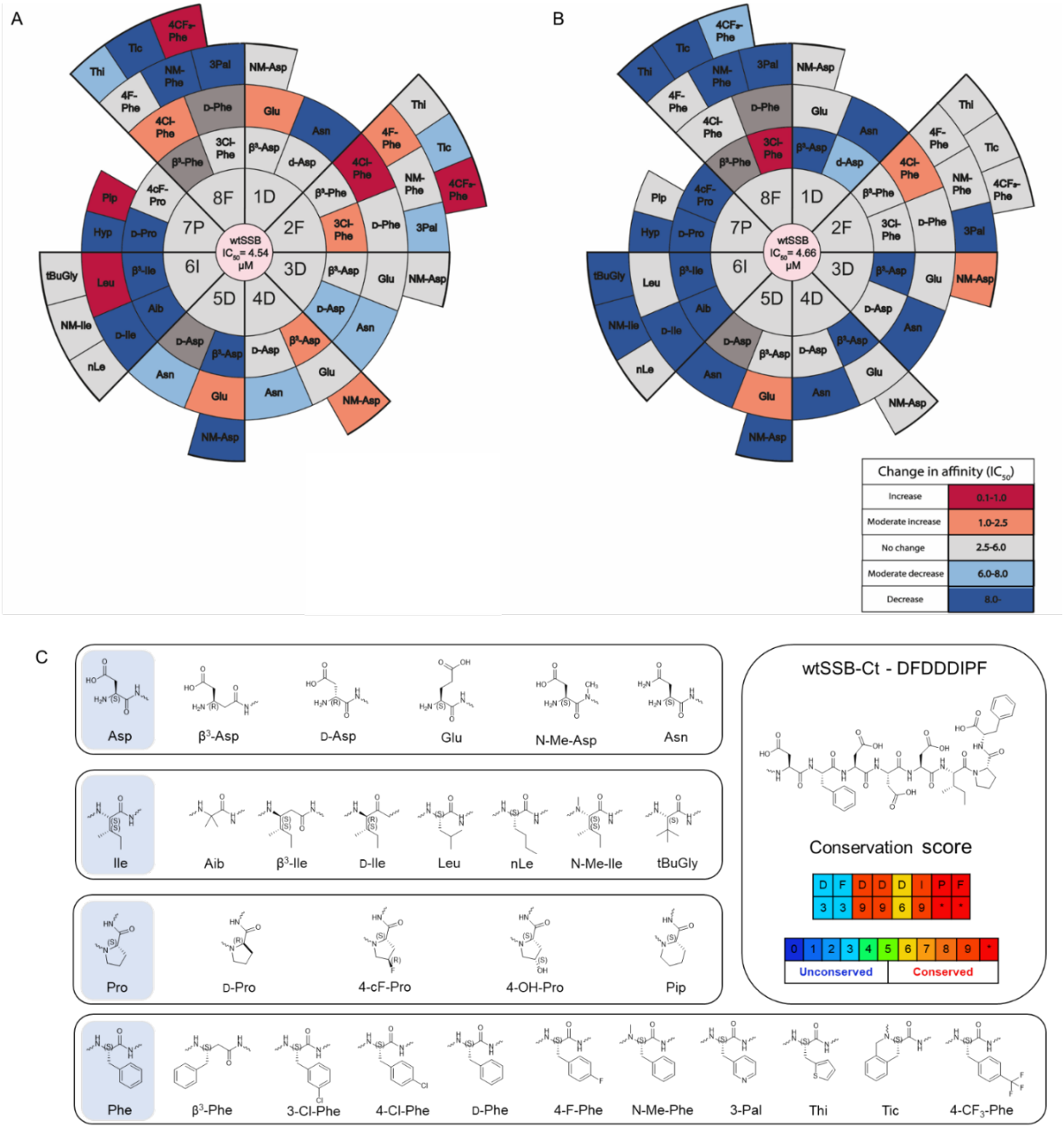

Figure 16. Figure 5. Screening and design of the mSSB-Ct library. Radial heat map showing competitive fluorescence anisotropy screening data of mSSBs on (**A**) ExoI and (**B**) RecO. IC_{50} values lower than 1 µM are highlighted in red. Moderate decreases in IC_{50} (1.0–2.5 µM) are shown in orange. Tolerated or non-significant changes are shown in light pink (IC_{50} = 2.5–6.0 µM). Moderate or significant increases in IC_{50} values (6.0–8.0 µM and >8.0 µM) are shown in light and dark blue, respectively. Modifications in dark grey were not evaluated due to compound impurity. (**C**) wtSSB-Ct sequence, conservation score, and applied modifications for each occurring amino acid in the wtSSB-Ct sequence

Table 4. Results of direct fluorescence anisotropy titrations with ExoI and RecO. IC50 values are shown for comparison.

mSSB	IC_{50} on ExoI (μM)	EC_{50} on ExoI (μM)	IC_{50} on RecO (μM)	EC_{50} on RecO (μM)
F-wtSSB	4.54 ± 0.34	0.35 ± 0.04	4.66± 0.24	0.37 ± 0.03
F-4Cl-Phe2	0.56 ± 0.03	0.15 ± 0.02	0.41 ± 0.27	0.26 ± 0.01
F-4F-Phe2	1.93 ± 0.81	0.36 ± 0.03	0.52 ± 0.28	0.23 ± 0.02
F-4CF3-Phe2	0.33 ± 0.02	0.04 ± 0.002	0.56 ± 0.42	0.18 ± 0.03
F-NM-Asp3	3.88 ± 0.13	0.86 ± 0.11	1.88 ± 0.09	0.27 ± 0.04
F-Glu5	1.49 ± 0.10	0.11 ± 0.01	1.41 ± 0.55	0.20 ± 0.02
F-Leu6	0.92 ± 0.06	0.20 ± 0.03	3.23 ± 0.16	0.31 ± 0.03
F-nLe6	2.68 ± 0.54	0.20 ± 0.01	9.36 ± 0.08	0.20 ± 0.01
F-Pip7	0.89 ± 0.14	0.09 ± 0.01	3.31 ± 0.44	0.25 ± 0.03
F-3Cl-Phe8	3.40 ± 2.41	0.25 ± 0.02	0.36 ± 0.04	0.22 ± 0.03
F-4CF3-Phe8	0.39 ± 0.07	0.09 ± 0.01	5.97 ± 2.43	0.26 ± 0.02

5.2.3. Modifications both in the proximal DFDD and the distal DIPF segments increase affinity

For ExoI, Asp to Glu mutation yielded a moderate increase in affinity at position 1, whereas no improvement was obtained at position 3. β^3-Asp and NM-Asp substitutions yielded moderately higher affinities at position 4. Significantly higher affinity was achieved at position 2 with residues 4-Cl-Phe and 4-CF_3-Phe. Other halogen-substituted phenylalanine modifications (3-Cl-Phe and 4-F-Phe) caused a moderate decrease in IC_{50}. At position 5, Asp to Glu mutation caused a moderate improvement. Isoleucine to leucine mutation at position 6 significantly increased affinity to ExoI (IC_{50} = 0.92 μM). Out of the several proline analogs, 4-F-proline was tolerated by ExoI, and 2-aminopipecolic acid increased affinity (IC_{50} = 0.89 μM). Halogen-substituted phenylalanine modifications at position 8 are well tolerated (3Cl-Phe and 4F-Phe) or increase the affinity. A moderate increase was observed for 4-Cl-Phe, while 4-CF_3-Phe in position 8 yielded an IC_{50} of 0.09 μM.

For RecO, positions 1 and 4 did not allow improvement. On the contrary, a moderate increase in affinity was obtained with 4-Cl-Phe and NM-Asp in positions 2 and 3, respectively. Asp to Glu mutation was moderately advantageous at position 5. Leucine and norleucine were tolerated by RecO in position 6, with IC50 values of 3.22 μM and 2.55 μM, respectively. Modifications at position 7 did not result in improved affinity. Except for Pip, all replacements were detrimental to

the binding. In stark contrast, Phe8 to 3-chlorophenyl alanine significantly decreased the IC50 value to 0.84 µM, whereas 4-F-Phe and 4-Cl-Phe were tolerated.

These results support our structural model that Phe2 in the proximal DFDD motif can be an extra anchor point for ExoI and RecO. Appropriate chemical modifications in the side chain of Phe2 increase affinity to the targets even if the native binding mode of wtSSB-Ct does not directly depend on the presence of Phe2. The behavior of the two proteins tested is not uniform in interacting with Phe2. This observation suggests that this variable feature explains the lower conservation score for Phe2.[9] Chemical mutation data on the essential DIPF segment revealed that binding sites have a certain level of residual flexibility that can adapt to the substituted residues so that the affinity increases.

5.2.4. Combined modifications yield high-affinity ligands.

In the next step, we tested if the combination of the favorable target-specific single mutations could increase the affinity any further. For ExoI, we synthesized E1-sSSB-Ct (Glu – 4-CF_3-Phe – Asp – β^3-Asp – Glu – Leu – Pip – 4-CF_3-Phe) and E2-sSSB-Ct (Asp – 4-CF_3-Phe – Asp – Asp – Asp – Ile – Pro – 4-CF_3-Phe, Figure 24). These sequences were probed against ExoI, and binding characteristics were measured by competitive fluorescent anisotropy and isothermal titration calorimetry.

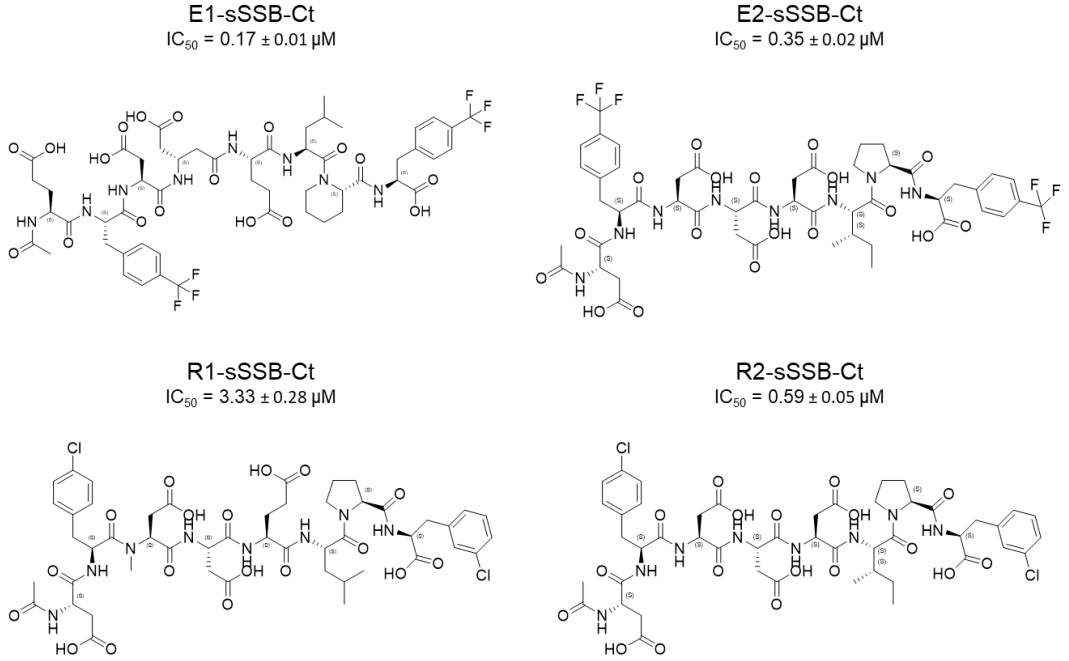

Figure 17. Combined modifications for SSB-Ct. E1-sSSB-Ct and E2-sSSB-Ct were tested on Exol. R1-sSSB-Ct and R2-sSSB-Ct were tested on RecO. IC_{50} values show the ability of the sequences to compete with F-wtSSB-Ct in competitive fluorescent anisotropy assay.

For ExoI, the combined modifications incorporated into E1-sSSB-Ct and E2-sSSB-Ct yielded further improvement; the IC_{50} value decreased to 0.17 μM and 0.35 μM, respectively. However, the effects of the mutations are not fully additive. In the framework of a two-anchor point binding model, the enthalpically stabilized contacts inevitably decrease the residual flexibility of the protein-ligand complex. To test the enthalpy-entropy compensation, we carried out isothermal titration calorimetric measurements and determined the stoichiometry and the thermodynamic parameters of the interactions (Figure 25 and Table 4). We found the stoichiometry of 1:1, supporting that the secondary binding sites do not bind an additional ligand under the conditions applied. In accord with the literature results, the wtSSB-Ct sequence without the fluorescent tag has a higher affinity to RecO (K_D = 0.14 μM) than to ExoI (K_D = 3.07 μM).[35] E1-sSSB-Ct and E2-sSSB-Ct displayed low nanomolar dissociation constants (K_D = 19 nM and 40 nM, respectively) with 1:1 stoichiometry. The main enthalpy gain can be contributed to the 4-CF_3-Phe–Phe replacements at positions 2 and 8, supporting the two anchor point scenario. However, the additional modifications (E1-sSSB-Ct) could only improve the binding affinity slightly. The binding enthalpy exhibited a marked increase for both peptides. Having additional modifications, peptide E1-sSSB-Ct compared to E2-sSSB-Ct shows an increase in enthalpy and a decrease in entropy, possibly due to the peptide forming more favorable contacts with the protein and losing flexibility in the process, hence the entropy loss. In contrast, we observed a binding entropy decrease relative to the native sequence in line with the enthalpy-entropy compensation effect. This finding strongly supports the enhanced proximal and distal anchor point scenario. To gain further support, we carried out 19F-NMR measurements with four peptides containing 4-CF_3-Phe residues; 4-CF_3-Phe2, 4-CF_3-Phe8, E1-sSSB-Ct and E2-sSSB-Ct. Fluor signals corresponding to both the proximal and distal segments disappeared upon adding Exol in a 1:1 ratio. This finding confirms the stabilization of the peptide – Exol interaction at both anchor points.

For RecO, R1-sSSB-Ct (Asp – 4Cl-Phe – NM-Asp – Asp – Glu – Leu – Pro – 3Cl-Phe) was synthesized first (Figure 24). R1-sSSB-Ct contains all favorable modifications, but the effects were non-additive. R1-sSSB-Ct failed to compete with wtSSB-Ct in the expected manner (IC_{50} = 3.33 μM). ITC titration of R1-sSSB-Ct showed no significant changes in the thermodynamic profile of binding (Figure 22D), resulting in a ΔG value similar to the wtSSB-Ct (Figure 25C-D). Therefore, we synthesized R2-sSSB-Ct (Asp – 4Cl-Phe – Asp – Asp – Asp – Ile – Pro – 3Cl-Phe), which contains modifications only at the terminal Phe residues (Figure 24). This peptide competed with wtSSB-Ct, having an IC_{50} value of 0.59 μM. ITC results confirmed the affinity increase (K_D = 35

nM) due to a marked increase in the enthalpic stabilization was damped by an enthalpy-entropy compensation effect (Figure 25F**).**

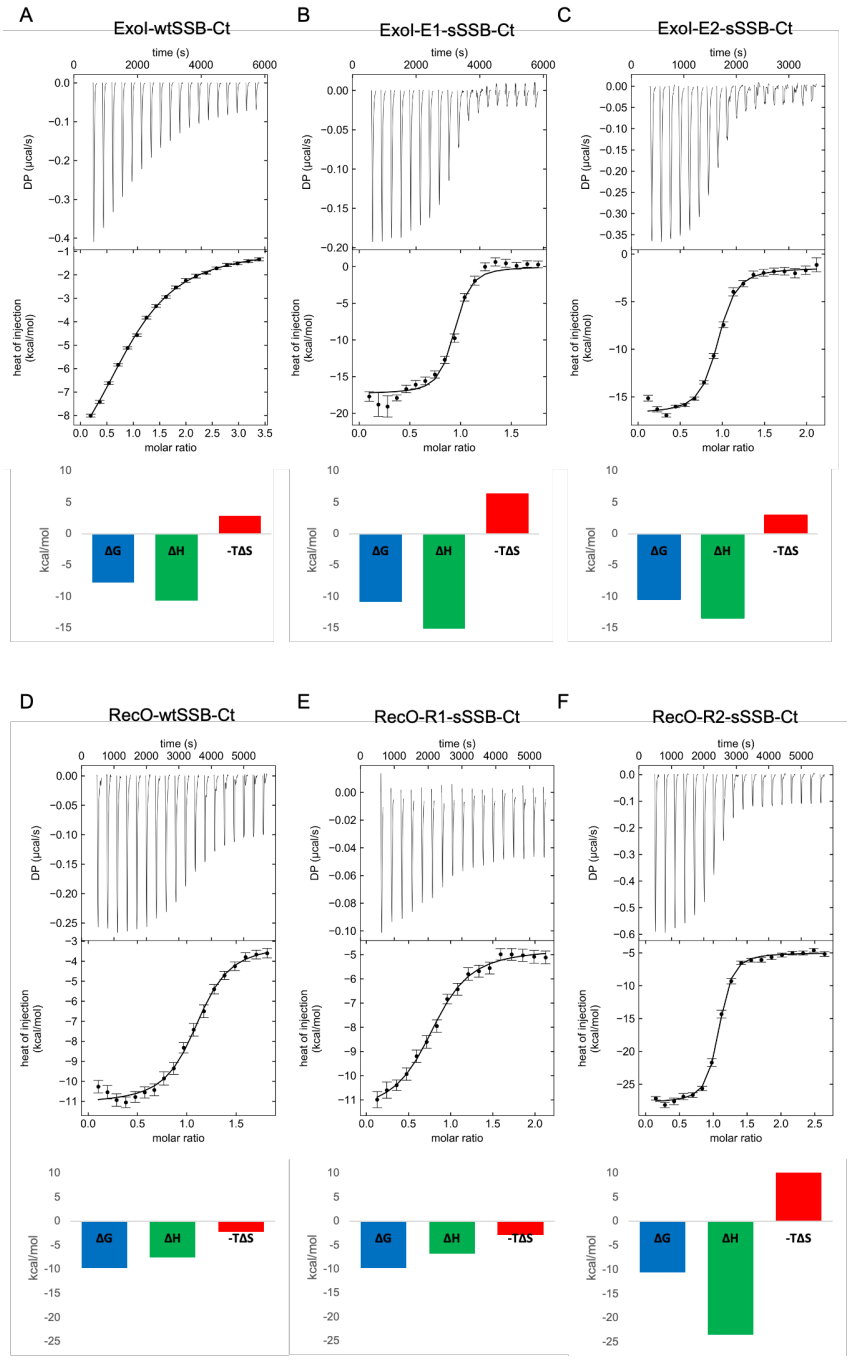

Figure 18. . Thermodynamic binding profiles to ExoI and RecO for the wild-type and modified SSB-Ct sequences. ITC data for ExoI was titrated with wtSSB-Ct (**A**), E1-sSSB-Ct (**B**), and E2-sSSB-Ct (**C**). ITC data for RecO was titrated with wtSSB-Ct (**D**), R1-sSSB-Ct (**E**), and R2-sSSB-Ct (**F**). The reconstructed thermograms displayed were obtained after global peak-shape analysis and singular value decomposition regularization implemented in the NITPIC program.

Table 5. – ITC binding data for the interactions of SSB-Ct variants with ExoI and RecO. Confidence intervals are shown for K_D and ΔH values.

	ExoI			RecO		
	wtSSB	**E1-sSSB**	**E2-sSSB**	**wtSSB**	**R1-sSSB**	**R2-sSSB**
N	0.9593	1.0000	0.8846	1.0032	0.9353	0.9999
K_D (μM)	3.07 (2.62–3.64)	0.02 (0.01–0.03)	0.04 (0.02-0.07)	0.14 (0.08–0.24)	0.13 (0.08–0.22)	0.04 (0.02–0.05)
ΔG (kcal/mol)	-7.78	-10.89	-10.43	-9.77	-9.71	-10.54
ΔH (kcal/mol)	-10.67 (-11.64– -9.97)	-17.35 (-18.34– -16.43)	-13.45 (-14.46– -12.49)	-7.54 (-8.28– -6.90)	-6.79 (-7.84– -6.05)	-23.48 (-24.34– -22.73)
-TΔS (kcal/mol)	2.90	6.46	3.02	-2.28	-2.91	12.32

We have investigated the ability of sSSB-Ct peptides to bind their targets in a complex environment. ExoI and RecO overexpressing BL21 $(DE)_3$ pLysS cells were diluted with non-overexpressing cells to have a 1 μM final concentration of recombinant protein. Pulldown assay was performed using cell lysate incubated with wtSSB-Ct, E1-sSSB-Ct, and R1-sSSB-Ct immobilized on streptavidin beads. Washed beads were digested, and tryptic peptides were detected using HPLC-MS. We detected twofold enrichment of ExoI on E1-sSSB and a tenfold enrichment of RecO on R1-sSSB compared to wtSSB-Ct (data not shown here).

5.2.5. Molecular dynamics simulations provide insight into the binding modes of E-sSSB-Ct and R2-sSSB-Ct.

We performed replica-exchange solute tempering simulations to test the changes in the residual flexibility of E-sSSB-Ct and R2-sSSB-Ct. In both cases, simulations yielded reduced residual flexibility in the bound state, according to the experimental findings (Figure 26). E-sSSB-Ct was anchored to site B through $4CF_3$-Phe2 and formed stable contact with site A through $4CF_3$-Phe8. This improved stability led to the root mean square fluctuation (RMSF) values of 2 and 3 Å at the terminals. R2-sSSB-Ct was connected to sites A and B through 4Cl-Phe2 and 3Cl-Phe8, respectively, with RMSF values around 4 and 2.5 Å, respectively. For RecO, similar to the wtSSB-Ct, the central segment (residues Asp4-Ile6) of R2-sSSB-Ct was the least flexible part compared with the terminals, and the RMSF values displayed an overall downward shift (Figure 22D). We

observed enhanced interactions for 3Cl-Phe8 with Tyr91 and 4Cl-Phe2 with Arg203 compared with wtSSB-ct.

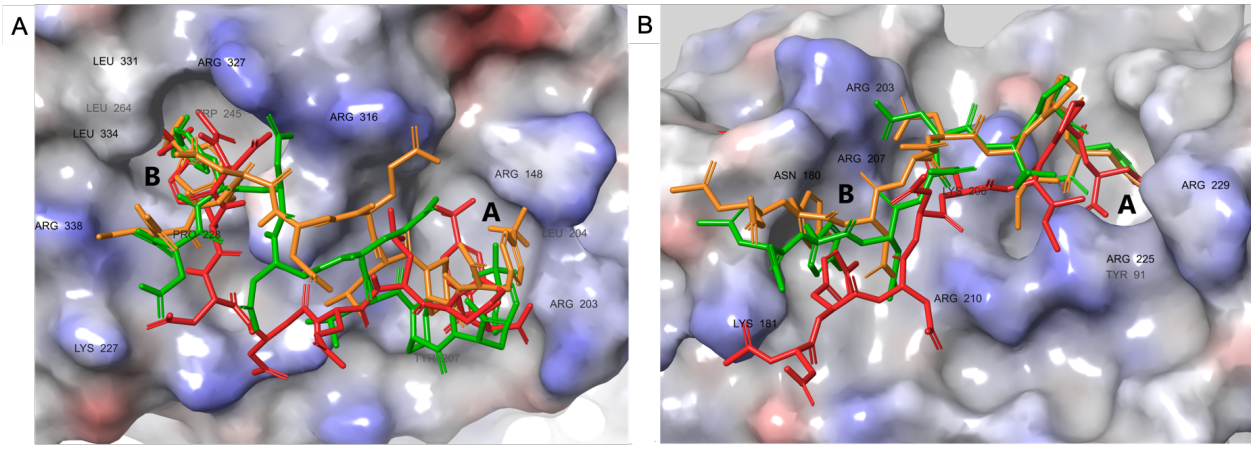

Figure 19. Representative structures for the complexes ExoI – E-sSSB-Ct (A) and RecO – R2-sSSB-Ct (B). RMSF values of backbone atom coordinates for ExoI (C) and RecO (D) displaying wtSSB-Ct (black squares), E-sSSB-Ct, and R2-sSSB-Ct (red triangles).

In conclusion, modifying the conserved Phe residues in SSB-Ct sequence increases the affinity of the peptide toward two SSB-interacting proteins. Owing to the modifications, the flexibility of the modified peptides could be decreased, supporting the two hot spot theory.

6. Conclusion

There is an ever-increasing demand for protein-protein inhibitors since many untreated diseases involve protein-protein interactions in their pathogenesis. Foldameric H14 helices consisting of β-amino acids fold into a helical shape that is stable in water and exhibits a large enough surface area (cca. 500 $Å^2$) to interact with other large molecules such as other foldameric helices or proteins.

Using H14 helices, we have created libraries utilizing the ability of these molecules to interact with protein surfaces. In addition to the H14 library, we have constructed another library consisting of H12 helices. We recorded the binding patterns of 512 LSM probes on five different proteins and analyzed the resulting interactions. LSM probes were particularly effective because of their ability to adapt to the local environment of protein hot spots. We found that the H14 library had a high number of hits; their binding was in correlation with the number of interacting partners of the proteins screened. The binding was driven by two proteinogenic side chains, and the scaffold served a shielding and templating function. Using the natural binding partners of the proteins in a competitive pull-down assay, we have shown that the H14 LSM probes can detect both orthosteric and non-orthosteric binding sites. Side chain preferences reflect the side chain chemistry of natural binding partners for each protein. Foldamer-protein side chain enrichments are highly similar to naturally occurring protein-protein interface enrichments. This work enables bottom-up large molecule drug discovery scenarios by linking surface fragments. It is a step toward using self-organizing protein mimetic LSM probes for molecular recognition, similar to natural antibodies.

Many top-down designed PPI inhibitors have been reported in the literature, using well-defined structural motifs (helices, sheets, turns) as starting points. In contrast, interfaces involving intrinsically disordered proteins (IDPs) are less understood. IDPs are involved in many diseases and are important targets for drug development. Single-stranded DNA-binding protein C-terminal octapeptide is the main site of physical interaction with ExoI and RecO. We have further clarified the importance of the octapeptide's proximal segment, despite its low conservation score and fuzzy binding to SIPs, using replica exchange solute tempering simulations and fluorescence anisotropy assay using the two halves of the octapeptide. We chemically modified the wild-type SSB-Ct peptide in search of affinity-improving modifications. The combination of advantageous modifications created high-affinity binders for ExoI and RecO, respectively. The thermodynamic profile of these systems revealed an enthalpy-stabilized binding of designer peptides with a marked enthalpy-entropy compensation. As an important effect of modifications, the stabilization of the protein-peptide complexes was assessed using replica exchange solute tempering simulations,

showing lower conformational flexibility of the proximal segment of the octapeptide. Our model strongly supports a two-hot spot binding model of SSB-Ct octapeptide to interacting partners. The importance of halogenating phenylalanine residues in the peptide was highlighted. These results may pave the way for effective SSB mimetic antibiotics design.

7. Summary

1. Foldameric H14 helices can act as local surface mimetics (LSMs); they are able to present side chains in a surface-compatible manner, and they can distinguish proteins based on their functions.
   1.1. Five proteins were tested with diverse functional features using H12 and H14 LSMs.
   1.2. Pulldown assays performed in the presence of natural ligands revealed that LSMs detect orthosteric and non-orthosteric hot spots.
   1.3. We have performed a database search to assess the interactome size for each protein tested here. We have found a correlation between interactome size and the number of bound foldamers.
   1.4. Protein similarities were calculated between the affinity pattern of each protein as covariances. We could conclude that LSMs distinguish proteins with different functions. Furthermore, we have found significant similarities between LSM side-chain enrichment data and general protein-protein interface side-chain enrichment data reported in the literature.
2. Designer SSB-Ct variants were synthesized that are nanomolar binders of their targets (ExoI and RecO). Molecular modeling studies show the stabilization of the proximal segment of modified peptides when in complex with their targets.
   2.1. Replica exchange solute tempering simulations and fluorescence anisotropy assays were performed on SSB-Ct and its interacting partners (ExoI and RecO) to demonstrate that the proximal, non-conserved segment is also important in the binding of SSB-Ct.
   2.2. Single mutant SSB-Ct variants were designed and synthesized using several non-natural amino acids to assess the structure-activity relationship of SSB-Ct. Peptides were screened on both proteins in a competitive fluorescence anisotropy assay. We could identify affinity-increasing modifications, proving that even conserved residues could be modified.
   2.3. Affinity-increasing modifications were combined into one peptide for each protein. The binding and competition of these peptides were assessed using isothermal titration calorimetry and fluorescence anisotropy. As an effect of combined modifications, the binding affinity of peptides was increased, and enthalpy-driven binding was determined with a marked enthalpy-entropy compensation.

2.4. The stabilization of the modified peptides bound to their targets was analyzed using replica exchange solute tempering simulations. The stabilization of the designer peptides is presented in the lower residual RMSD values compared to the wild-type SSB peptide.

8. Acknowledgments

I would like to express my sincere gratitude to my Ph.D. supervisors, Prof. Tamas Martinek and Gerda Szakonyi, for their invaluable guidance, support, and expertise throughout my doctoral journey. Their encouragement, patience, and feedback have been instrumental in shaping my research and academic development. I am truly grateful for their mentorship.

Although they are not official mentors, I look up to Éva Karolina Bartus, Edit Wéber, and Zsófia Hegedűs as if they were my mentors. I would like to thank them for always being there when I needed scientific guidance.

I would like to extend my gratitude to all the individuals who fueled my love for science and research. First, I would like to thank Professor Ferenc Fülöp and Professor Enikő Forró for letting me start my academic journey in the Department of Pharmaceutical Chemistry.

My sincere gratitude goes to Diana Szucs for being my mentor when I was an undergraduate researcher in the Department of Pharmaceutical Analysis and keep helping and supporting me until and after the end of my Ph.D. studies. Kudos to Professor György Dombi, who has always been a great inspiration and good friend to me. Thanks to Erika Kunosné Tóth for being the best lab tech and helping me with my molecular biology experiments. I want to thank all the help with molecular biology to Katalin Jósvay.

I would like to thank my colleagues at the Department of Medical Chemistry for their co-operations, especially Ferenc Bogár and Gábor Paragi, for performing molecular dynamics simulations. My gratitude goes to Gábor Kecskeméti and Zoltán Szabó for performing proteomics studies.

I would like to thank my former and present colleagues for being my friends and accompanying me when I needed to let off steam conjured in the enigmatic engine of academic research. A list of these special people in chronological order: Ádám Georgiádes, Helga Fekete, Norbert Imre, Brigitta Csipak, Beáta Mag, Brigitta Bodnár, Zita Ibolya Papp, Arijit Sakar, and Gaurav Sharma.

I would also like to thank my family for their unwavering love and support. Their sacrifices, understanding, and encouragement have been critical to my success. I am particularly grateful to the most important people in my life, my grandma, my mom, and my sister, for their constant encouragement and belief in me.

My friends have also been a significant source of support, and I am grateful for their friendship and encouragement. They have kept me sane and provided a much-needed distraction from academic pressures. I am particularly grateful to János Gera for his unwavering support and for being there

for me in good times and bad. I am grateful to my childhood friends Andor Okos Rigó, Attila Kovács, Dániel Benedek, Gábor Csernus-Lukács, and Attila Gránicz for not forgetting me while I was distracted collecting impact factors.

I want to thank my friend, Kaushik, for keeping me company and introducing me to Indian cuisine: my most important ATP source ever since he taught me how to cook.

Finally, I am grateful to my students: teaching has always been a huge inspiration and source of motivation for me. Special thank you to Vencel Petrovicz and Martin Cseh; seeing my student grow and become colleagues has been wonderful.

Once again, I extend my heartfelt thanks to everyone who has contributed to my academic and personal growth.

9. References

1 Gao, M. & Skolnick, J. Structural space of protein–protein interfaces is degenerate, close to complete, and highly connected. **107**, 22517-22522, doi:doi:10.1073/pnas.1012820107 (2010).

2 Otto, S., Furlan, R. L. E. & Sanders, J. K. M. Dynamic combinatorial chemistry. *Drug Discovery Today* **7**, 117-125, doi:https://doi.org/10.1016/S1359-6446(01)02086-4 (2002).

3 Yan, C., Wu, F., Jernigan, R. L., Dobbs, D. & Honavar, V. Characterization of protein-protein interfaces. *The protein journal* **27**, 59-70, doi:10.1007/s10930-007-9108-x (2008).

4 Schilder, J. & Ubbink, M. Formation of transient protein complexes. *Current opinion in structural biology* **23**, 911-918, doi:10.1016/j.sbi.2013.07.009 (2013).

5 De Las Rivas, J. & Fontanillo, C. Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS computational biology* **6**, e1000807, doi:10.1371/journal.pcbi.1000807 (2010).

6 Goodsell, D. S. & Olson, A. J. Structural symmetry and protein function. *Annual review of biophysics and biomolecular structure* **29**, 105-153, doi:10.1146/annurev.biophys.29.1.105 (2000).

7 Mei, G., Di Venere, A., Rosato, N. & Finazzi-Agrò, A. The importance of being dimeric. *The FEBS journal* **272**, 16-27, doi:10.1111/j.1432-1033.2004.04407.x (2005).

8 Fischer, E. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft* **27**, 2985-2993, doi:https://doi.org/10.1002/cber.18940270364 (1894).

9 Massova, I. & Kollman, P. A. Computational Alanine Scanning To Probe Protein−Protein Interactions: A Novel Approach To Evaluate Binding Free Energies. *Journal of the American Chemical Society* **121**, 8133-8143, doi:10.1021/ja990935j (1999).

10 Koshland, D. E. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **44**, 98-104, doi:10.1073/pnas.44.2.98 (1958).

11 Creighton, T. E. *Proteins: Structures and Molecular Properties*. (W. H. Freeman, 1993).

12 Phillip, Y. *et al.* Contrasting factors on the kinetic path to protein complex formation diminish the effects of crowding agents. *Biophysical journal* **103**, 1011-1019, doi:10.1016/j.bpj.2012.08.009 (2012).

13 Volkov, A. N., Bashir, Q., Worrall, J. A., Ullmann, G. M. & Ubbink, M. Shifting the equilibrium between the encounter state and the specific form of a protein complex by interfacial point mutations. *Journal of the American Chemical Society* **132**, 11487-11495, doi:10.1021/ja100867c (2010).

14 Chakrabarti, P. & Janin, J. Dissecting protein-protein recognition sites. *Proteins* **47**, 334-343, doi:10.1002/prot.10085 (2002).

15 Zerbe, B. S., Hall, D. R., Vajda, S., Whitty, A. & Kozakov, D. Relationship between hot spot residues and ligand binding hot spots in protein-protein interfaces. *Journal of chemical information and modeling* **52**, 2236-2244, doi:10.1021/ci300175u (2012).

16 Dyson, H. J. Making Sense of Intrinsically Disordered Proteins. *Biophysical journal* **110**, 1013-1016, doi:10.1016/j.bpj.2016.01.030 (2016).

17 Lehne, B. & Schlitt, T. Protein-protein interaction databases: keeping up with growing interactomes. *Human Genomics* **3**, 291, doi:10.1186/1479-7364-3-3-291 (2009).

18 Fields, S. & Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246, doi:10.1038/340245a0 (1989).

19 Rigaut, G. *et al.* A generic protein purification method for protein complex characterization and proteome exploration. *Nature biotechnology* **17**, 1030-1032, doi:10.1038/13732 (1999).

20 Modell, A. E., Blosser, S. L. & Arora, P. S. Systematic Targeting of Protein-Protein Interactions. *Trends in pharmacological sciences* **37**, 702-713, doi:10.1016/j.tips.2016.05.008 (2016).

21 Lokey, R. S. Forward chemical genetics: progress and obstacles on the path to a new pharmacopoeia. *Current opinion in chemical biology* **7**, 91-96, doi:10.1016/s1367-5931(02)00002-9 (2003).

22 Lopez-Girona, A. *et al.* Cereblon is a direct protein target for immunomodulatory and antiproliferative activities of lenalidomide and pomalidomide. *Leukemia* **26**, 2326-2335, doi:10.1038/leu.2012.119 (2012).

23 Hennemann, H., Wirths, S. & Carl, C. Cell-based peptide screening to access the undruggable target space. *European journal of medicinal chemistry* **94**, 489-496, doi:10.1016/j.ejmech.2014.10.038 (2015).

24 Kasper, A. C., Baker, J. B., Kim, H. & Hong, J. The end game of chemical genetics: target identification. *Future medicinal chemistry* **1**, 727-736, doi:10.4155/fmc.09.52 (2009).

25 Moffat, J. G., Rudolph, J. & Bailey, D. Phenotypic screening in cancer drug discovery - past, present and future. *Nat Rev Drug Discov* **13**, 588-602, doi:10.1038/nrd4366 (2014).

26 Tanrikulu, Y., Krüger, B. & Proschak, E. The holistic integration of virtual screening in drug discovery. *Drug discovery today* **18**, 358-364, doi:10.1016/j.drudis.2013.01.007 (2013).

27 Mullard, A. Pioneering apoptosis-targeted cancer drug poised for FDA approval. *Nature reviews. Drug discovery* **15**, 147-149, doi:10.1038/nrd.2016.23 (2016).

28 Rajamani, D., Thiel, S., Vajda, S. & Camacho, C. J. Anchor residues in protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 11287-11292, doi:10.1073/pnas.0401942101 (2004).

29 Rooklin, D., Wang, C., Katigbak, J., Arora, P. S. & Zhang, Y. AlphaSpace: Fragment-Centric Topographical Mapping To Target Protein-Protein Interaction Interfaces. *Journal of chemical information and modeling* **55**, 1585-1599, doi:10.1021/acs.jcim.5b00103 (2015).

30 Pelay-Gimeno, M., Glas, A., Koch, O. & Grossmann, T. N. Structure-Based Design of Inhibitors of Protein-Protein Interactions: Mimicking Peptide Binding Epitopes. *Angew Chem Int Ed Engl* **54**, 8896-8927, doi:10.1002/anie.201412070 (2015).

31 Guo, Z. *et al.* Identification of Protein–Ligand Binding Sites by the Level-Set Variational Implicit-Solvent Approach. *J. Chem. Theory Comput.* **11**, 753-765, doi:10.1021/ct500867u (2015).

32 Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research* **46**, W296-w303, doi:10.1093/nar/gky427 (2018).

33 Rodrigues, J. P. *et al.* Defining the limits of homology modeling in information-driven protein docking. *Proteins* **81**, 2119-2128, doi:10.1002/prot.24382 (2013).

34 Evans, D. A. *et al.* A General Method for the Synthesis of Enantiomerically Pure β-Substituted, β-Amino Acids through α-Substituted Succinic Acid Derivatives. *The Journal of Organic Chemistry* **64**, 6411-6417, doi:10.1021/jo990756k (1999).

35 Yin, Z. *et al.* Tailor-made amino acid-derived pharmaceuticals approved by the FDA in 2019. *Amino acids* **52**, 1227-1261, doi:10.1007/s00726-020-02887-4 (2020).

36 Cabri, W. *et al.* Therapeutic Peptides Targeting PPI in Clinical Development: Overview, Mechanism of Action and Perspectives. *Front Mol Biosci* **8**, 697586, doi:10.3389/fmolb.2021.697586 (2021).

37 Andersson, E. *et al.* D-cycloserine vs placebo as adjunct to cognitive behavioral therapy for obsessive-compulsive disorder and interaction with antidepressants: a randomized clinical trial. **72**, 659-667 (2015).

38 Patil, S. T. *et al.* Activation of mGlu2/3 receptors as a new approach to treat schizophrenia: a randomized Phase 2 clinical trial. *Nature medicine* **13**, 1102-1107, doi:10.1038/nm1632 (2007).

39 Alexandre, F.-R. *et al.* Synthesis and antiviral evaluation of a novel series of homoserine-based inhibitors of the hepatitis C virus NS3/4A serine protease. **25**, 3984-3991 (2015).

40 Parsek, M. R., Val, D. L., Hanzelka, B. L., Cronan, J. E., Jr. & Greenberg, E. P. Acyl homoserine-lactone quorum-sensing signal generation. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 4360-4365, doi:10.1073/pnas.96.8.4360 (1999).

41 Blackwell, H. E. *et al.* Ring-closing metathesis of olefinic peptides: design, synthesis, and structural characterization of macrocyclic helical peptides. **66**, 5291-5302 (2001).

42 Saro, D., Klosi, E., Paredes, A. & Spaller, M. R. J. O. l. Thermodynamic analysis of a hydrophobic binding site: probing the PDZ domain with nonproteinogenic peptide ligands. **6**, 3429-3432 (2004).

43 Lorenz, M., Evers, A., Wagner, M. J. B. & letters, m. c. Recent progress and future options in the development of GLP-1 receptor agonists for the treatment of diabesity. **23**, 4011-4018 (2013).

44 Fragiadaki, M., Magafa, V., Slaninová, J. & Cordopatis, P. Synthesis and biological evaluation of oxytocin analogues containing l-α-t-butylglycine [Gly(But)] in positions 8 or 9. *Peptides* **24**, 1425-1431, doi:https://doi.org/10.1016/j.peptides.2003.09.007 (2003).

45 Feelders, R. A., Yasothan, U. & Kirkpatrick, P. J. N. R. D. D. Pasireotide. **11**, 597-599 (2012).

46 Llinas-Brunet, M. *et al.* Discovery of a potent and selective noncovalent linear inhibitor of the hepatitis C virus NS3 protease (BI 201335). **53**, 6466-6476 (2010).

47 Scola, P. M. *et al.* The discovery of asunaprevir (BMS-650032), an orally efficacious NS3 protease inhibitor for the treatment of hepatitis C virus infection. **57**, 1730-1752 (2014).

48 Evans, D. A. & Weber, A. E. J. J. o. t. A. C. S. Synthesis of the cyclic hexapeptide echinocandin D. New approaches to the asymmetric synthesis of. beta.-hydroxy. alpha.-amino acids. **109**, 7151-7157 (1987).

49 Burlacu, C. L., Buggy, D. J. J. T. & management, c. r. Update on local anesthetics: focus on levobupivacaine. **4**, 381 (2008).

50 Gurk-Turner, C. in *Baylor University Medical Center Proceedings.* 83-86 (Taylor & Francis).

51 Kirkpatrick, P., Raja, A., LaBonte, J. & Lebbos, J. J. N. r. D. d. Daptomycin. **2**, 943-944 (2003).

52 Gatto, G. J., Boyne, M. T., Kelleher, N. L. & Walsh, C. T. Biosynthesis of Pipecolic Acid by RapL, a Lysine Cyclodeaminase Encoded in the Rapamycin Gene Cluster. *Journal of the American Chemical Society* **128**, 3838-3847, doi:10.1021/ja0587603 (2006).

53 Tran, T. T., Patino, N., Condom, R., Frogier, T. & Guedj, R. Fluorinated peptides incorporating a 4-fluoroproline residue as potential inhibitors of HIV protease. *Journal of Fluorine Chemistry* **82**, 125-130, doi:https://doi.org/10.1016/S0022-1139(96)03568-3 (1997).
54 Bork, K., Yasothan, U. & Kirkpatrick, P. J. N. R. D. D. Icatibant. **7**, 801-803 (2008).
55 Asami, T. *et al.* Design, synthesis, and biological evaluation of novel investigational nonapeptide KISS1R agonists with testosterone-suppressive activity. **56**, 8298-8307 (2013).
56 Crounse, R., Maxwell, J. & Blank, H. J. N. Inhibition of growth of hair by mimosine. **194**, 694-695 (1962).
57 Kleemann, A., Hoppe, B. & Tanner, H. (VCH: Weinheim, Federal Republic of Germany, 1985).
58 Amer, A. *et al.* 5-Hydroxy-L-tryptophan suppresses food intake in food-deprived and stressed rats. **77**, 137-143 (2004).
59 Horne, W. S. & Grossmann, T. N. Proteomimetics as protein-inspired scaffolds with defined tertiary folding patterns. *Nature Chemistry* **12**, 331-337, doi:10.1038/s41557-020-0420-9 (2020).
60 Tsomaia, N. Peptide therapeutics: targeting the undruggable space. *European journal of medicinal chemistry* **94**, 459-470, doi:10.1016/j.ejmech.2015.01.014 (2015).
61 Chang, Y. S. *et al.* Stapled α-helical peptide drug development: a potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E3445-3454, doi:10.1073/pnas.1303002110 (2013).
62 Johnson, L. M. *et al.* A potent α/β-peptide analogue of GLP-1 with prolonged action in vivo. *J Am Chem Soc* **136**, 12848-12851, doi:10.1021/ja507168t (2014).
63 Jochim, A. L. & Arora, P. S. Systematic analysis of helical protein interfaces reveals targets for synthetic inhibitors. *ACS chemical biology* **5**, 919-923, doi:10.1021/cb1001747 (2010).
64 Bullock, B. N., Jochim, A. L. & Arora, P. S. Assessing Helical Protein Interfaces for Inhibitor Design. *Journal of the American Chemical Society* **133**, 14220-14223, doi:10.1021/ja206074j (2011).
65 Lao, B. B. *et al.* Rational design of topographical helix mimics as potent inhibitors of protein-protein interactions. *Journal of the American Chemical Society* **136**, 7877-7888, doi:10.1021/ja502310r (2014).
66 Azzarito, V., Long, K., Murphy, N. S. & Wilson, A. J. Inhibition of α-helix-mediated protein-protein interactions using designed molecules. *Nat Chem* **5**, 161-173, doi:10.1038/nchem.1568 (2013).
67 Walensky, L. D. & Bird, G. H. Hydrocarbon-stapled peptides: principles, practice, and progress. *Journal of medicinal chemistry* **57**, 6275-6288, doi:10.1021/jm4011675 (2014).
68 Sawada, T. & Gellman, S. H. Structural mimicry of the α-helix in aqueous solution with an isoatomic α/β/γ-peptide backbone. *J Am Chem Soc* **133**, 7336-7339, doi:10.1021/ja202175a (2011).
69 Bartus, É. *et al.* De Novo Modular Development of a Foldameric Protein–Protein Interaction Inhibitor for Separate Hot Spots: A Dynamic Covalent Assembly Approach. *ChemistryOpen* **6**, 236-241, doi:doi:10.1002/open.201700012 (2017).
70 Angelo, N. G. & Arora, P. S. Nonpeptidic foldamers from amino acids: synthesis and characterization of 1,3-substituted triazole oligomers. *J Am Chem Soc* **127**, 17134-17135, doi:10.1021/ja056406z (2005).
71 Freire, F. & Gellman, S. H. Macrocyclic design strategies for small, stable parallel beta-sheet scaffolds. *J Am Chem Soc* **131**, 7970-7972, doi:10.1021/ja902210f (2009).
72 Hegedüs, Z. *et al.* Identification of β-strand mediated protein–protein interaction inhibitors using ligand-directed fragment ligation. *Chemical Science* **12**, 2286-2293, doi:10.1039/D0SC05694D (2021).
73 Guharoy, M. & Chakrabarti, P. Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein-protein interactions. *Bioinformatics (Oxford, England)* **23**, 1909-1918, doi:10.1093/bioinformatics/btm274 (2007).
74 Villar, E. A. *et al.* How proteins bind macrocycles. *Nature chemical biology* **10**, 723-731, doi:10.1038/nchembio.1584 (2014).
75 Gilbreth, R. N. & Koide, S. Structural insights for engineering binding proteins based on non-antibody scaffolds. *Current opinion in structural biology* **22**, 413-420, doi:10.1016/j.sbi.2012.06.001 (2012).
76 Hosse, R. J., Rothe, A. & Power, B. E. A new generation of protein display scaffolds for molecular recognition. *Protein science : a publication of the Protein Society* **15**, 14-27, doi:10.1110/ps.051817606 (2006).
77 Checco, J. W. *et al.* Targeting diverse protein-protein interaction interfaces with α/β-peptides derived from the Z-domain scaffold. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 4552-4557, doi:10.1073/pnas.1420380112 (2015).
78 Tavenor, N. A., Reinert, Z. E., Lengyel, G. A., Griffith, B. D. & Horne, W. S. Comparison of design strategies for α-helix backbone modification in a protein tertiary fold. *Chemical communications (Cambridge, England)* **52**, 3789-3792, doi:10.1039/c6cc00273k (2016).
79 Hewitt, S. H. *et al.* Protein Surface Mimetics: Understanding How Ruthenium Tris(Bipyridines) Interact with Proteins. **18**, 223-231, doi:https://doi.org/10.1002/cbic.201600552 (2017).

80 Hewitt, S. H. & Wilson, A. J. Metal complexes as “protein surface mimetics”. *Chemical Communications* **52**, 9745-9756, doi:10.1039/C6CC03457H (2016).
81 Werner, H. M. & Horne, W. S. Folding and function in α/β-peptides: targets and therapeutic applications. *Current opinion in chemical biology* **28**, 75-82, doi:10.1016/j.cbpa.2015.06.013 (2015).
82 Martinek, T. A. & Fülöp, F. Side-chain control of beta-peptide secondary structures. *European journal of biochemistry* **270**, 3657-3666, doi:10.1046/j.1432-1033.2003.03756.x (2003).
83 Kritzer, J. A., Stephens, O. M., Guarracino, D. A., Reznik, S. K. & Schepartz, A. beta-Peptides as inhibitors of protein-protein interactions. *Bioorganic & medicinal chemistry* **13**, 11-16, doi:10.1016/j.bmc.2004.09.009 (2005).
84 Sadowsky, J. D. *et al.* (α/β+α)-Peptide Antagonists of BH3 Domain/Bcl-xL Recognition: Toward General Strategies for Foldamer-Based Inhibition of Protein−Protein Interactions. *Journal of the American Chemical Society* **129**, 139-154, doi:10.1021/ja0662523 (2007).
85 Kritzer, J. A., Lear, J. D., Hodsdon, M. E. & Schepartz, A. Helical beta-peptide inhibitors of the p53-hDM2 interaction. *Journal of the American Chemical Society* **126**, 9468-9469, doi:10.1021/ja031625a (2004).
86 Haase, H. S. *et al.* Extending foldamer design beyond α-helix mimicry: α/β-peptide inhibitors of vascular endothelial growth factor signaling. *Journal of the American Chemical Society* **134**, 7652-7655, doi:10.1021/ja302469a (2012).
87 Seebach, D., Beck, A. K., Bierbaum, D. J. J. C. & biodiversity. The world of β-and γ-peptides comprised of homologated proteinogenic amino acids and other components. **1**, 1111-1239 (2004).
88 Peggion, E. *et al.* Structure− Function Studies of Analogues of Parathyroid Hormone (PTH)-1− 34 Containing β-Amino Acid Residues in Positions 11− 13. **41**, 8162-8175 (2002).
89 Denton, E. V. *et al.* A β-peptide agonist of the GLP-1 receptor, a class B GPCR. **15**, 5318-5321 (2013).
90 Guharoy, M. & Chakrabarti, P. Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein–protein interactions. *Bioinformatics* **23**, 1909-1918, doi:10.1093/bioinformatics/btm274 %J Bioinformatics (2007).
91 Mándity, I. M. *et al.* Building β-Peptide H10/12 Foldamer Helices with Six-Membered Cyclic Side-Chains: Fine-Tuning of Folding and Self-Assembly. *Organic Letters* **12**, 5584-5587, doi:10.1021/ol102494m (2010).
92 Rathore, N., Gellman, S. H. & de Pablo, J. J. Thermodynamic Stability of β-Peptide Helices and the Role of Cyclic Residues. *Biophysical journal* **91**, 3425-3435, doi:https://doi.org/10.1529/biophysj.106.084491 (2006).
93 Hetényi, A. *et al.* Sculpting the β-peptide foldamer H12 helix via a designed side-chain shape. *Chemical Communications*, 177-179, doi:10.1039/B812114A (2009).
94 Hruby, V. J., al-Obeidi, F. & Kazmierski, W. Emerging approaches in the molecular design of receptor-selective peptide ligands: conformational, topographical and dynamic considerations. *The Biochemical journal* **268**, 249-262, doi:10.1042/bj2680249 (1990).
95 Jones, R. M., Boatman, P. D., Semple, G., Shin, Y. J. & Tamura, S. Y. Clinically validated peptides as templates for de novo peptidomimetic drug design at G-protein-coupled receptors. *Curr Opin Pharmacol* **3**, 530-543, doi:10.1016/j.coph.2003.06.003 (2003).
96 Lang, M. & Beck-Sickinger, A. G. Structure-activity relationship studies: methods and ligand design for G-protein coupled peptide receptors. *Current protein & peptide science* **7**, 335-353, doi:10.2174/138920306778017981 (2006).
97 Ripka, A. S. & Rich, D. H. Peptidomimetic design. *Current opinion in chemical biology* **2**, 441-452, doi:10.1016/s1367-5931(98)80119-1 (1998).
98 Perez, J. J. Designing Peptidomimetics. *Current topics in medicinal chemistry* **18**, 566-590, doi:10.2174/1568026618666180522075258 (2018).
99 Cheloha, R. W. *et al.* Development of Potent, Protease-Resistant Agonists of the Parathyroid Hormone Receptor with Broad β Residue Distribution. *J. Med. Chem.* **60**, 8816-8833, doi:10.1021/acs.jmedchem.7b00876 (2017).
100 Liu, S., Cheloha, R. W., Watanabe, T., Gardella, T. J. & Gellman, S. H. Receptor selectivity from minimal backbone modification of a polypeptide agonist. **115**, 12383-12388, doi:doi:10.1073/pnas.1815294115 (2018).
101 Cheloha, R. W. *et al.* Recognition of Class II MHC Peptide Ligands That Contain β-Amino Acids. *Journal of immunology (Baltimore, Md. : 1950)* **203**, 1619-1628, doi:10.4049/jimmunol.1900536 (2019).
102 Peterson-Kaufman, K. J. *et al.* Residue-Based Preorganization of BH3-Derived α/β-Peptides: Modulating Affinity, Selectivity and Proteolytic Susceptibility in α-Helix Mimics. *ACS Chemical Biology* **10**, 1667-1675, doi:10.1021/acschembio.5b00109 (2015).

103 Sadowsky, J. D. *et al.* Chimeric (α/β + α)-Peptide Ligands for the BH3-Recognition Cleft of Bcl-xL: Critical Role of the Molecular Scaffold in Protein Surface Recognition. *Journal of the American Chemical Society* **127**, 11966-11968, doi:10.1021/ja053678t (2005).

104 George, K. L. & Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *Journal of the American Chemical Society* **139**, 7931-7938, doi:10.1021/jacs.7b03114 (2017).

105 Gopalakrishnan, R., Frolov, A. I., Knerr, L., Drury, W. J., III & Valeur, E. Therapeutic Potential of Foldamers: From Chemical Biology Tools To Drug Candidates? *J. Med. Chem.* **59**, 9599-9621, doi:10.1021/acs.jmedchem.6b00376 (2016).

106 Horne, W. S. & Gellman, S. H. Foldamers with Heterogeneous Backbones. *Accounts of Chemical Research* **41**, 1399-1408, doi:10.1021/ar800009n (2008).

107 Ferrero, E. *et al.* Ten simple rules to power drug discovery with data science. *PLoS computational biology* **16**, e1008126, doi:10.1371/journal.pcbi.1008126 (2020).

108 Lenci, E. & Trabocchi, A. Peptidomimetic toolbox for drug discovery. *Chemical Society Reviews* **49**, 3262-3277, doi:10.1039/D0CS00102C (2020).

109 Murray, C. W. & Rees, D. C. The rise of fragment-based drug discovery. *Nature Chemistry* **1**, 187-192, doi:10.1038/nchem.217 (2009).

110 Erlanson, D. A., McDowell, R. S. & O'Brien, T. Fragment-Based Drug Discovery. *J. Med. Chem.* **47**, 3463-3482, doi:10.1021/jm040031v (2004).

111 Simpson, M. G., Pittelkow, M., Watson, S. P. & Sanders, J. K. Dynamic combinatorial chemistry with hydrazones: libraries incorporating heterocyclic and steroidal motifs. *Org Biomol Chem* **8**, 1181-1187, doi:10.1039/b917146k (2010).

112 Huc, I. & Nguyen, R. Dynamic combinatorial chemistry. *Combinatorial chemistry & high throughput screening* **4**, 53-74, doi:10.2174/1386207013331273 (2001).

113 Corbett, P. T. *et al.* Dynamic Combinatorial Chemistry. *Chemical reviews* **106**, 3652-3711, doi:10.1021/cr020452p (2006).

114 Klepel, F. & Ravoo, B. J. Dynamic covalent chemistry in aqueous solution by photoinduced radical disulfide metathesis. *Organic & Biomolecular Chemistry* **15**, 3840-3842, doi:10.1039/C7OB00667E (2017).

115 Li, L., Feng, W., Welle, A. & Levkin, P. A. UV-Induced Disulfide Formation and Reduction for Dynamic Photopatterning. *Angewandte Chemie (International ed. in English)* **55**, 13765-13769, doi:10.1002/anie.201607276 (2016).

116 Scheuermann, T. H. & Brautigam, C. A. High-precision, automated integration of multiple isothermal titration calorimetric thermograms: New features of NITPIC. *Methods* **76**, 87-98, doi:10.1016/j.ymeth.2014.11.024 (2015).

117 Zhao, H., Piszczek, G. & Schuck, P. SEDPHAT – A platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods* **76**, 137-148, doi:10.1016/j.ymeth.2014.11.012 (2015).

118 Hetényi, A. *et al.* Sculpting the beta-peptide foldamer H12 helix via a designed side-chain shape. *Chemical communications (Cambridge, England)*, 177-179, doi:10.1039/b812114a (2009).

119 Li, I. T., Ranjith, K. R. & Truong, K. Sequence reversed peptide from CaMKK binds to calmodulin in reversible Ca2+ -dependent manner. *Biochemical and biophysical research communications* **352**, 932-935, doi:10.1016/j.bbrc.2006.11.123 (2007).

120 Gonzalez, L. L., Garrie, K. & Turner, M. D. Role of S100 proteins in health and disease. *Biochimica et biophysica acta. Molecular cell research* **1867**, 118677, doi:10.1016/j.bbamcr.2020.118677 (2020).

121 Bresnick, A. R., Weber, D. J. & Zimmer, D. B. S100 proteins in cancer. *Nature Reviews Cancer* **15**, 96-109, doi:10.1038/nrc3893 (2015).

122 Wang, D. D. & Bordey, A. The astrocyte odyssey. *Progress in neurobiology* **86**, 342-367, doi:10.1016/j.pneurobio.2008.09.015 (2008).

123 Camby, I., Le Mercier, M., Lefranc, F. & Kiss, R. Galectin-1: a small protein with major functions. *Glycobiology* **16**, 137R-157R, doi:10.1093/glycob/cwl025 %J Glycobiology (2006).

124 Gupta, D., Cho, M., Cummings, R. D. & Brewer, C. F. Thermodynamics of carbohydrate binding to galectin-1 from Chinese hamster ovary cells and two mutants. A comparison with four galactose-specific plant lectins. *Biochemistry* **35**, 15236-15243, doi:10.1021/bi961458+ (1996).

125 Miller, M. C. *et al.* Binding of polysaccharides to human galectin-3 at a noncanonical site in its carbohydrate recognition domain. *Glycobiology* **26**, 88-99, doi:10.1093/glycob/cwv073 (2016).

126 Gabius, H. J. Non-carbohydrate binding partners/domains of animal lectins. *The International journal of biochemistry* **26**, 469-477, doi:10.1016/0020-711x(94)90002-7 (1994).

127 Blois, S. M. *et al.* Pregnancy Galectinology: Insights Into a Complex Network of Glycan Binding Proteins. *Frontiers in immunology* **10**, 1166, doi:10.3389/fimmu.2019.01166 (2019).

128 Umezu, K., Nakayama, K. & Nakayama, H. Escherichia coli RecQ protein is a DNA helicase. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 5363-5367, doi:10.1073/pnas.87.14.5363 (1990).
129 Bernstein, D. A., Zittel, M. C. & Keck, J. L. High-resolution structure of the E.coli RecQ helicase catalytic core. *The EMBO journal* **22**, 4910-4921, doi:10.1093/emboj/cdg500 (2003).
130 Shereda, R. D., Bernstein, D. A. & Keck, J. L. A Central Role for SSB in Escherichia coli RecQ DNA Helicase Function. *Journal of Biological Chemistry* **282**, 19247-19258, doi:10.1074/jbc.M608011200 (2007).
131 Tidow, H. & Nissen, P. Structural diversity of calmodulin binding to its target sites. *The FEBS journal* **280**, 5551-5565, doi:10.1111/febs.12296 (2013).
132 Stark, C. *et al.* BioGRID: a general repository for interaction datasets. *Nucleic acids research* **34**, D535-539, doi:10.1093/nar/gkj109 (2006).
133 Orii, N. & Ganapathiraju, M. K. Wiki-pi: a web-server of annotated human protein-protein interactions to aid in discovery of protein function. *PLoS ONE* **7**, e49029, doi:10.1371/journal.pone.0049029 (2012).
134 Fahey, M. E. *et al.* GPS-Prot: A web-based visualization platform for integrating host-pathogen interaction data. *BMC Bioinformatics* **12**, 298, doi:10.1186/1471-2105-12-298 (2011).
135 Orchard, S. *et al.* The MIntAct project--IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic acids research* **42**, D358-363, doi:10.1093/nar/gkt1115 (2014).
136 Alonso-López, D. *et al.* APID database: redefining protein-protein interaction experimental evidences and binary interactomes. *Database : the journal of biological databases and curation* **2019**, doi:10.1093/database/baz005 (2019).
137 Hajduk, P. J. & Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov* **6**, 211-219, doi:10.1038/nrd2220 (2007).
138 Gógl, G. *et al.* Structural Basis of Ribosomal S6 Kinase 1 (RSK1) Inhibition by S100B Protein: MODULATION OF THE EXTRACELLULAR SIGNAL-REGULATED KINASE (ERK) SIGNALING CASCADE IN A CALCIUM-DEPENDENT WAY. *J Biol Chem* **291**, 11-27, doi:10.1074/jbc.M115.684928 (2016).
139 Asensio, J. L. *et al.* Bovine Heart Galectin-1 Selects a Unique (Syn) Conformation of C-Lactose, a Flexible Lactose Analogue. *Journal of the American Chemical Society* **121**, 8995-9000, doi:10.1021/ja990601u (1999).
140 Shereda, R. D., Bernstein, D. A. & Keck, J. L. A central role for SSB in Escherichia coli RecQ DNA helicase function. *J Biol Chem* **282**, 19247-19258, doi:10.1074/jbc.M608011200 (2007).
141 Gómez, A. *et al.* Do protein-protein interaction databases identify moonlighting proteins? *Molecular bioSystems* **7**, 2379-2382, doi:10.1039/c1mb05180f (2011).
142 Watkins, A. M., Bonneau, R. & Arora, P. S. Side-Chain Conformational Preferences Govern Protein-Protein Interactions. *Journal of the American Chemical Society* **138**, 10386-10389, doi:10.1021/jacs.6b04892 (2016).
143 Bogan, A. A. & Thorn, K. S. Anatomy of hot spots in protein interfaces. *J Mol Biol* **280**, 1-9, doi:10.1006/jmbi.1998.1843 (1998).
144 Mruk, K., Farley, B. M., Ritacco, A. W. & Kobertz, W. R. Calmodulation meta-analysis: predicting calmodulin binding via canonical motif clustering. *The Journal of general physiology* **144**, 105-114, doi:10.1085/jgp.201311140 (2014).
145 Ikura, M. & Yap, K. L. Where cancer meets calcium — p53 crosstalk with EF-hands. *Nature Structural Biology* **7**, 525-527, doi:10.1038/76721 (2000).
146 Rustandi, R. R., Baldisseri, D. M. & Weber, D. J. Structure of the negative regulatory domain of p53 bound to S100B(betabeta). *Nat Struct Biol* **7**, 570-574, doi:10.1038/76797 (2000).
147 Zimmer, D. B. & Weber, D. J. The Calcium-Dependent Interaction of S100B with Its Protein Targets. *Cardiovascular psychiatry and neurology* **2010**, doi:10.1155/2010/728052 (2010).
148 Lu, D. & Keck, J. L. Structural basis of Escherichia coli single-stranded DNA-binding protein stimulation of exonuclease I. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 9169-9174, doi:10.1073/pnas.0800741105 (2008).
149 Ryzhikov, M., Koroleva, O., Postnov, D., Tran, A. & Korolev, S. Mechanism of RecO recruitment to DNA by single-stranded DNA binding protein. *Nucleic acids research* **39**, 6305-6314, doi:10.1093/nar/gkr199 (2011).
150 Cadman, C. J. & McGlynn, P. PriA helicase and SSB interact physically and functionally. *Nucleic acids research* **32**, 6378-6387, doi:10.1093/nar/gkh980 (2004).
151 Wu, H.-Y., Lu, C.-H. & Li, H.-W. RecA-SSB Interaction Modulates RecA Nucleoprotein Filament Formation on SSB-Wrapped DNA. *Sci Rep* **7**, 11876, doi:10.1038/s41598-017-12213-w (2017).
152 Bianco, P. R. The tale of SSB. *Progress in Biophysics and Molecular Biology* **127**, 111-118, doi:10.1016/j.pbiomolbio.2016.11.001 (2017).

153 Wu, H. Y., Lu, C. H. & Li, H. W. RecA-SSB Interaction Modulates RecA Nucleoprotein Filament Formation on SSB-Wrapped DNA. *Sci Rep* **7**, 11876, doi:10.1038/s41598-017-12213-w (2017).
154 Lehman, I. R. & Nussbaum, A. L. THE DEOXYRIBONUCLEASES OF ESCHERICHIA COLI. V. ON THE SPECIFICITY OF EXONUCLEASE I (PHOSPHODIESTERASE). *J Biol Chem* **239**, 2628-2636 (1964).
155 Sakai, A. & Cox, M. M. RecFOR and RecOR as Distinct RecA Loading Pathways*. *Journal of Biological Chemistry* **284**, 3264-3272, doi:https://doi.org/10.1074/jbc.M807220200 (2009).
156 Voter, A. F. *et al.* A High-Throughput Screening Strategy to Identify Inhibitors of SSB Protein–Protein Interactions in an Academic Screening Facility. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **23**, 94-101, doi:10.1177/2472555217712001 (2018).
157 Chilingaryan, Z. *et al.* Fragment-Based Discovery of Inhibitors of the Bacterial DnaG-SSB Interaction. *Antibiotics* **7**, 14, doi:10.3390/antibiotics7010014 (2018).
158 Lu, D., Bernstein, D. A., Satyshur, K. A. & Keck, J. L. Small-molecule tools for dissecting the roles of SSB/protein interactions in genome maintenance. *Proceedings of the National Academy of Sciences* **107**, 633-638, doi:10.1073/pnas.0909191107 (2010).
159 Marceau, A. H. *et al.* Protein Interactions in Genome Maintenance as Novel Antibacterial Targets. *PLoS ONE* **8**, e58765, doi:10.1371/journal.pone.0058765 (2013).
160 Alnammi, M. *et al.* Evaluating scalable supervised learning for synthesize-on-demand chemical libraries. (Chemistry, 2021).
161 Lu, D. & Keck, J. L. Structural basis of Escherichia coli single-stranded DNA-binding protein stimulation of exonuclease I. *Proceedings of the National Academy of Sciences* **105**, 9169-9174, doi:10.1073/pnas.0800741105 (2008).
162 Ryzhikov, M., Koroleva, O., Postnov, D., Tran, A. & Korolev, S. Mechanism of RecO recruitment to DNA by single-stranded DNA binding protein. *Nucleic acids research* **39**, 6305-6314, doi:10.1093/nar/gkr199 (2011).
163 Wang, L., Friesner, R. A. & Berne, B. J. Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). *The journal of physical chemistry. B* **115**, 9431-9438, doi:10.1021/jp204407d (2011).
164 Liu, P., Kim, B., Friesner, R. A. & Berne, B. J. Replica exchange with solute tempering: A method for sampling biological systems in explicit water. **102**, 13749-13754, doi:doi:10.1073/pnas.0506346102 (2005).
165 Schrödinger Release 2021-1 (Schrödinger, LLC, New York, NY, 2021).

10. Appendix