Double-CLIPS technology for the mimicry of structurally complex antibody binding sites on proteins
Smeenk, L.E.J.

Citation for published version (APA):
Smeenk, L. E. J. (2013). Double-CLIPS technology for the mimicry of structurally complex antibody binding sites on proteins

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
THE ART OF PROTEIN MIMICRY
1.1 GENERAL INTRODUCTION

In 1989, Mutter and Vuilleumier posed the rather philosophical question “what sense it makes to redesign and improve nature on a molecular level?”\(^1\) Apart from the main scientific challenge for the synthetic chemist to devise and synthesize new molecules with well-defined and interesting properties, there is a real need for tailor-made molecules that accomplish tasks comparable to those found in nature by large biomolecules such as proteins. Examples of such tasks that involve protein-protein interactions include the immune response,\(^2\) apoptosis,\(^3\) DNA replication and signal transduction.\(^4\) It will therefore not be surprising that disturbance in the regulation of these interactions are the cause of many life threatening diseases e.g. cancer, malaria, HIV, Huntington’s disease, osteoporosis, leukemia and other auto-immune or infectious diseases.\(^5,6,7,8,9\) Since the interaction between proteins forms the basis of essential biological processes, understanding the exact nature of the non-covalent binding site between proteins could lead to the definition of new targets for the design of potential medicines for the above diseases.

Small molecule inhibitors obeying Lipinski’s rule of five\(^10\) have long been the only synthetic molecules available to modulate and/or interfere with protein-protein interactions, but ensuring appropriate activity and robust selectivity was found to be highly challenging for these small alternatives. Although the principle of vaccination was already discovered and investigated by Jenner and Pasteur in the eighteenth century, the idea of using weakened or dead viruses as medicines, to elicit an immune response to generate antibodies, only originates from the last 20 years. The increasing knowledge about the mechanism of the human defense system and the role of protein-protein interactions therein resulted in bio-inspired mimics, and their mode of action is purely based on selective interference with protein-protein interactions. Moreover, synthetic antibodies have already proven to be highly selective and therefore successful in the treatment of several life-threatening diseases\(^11,12,13\) and are an important topic of current research.\(^14,15,16\)

Nevertheless, drug development nowadays is shifting its focus towards protein mimics of sizes in between the group of small molecule inhibitors and large synthetic antibodies. This gives access to a group of molecules that is able to reproduce key protein contacts, while taking advantage of the chemical properties inherent to synthetic molecules. The concept of protein mimicry using peptides, which are directly derived from proteins, can only be successful if they also represent the correct 3D-topology. Because of these structural requirements, the design and synthesis of such peptide-based protein mimics remains far more complicated than the current application of small organic molecules used in classical approaches for drug design. Furthermore, little is known about the minimum structural requirement of protein mimics that may play an important role for the production of successful mimics. This chapter describes the development over the past few decades of peptide-based molecules that imitate the secondary and/or tertiary structures of proteins in order to obtain potential candidates that can function as therapeutics, or ultimately, as synthetic vaccines.
1.2 EXPLORING PROTEIN-PROTEIN INTERACTIONS (PPIs)

If proteins play such an important role in many processes in life and related diseases, why not use them directly as medicines? It has been generally accepted that proteins are rather poor drug candidates, due to their unpredictable bioavailability and the uncertainties regarding their exact mode of action. From a recent point of view, synthetic peptides could serve as exogenous structural and functional mimics of the highly complex bioactive regions of proteins\(^\text{17}\). However, with some exceptions\(^\text{18,19,20}\), few examples of such classical small molecules that interfere with protein interactions are known. The reason for this problem lies in the nature of protein binding sites\(^\text{21}\).

1.2.1 The Definition of Protein Binding Sites

In general, the understanding of protein-protein interactions at the atomic level is rather limited, so a precise description of all non-covalent binding sites is not possible. The contact region mostly encompasses a large surface area (approximately 750-1500 Å, Figure 1.1),\(^\text{22}\) based on non-covalent interactions such as hydrogen bonds, electrostatic interactions, hydrophobic interactions, metal coordination and Van der Waals interactions. Covering these large areas with small molecules is a huge challenge, because the non-covalent interface available in small molecules is not sufficiently large to make them selective for one protein interaction site. The discovery of protein ‘hot-spots’ (i.e. the most important residues or areas of a protein associated in binding) simplified the task of protein mimicry to some extent, because it clarified the highly ordered structural motifs found within these hot-spots and thus stimulated the mimicry of protein secondary structures.\(^\text{23}\) Still, the most important reason for the failure of small molecules to interfere with protein-protein interactions is found in the structural complexity of the majority of protein binding sites. Proteins usually fold into complex 3D structures, whereby only parts of their solvent-exposed surfaces play a role in protein-protein interactions. The structural topologies of these protein parts determine the complexity of the binding site. The exact portion of the protein that is recognized by its antibody is defined as the epitope. A minority of proteins express their epitopes via a linear sequence on the protein, defined as a “continuous epitope” (Figure 1.2), but the majority of protein binding interactions are found to be composed of two or more continuous parts that are located far away from each other on the linear protein chain, but brought in close proximity during folding to form a “discontinuous

\[\text{Figure 1.1: X-ray structure of a protein-protein interaction site.}\]

\[\text{Figure 1.2: Schematic representation of a continuous and discontinuous epitope.}\]
epitope\textsuperscript{a} \textsuperscript{24} (Figure 1.2). For the description and mimicry of epitopes, three levels may be distinguished. With respect to the continuous epitopes, level 1 is assigned to linear, unstructured binding sites of proteins, because they often can be successfully mimicked by linear synthetic peptides (Figure 1.3).

![Figure 1.3: Three classes of epitope mimicry](image)

The continuous linear epitope on Protein A (level 1) is best mimicked by a corresponding linear peptide for binding to protein B. The conformational epitope presented on Protein C (level 2) requires at least a correct structure of a presented peptide mimic for strong binding to Protein D. The discontinuous conformational epitope of Protein E (level 3) can only be successfully mimicked when both separate peptide fragments are connected in a peptide mimic.

Clearly, when continuous epitopes are more structurally defined, e.g. because they contain an \( \alpha \)-helical, a \( \beta \)-sheet or a \( \beta \)-turn element, its mimicry is better defined when the relevant peptide fragments are constrained into their native structure. This level of conformational continuous protein epitopes is considered level 2, because linear unstructured peptides are not able to mimic the protein anymore, while correctly constrained peptides are more successful (Figure 1.3). Nevertheless, when a peptide is able to react with an antibody, the description of ‘continuous epitope’ is often incorrectly used for the protein epitope, because in most cases the epitope is much more complex and consists of several separated and folded peptide fragments and this epitope should therefore be considered a discontinuous rather than a continuous epitope.\textsuperscript{25} For example, the epitope of the Fab fragment of mAb 184.1 on the outer surface protein A (OspA) has been clarified by X-ray structures and was found to consist of the residues 30, 33-35, 42-46, 52, 69-71, 92-95 and 117-119, because these residues are in contact with the antibody.\textsuperscript{26} Peptides covering the residues 30-38 and 65-73 and constrained to present their native secondary structure may very well bind (usually only weakly) to mAb 184.1 since these residues are part of the interaction surface. Following the above description of epitopes, these peptides will be labeled “conformational (level 2) continuous epitopes” of the protein. However, together they only contain a few of the residues that correspond to the discontinuous epitope (level 3) recognized by mAb 184.1 (Figure 1.3).

1.2.2 Inhibition of Protein-Protein Interactions (PPIs)

Because protein-protein interactions play such an important role within for example immune responses, it becomes highly attractive to develop molecules that are able to disturb or even inhibit these interactions in order to treat related disorders. As stated earlier, success within the strategy of interfering protein-protein interactions relies on the correct presentation of the majority of discontinuous binding sites recognized by our natural
antibodies. The in 1975 developed hybridoma technology for the production of monoclonal antibodies, clearly revolutionized science, because ten years later, the first (monoclonal) therapeutic antibody muromonab reached the market. This monoclonal antibody targets the CD3-receptor, a membrane protein on the surface of T-cells to reduce acute rejection of organ transplants. After muromonab, the vascular endothelial growth factor specific antibody bevacizumab (Avastin; Genentech/Roche) and the human epidermal growth factor receptor 2 specific antibody trastuzumab (Herceptin; Genentech/Roche) reached the market and are currently used for the treatment of several types of cancer. In addition, the tumor necrosis factor specific antibodies infliximab (Remicade; Centocor/Merck) and adalimumab (Humira; Trudexa/Abbott) are nowadays applied to treat rheumatoid arthritis. Undeniably, the advantage of using nature-based antibodies for treatment of life-threatening diseases is mostly found in their high specificity in blocking a specific protein-protein interaction, because they recognize highly discontinuous binding sites. But despite this, other challenges to enhance the application of monoclonal antibodies have to be conquered. For example, synthetic antibodies are quite unstable under biological conditions and their oral bioavailability is still difficult to realize. Moreover, as a consequence of the huge technological requirements and long production timescale, the costs for treating a patient with monoclonal antibodies are extremely high. Investigation of the first generation of therapeutic antibodies for a deeper understanding of their mode of action and optimization of the current treatments forms the next step towards more specific and safer therapies, but goes far beyond the scope of this literature study.

The mode of action of synthetic antibodies could in a way be copied by using small antibody fragments, that imitate their discontinuous complementarity-determining regions (CDRs), e.g. by using peptides or protein fragments. The CDRs of a dominant idiotypic antibody (T15) were the basis of the first CDR-derived peptide-based antibody mimics reported by Kang et al. in 1988. However, it has been shown that there is a crucial difference between the binding mode of monoclonal antibodies and their CDR-derived synthetic binders. A recent study by Timmerman & Altschuh showed that the selectivity and high-affinity binding expressed by monoclonal antibodies stands in sharp contrast with the binding mode of their CDR-derived peptide mimics, which mainly stick to their targets via a combination of (non-selective) electrostatic and hydrophobic interactions. As a result, there is little relation between the binding mode of antibodies and their CDR-derived mimics. It should therefore be concluded that the CDR's of monoclonal antibodies do not necessarily provide the optimal basis for functional antibody mimics design.

As an alternative to CDR-derived peptides, mimics derived from proteins other than antibodies can sometimes also fulfill the binding requirements for inhibiting protein-protein interactions involved in other biological processes, such as signal transduction and apoptosis. Nature provides countless examples of peptides functioning as enzyme inhibitors, receptor agonists or antagonists, hormones and as neurotransmitters, and synthetic peptides nowadays have found applications in for example replacing proteins as reagents for diagnosis of viral and autoimmune diseases. Constraining unstructured linear synthetic peptides into similar protein secondary structures like α-helices or β-sheets forms the basis of the development of new therapeutic peptides, as recently reviewed by Hamilton. Clearly, the most promising work in this field was done by the Walensky, Verdine,
Cowburn,\textsuperscript{41} and Fairlie\textsuperscript{42} groups, showing how ‘stapled peptides’, i.e. peptides that are covalently linked between several residues to induce helix formation, provide an outstanding example of bypassing typical problems associated with peptide drugs by using constrained $\alpha$-helical peptides (Figure 1.4). The application of stapled peptides by Aileron Therapeutics highlighted their resistance to proteases and their ability to penetrate cells, while retaining their specificity and multitarget recognition properties.

![Figure 1.4: Schematic representation of three stapled peptides.\textsuperscript{39}](image)

A novel strategy that combines the application of monoclonal antibodies and peptide-based therapeutics to a different level is to switch from the so far described ‘direct’ inhibitions of protein-protein interactions, to a nature-inspired strategy for ‘indirect’ blockade of protein interactions. This approach aims at (small) proteomimetic synthetic immunogens that are able to trigger the (natural or artificial) production of antibodies in organisms. Research towards synthetic \textit{therapeutic vaccines} can eventually increase our knowledge of immunological specificity and give us the opportunity to influence the immune system, by triggering the production of new (anti-peptide) antibodies that cross-react with the corresponding native proteins to block their activities.\textsuperscript{25} Peptide-based protein mimics have already been developed and used to serve as vaccines by immunizing animals to induce protective immunity against viruses and bacteria, like the foot-and-mouth disease,\textsuperscript{43} polio,\textsuperscript{44,45,46} influenza\textsuperscript{47} and hepatitis B,\textsuperscript{48,49,50} as well as against the cholera bacterium.\textsuperscript{51} However, although peptides may have the required complexity to be antigenic (i.e. have the ability to bind specifically with T-cell receptors or antibodies), their small sizes usually render them ineffective as immunogens (i.e. they are unable to induce an adaptive immune response for producing cross-reactive antibodies). Even if the peptides show complementary structural and chemical interactions with an antibody, their ability to give rise to an immune response depends on many other extrinsic factors.\textsuperscript{52} Fairlie and coworkers, for example, underscored the lack of peptide-based vaccines on the market today in a review article.\textsuperscript{53} From the 125 peptides that entered clinical trials to serve as vaccine for human use, only 30 made it to phase II, but not a single one was able to pass phase III. Several reasons for these failures were given by van Regenmortel,\textsuperscript{25} such as the over-confidence in the specificity of antibodies, the underestimation of the difference between antigenicity and immunogenicity and most importantly, the strong belief that continuous instead of discontinuous binding sites may be used in the design of vaccine candidates.
1.3 SYNTHETIC PROTEIN MIMICS

It is commonly known that efficient binding between constrained peptides and antibodies raised against a native protein is no guarantee for the peptide to produce antibodies that can in turn react with this protein. However, the chances increase when the peptide used for immunization has a similar conformation as the corresponding region within the protein immunogen. (Figure 1.5).

![Diagram showing different levels of protein binding sites: Linear Peptide mimicking a “Level 1” epitope, Continuous (Linear) Epitope, Continuous (Conformational) Epitope, Discontinuous (Conformational) Epitope, Scaffold, Looped Peptide mimicking a “Level 2” epitope, Double-looped Peptide mimicking a “Level 3” epitope.]

Figure 1.5: Three levels of protein binding sites. Mimicry requires a linear peptide (level 1), a conformational scaffolded peptide (level 2) or scaffolded discontinuous peptide fragments (level 3).

In this paragraph, epitopes of increasing complexity (level 1-3) are described that function as target for the design and synthesis of peptide-based protein mimics. First of all, level 1 linear continuous epitopes have served as target for synthetic peptides as antigen mimics that even resulted in the synthesis of linear peptides functioning as synthetic vaccines. When protein mimicry concerns peptide design based on level 2 epitopes, the application of smaller designer molecules (i.e. “scaffolds”) for constraining peptides into their natural conformation becomes inevitable. Scaffolds can be applied to mimic protein secondary structures. But by far the most challenging epitopes for protein mimics are discontinuous in nature, which can only be imitated effectively when the individual peptide fragments involved in the epitope are brought together and constrained properly, so that they correctly represent their native 3D-topology. Protein mimicry based on these level 3 epitopes still suffers from complex synthetic routes or lack of information about the precise fragments involved in the epitope and therefore new simplified strategies in this challenging field may open a route towards a better clarification of discontinuous epitopes and successful mimicry thereof.

1.3.1 Level 1: Linear Protein Binding Sites as Targets for Synthetic Protein Mimics

Occasionally, short linear peptides can mimic the binding properties of a protein, but this approach is largely limited to the unstructured N- and C-termini of proteins. Already in 1963, Anderer and coworkers showed that short C-terminal peptides derived from a viral
coat protein could generate antibodies against the protein.\textsuperscript{54} Despite this important observation, the work was not given the right attention, probably because this example involved a plant virus and the immunological response was therefore considered irrelevant for humans.

Almost two decades later, Anderer's work was followed by publications reporting the syntheses of level 1 peptide mimics relevant for newly identified bacterial and viral proteins. The Audibert\textsuperscript{55} and Beachey\textsuperscript{56} groups, for example, published independently the principle of using linear synthetic peptides to elicit antitoxin activities in bacterial infections. Audibert used a 14-residue peptide to generate protective antibodies against the N-terminal region of \textit{diphtheria} toxin when administered to guinea pigs while Beachey synthesized peptides that elicit protection against bacterial infections by showing the type-specific protective antibody production against \textit{Streptococcus pyogenes} by a 35-mer peptide. Other examples of bacterial antigen mimics were published by Jacob et al. for \textit{Vibrio cholera} toxin,\textsuperscript{51} by Harari et al. for \textit{Shigella dysenteriae} toxin\textsuperscript{57} and by Ogawa for \textit{Porphyromonas gingivalis}.\textsuperscript{58}

Later, the mimicry of viral proteins received increasing attention when compared to the bacterial proteins, as the structures of viral proteins were less complex than those of bacteria. Peptides able to induce antibody production against viruses were discovered and their design was based on either structurally disordered continuous loops or parts of discontinuous binding sites of viral antigens. The high mobility regions of the C- or N-terminal chains mostly functioned as targets, because their conformational resemblance with free linear peptide mimics resulted in antigenic cross-reactivity as expected.\textsuperscript{59,60}

A variable and continuous target used for vaccination is the \textit{Foot-and-mouth disease virus} (FMDV) protein VP1, clarified by X-ray crystallography.\textsuperscript{61} The binding site consists of a loop, known as the GH loop that contains the three-residue sequence Arg-Gly-Asp (RGD), which is highly responsible for binding of a virus onto a cell. Bittle\textsuperscript{62,63} and Dimarchi\textsuperscript{43} introduced such RGD-based peptide mimics and observed that with these peptides better neutralizing antibodies were obtained as compared to the intact VP1 protein, though the immunological activities of these peptide mimics were not as good as those of the inactivated viruses.

It was not until 1994 when Langeveld et al\textsuperscript{44} developed the first successful synthetic vaccine, based on the continuous N-terminus binding site of the VP2 protein from the \textit{canine parvovirus} (CPV). The linear peptide was shown to induce protection in target animals and accordingly, several groups started to explore the field of continuous binding site mimics. The Talwar group for example, showed that immunization with the C-terminal peptide of hCG (CTP-37), the pregnancy hormone, renders women temporarily infertile.\textsuperscript{65,66} Other examples using linear peptides as vaccines were found for the \textit{hepatitis B virus},\textsuperscript{48,49,50} \textit{respiratory syncytial virus},\textsuperscript{67,68} \textit{human papilloma virus},\textsuperscript{69,70} \textit{bovine leukemia virus},\textsuperscript{71,72} \textit{feline immunodeficiency virus}\textsuperscript{75} and \textit{hepatitis C virus},\textsuperscript{74} though examples of effective protein mimics based on level 1-type interaction sites are limited.
1.3.2 Level 2: Conformational Protein Binding Sites as Targets for Synthetic Protein Mimics

Protein structures determine protein functions. Because linear synthetic peptides do not spontaneously fold into the structure of the native protein, the idea of constraining peptides into their natural conformation using smaller designer molecules, called templates or scaffolds, was first introduced and applied by Hirschmann and co-workers in 1979.\textsuperscript{75} Mimicry of protein surfaces by peptide-based structures using scaffolds is a major focus of current research, as reviewed by Hamilton.\textsuperscript{36} By reducing the unfavorable entropy term for spontaneous folding of linear peptides and at the same time decreasing the enthalpy contribution by presenting the correct structure to a peptide, scaffolds can lower the total free energy and force a linear peptide to fold into a defined secondary structure.

The Kaumaya group nicely illustrated that continuous binding site mimics, like the constrained peptides of the cysteine-rich regions of the human epidermal growth factor receptor 2 (Her2), a well-known breast cancer antigen, showed better antibody responses than their linear counterparts.\textsuperscript{76} Similar attempts were made targeting the gp41 protein of human immunodeficiency virus 1 (HIV-1), using a heptapeptide derived from its external regions.\textsuperscript{77,78,79} Despite the fact that several constrained peptides had a higher affinity for the antibody 2F5 than the unconstrained heptapeptide, none of them was able to induce antibody formation with detectable neutralizing capacity when used as immunogen. These results again indicate the difference between antigenicity and immunological activity of peptides. Because the secondary structures of proteins (e.g. \(\alpha\)-helices, \(\beta\)-sheets, \(\beta\)-strands and \(\beta\)-turns) determine the three-dimensional topology and function of a protein, it is likely that these entities are essential for both antigenic and immunological effects. Therefore, several scaffolds have been developed over the past three decades that facilitate the formation of secondary protein structures in peptides.

\textit{Mimicry of \(\beta\)-Turn and \(\beta\)-Sheet Conformations}

In the late seventies, the group of Hirschmann demonstrated the concept of introducing a conformational rigid molecule such as a \(\gamma\)-lactam (1 in Figure 1.6) in the backbone to give a peptide with one single bioactive conformation.\textsuperscript{75} The \(\gamma\)-lactam induced a \(\beta\)-turn, resulting in a more active analog of the luteinizing hormone-releasing hormone than the parent protein, thus providing evidence for the bioactive conformation.\textsuperscript{80} These findings induced the development of secondary structure \textit{nucleators} as small molecules that are able to promote the folding of an attached polypeptide sequence, thus affording a monomeric \(\beta\)-turn, \(\beta\)-sheet or \(\alpha\)-helix in aqueous solutions.\textsuperscript{81} In the following years, several \(\beta\)-turn mimetics were developed, including the chemical structures shown in Figure 1.6.\textsuperscript{82,83,84} These examples illustrate the development from simple molecular turn inducers (1-4 in Figure 1.6) with simply two attached peptide chains (R\textsubscript{1} and R\textsubscript{2}), towards more complex tetrapeptide-based mimics (5-6 in Figure 1.6) that highlight the potential to incorporate both the natural amino acid side chains as well as several NH or C=O functionalities already within the template.\textsuperscript{85,86,87}
Figure 1.6: Scaffolds for mimicry of protein secondary structures: β-turn and β-sheet nucleators.

Apart from β-turn imitations, β-sheet structures form a more challenging target because the thermodynamic stability of the amino acid residues in a sheet is much more context dependent than in an analogous turn.88,89,90,91 Ideally, template-induced folding to give a β-sheet should be preceded by intramolecular self-association, affording a homogeneous β-sheet structure. However, intermolecular folding and self-assembly generally show comparable rates resulting in competitive formation of several heterogeneous high molecular weight β-sheet structures. Kemp and co-workers presented a controlled β-sheet nucleation combined with β-turn initiation, by incorporating an epindolidione moiety (7 in Figure 1.6) into a nonapeptide sequence.92,93

Figure 1.7: Example of β-sheet nucleation combined with β-turn initiation by hydrogen bond formation, using an epindolidione moiety (7) that favors a parallel β-sheet structure when two proline residues were incorporated, and an anti-parallel β-sheet conformation when two urea moieties were introduced.
This rigid template is not only able to form three intramolecular hydrogen bonds with the attached peptide chains. It was also decorated using turn inducers, like Pro-D-Ala residues, to favor the parallel β-sheet conformation (Figure 1.7). Similarly, introduction of two urea moieties afforded an anti-parallel β-sheet, indicating that templated β-sheet nucleation was indeed possible. Ever since, templates for β-turn induction were more commonly used to both inverse peptide chains and to promote intramolecular hydrogen bonding to result in spontaneous β-sheet formation. For example, the dibenzofuran-based template 8 (Figure 1.6) functions as a β-turn nucleator, whereby a linear chain of 13 residues (with an overall charge of +7) was folded into a β-sheet structure (shown by NMR and near UV CD measurements) even though repulsion between the positively charged residues was thought to be highly unfavorable for β-sheet formation.\(^{94}\) Similarly, a dibenzofuran-based diacid-functionalized scaffold was applied for β-sheet tertiary structure formation in aqueous buffers, inducing formation of an amphiphilic antiparallel β-sheet structure.\(^{95}\) Finally, bipyridine based β-turn inducers like 9 (Figure 1.6) have been described, based on the addition of a metal ion to switch from a transoid bipyridine conformation towards a cisoid form, also resulting in β-sheet formation (Figure 1.8).\(^{96}\)

![Figure 1.8](image-url)

**Figure 1.8:** Example of a bipyridine based β-turn inducer (9) illustrating the effect of the metal in the transition of transoid conformation into cisoid structure.

**Mimicry of Looped Conformations**

Another protein conformation is found in binding sites that consist of a combination of β-sheets and β-turns, or simply referred to as peptide ‘sheet-turn-sheet’ loops. This structure is for example seen on the membrane proteins of the *poliovirus*, consisting of four main proteins, VP1 to VP4 (Figure 1.9), whereby three of these proteins have similar structures that are built up from several loops aligned as eight antiparallel β-sheets. Cyclization of a peptide corresponding to these loops was found to be a minimum requirement for antibody-like affinities.\(^{97}\) Indeed, cyclized peptides showed enhanced immunogenicity and antibody generation as compared to their linear counterparts and thereby also revealed cross-reaction with the parent protein.\(^{98, 99, 100, 101}\) Peptide cyclization strategies include backbone-to-backbone or side chain-to-side chain cyclizations,\(^{102}\) for which reactions like olefin metathesis,\(^{103, 104}\) macrocyclization by conventional lactamation chemistry\(^{105}\) and the
Cu(I) catalyzed [2+3] cycloaddition were used. Other structural fixations of peptides are described using S₅Ar displacements or tryptophan zippers. Undesirably, all of these methods require either special protection/deprotection protocols or the addition of a catalyst to increase the reaction rates, severely limiting their applicability for binding site reconstruction. In 2005, the CLIPS™-technology for peptide cyclization was developed by Timmerman et al. (Figure 1.10).

\[ \text{Figure 1.10: The CLIPS reaction.} \]

CLIPS stands for Chemical Linkage of Peptides onto Scaffolds and the reaction has proven to be a powerful route for chemoselective cyclization of side chain unprotected dithiol-containing peptides. Peptides bearing cysteine residues at both ends of the linear chain are constrained using a scaffold that is based on one of the commercially available benzylic halide type compounds 1,3-dibromoxylene mT2 ("m" for "meta" bromide positions, "T" for "Template" and "2" for the number of bromides), 1,3,5-tribromomesitylene (T3) and 1,2,4,5-tetramodurene (T4) (Figure 1.11).

\[ \text{Figure 1.11: Structures of the commercially available CLIPS-scaffolds 1,3-dibromoxylene (mT2), 1,3,5-tribromomesitylene (T3) and 1,2,4,5-tetramodurene (T4).} \]

The cyclization reaction is so fast, that the linear intermediate is hardly ever detected by HPLC. Moreover, the technique is unique, because the reaction conditions are extremely mild (<15 min. at 25 °C, aqueous-buffered solutions, pH 7.8) and no side-reactions occur with other amino acid side-chain functionalities (NH₂, COOH, OH, imidazole, etc.). The technology was successfully applied for double-constrained peptides (using CLIPS in combination with disulfide bond formation), derived from the hFsh β-loop that generated antisera with high cross-reactivity to hFsh, where the corresponding linear peptides totally failed. In addition, the group of Heinis and Winter used the CLIPS technology in phage display libraries for the selection of ligands based on bicyclic peptides. Synthetic linear phage-encoded peptides with three cysteine residues were fused on the gene-3-protein and cyclized into double looped peptides using T3. The combined CLIPS/phage display technique yielded inhibitors against protein targets such as the kallikrein family serine proteases in the low-nanomolar range.
Mimicry of α-Helices and α-Helical Bundle Conformations

Finally, apart from β-sheet/turn and looped secondary structures, α-helical motifs are widely found in the 3D architecture of proteins in nature and are responsible for highly specific protein-protein interactions. The α-helical structures are the inspiration of significant efforts to modulate protein-protein interactions, with a particular focus on therapeutically interesting targets. Stabilizing peptides that bind their receptors in a helical conformation are expected to favor receptor binding by virtue of preorganization. Furthermore, the exposure of the polar amide backbone is reduced due to intramolecular hydrogen bonding associated with helix formation, increasing both protease resistance and membrane penetration. Kemp and co-workers developed the N-terminal helix template 10 (Figure 1.13), containing three carbonyl moieties oriented in such a way that they form a mimic of the pitch and spacing of a right-handed α-helix. When covalently attached to the N-terminus of a peptide sequence, the carbonyl groups form hydrogen bonds with the amide NH protons of the attached peptide sequence (Figure 1.12). Similarly, by incorporating a hydrazineethylene bridge that functions as a covalent replacement of a backbone hydrogen bond between the i and i+4 residues of an α-helix, α-helix nucleator 11 (Figure 1.13) was developed that mimics one turn of an α-helix.

More recently, Verdine and co-workers at Aileron Therapeutics studied the α-helical mimicry in an all-hydrocarbon cross-linking system that was able to greatly increase the helical propensity and metabolic stability of peptides, making them much better membrane penetrators and increasing their stability compared to their unraveled form. Ruthenium-catalyzed ring closing metathesis (RCM) was used to covalently connect O-allyl serine residues (12, Figure 1.13) located on adjacent helical turns in a peptide. These so called
'stapled peptides' retain the specificity and natural multitarget recognition capabilities of therapeutic proteins with very few limitations in their ability to address extracellular and intracellular targets.\textsuperscript{37,38,104} Other methods that have been reported to stabilize $\alpha$-helical macrocyclic peptides are comprehensively reviewed by Schneider and Kelly\textsuperscript{96} and include incorporation of salt bridges,\textsuperscript{117} metal chelates (such as 13 in Figure 1.13),\textsuperscript{118,119} amide bonds,\textsuperscript{120} disulfide bridges\textsuperscript{121} and lactam bridges\textsuperscript{122,123} to connect the i and i+4 position within an $\alpha$-helix. And with the technique to mimic single $\alpha$-helical structures in hand, Kaiser and Sasaki were challenged to develop a tetrasubstituted coporphyrin scaffold 14 (Figure 1.13) for the mimicry of binding sites that consist of several bundled $\alpha$-helices (Figure 1.14). An example is found for cytochrome P-450, whose porphyrinbased active site is surrounded by several $\alpha$-helices that serve as a hydrophobic pocket.\textsuperscript{124} The template successfully afforded a four-helix bundle with four identical linear peptides. As shown by CD studies, the coporphyrin proved to be responsible for the stabilization of the $\alpha$-helical structures.

Figure 1.14: The principle of metal-mediated helix bundling using the coporphyrin scaffold 14 (left), or the bipyridine-based scaffold 16 (right).

Later, DeGrado, Nishino and Mihara independently improved this technology by developing tetraphenylporphyrins, directing the formation of four-helix bundles that could insert into membranes to serve as ion channels\textsuperscript{125} or penetrate lipid bilayers.\textsuperscript{126} In addition, controlled helical bundle formation was independently achieved by Sasaki\textsuperscript{127} and Ghadiri,\textsuperscript{128,129} demonstrating that metal-mediated constraining using (bi)pyridine-based scaffolds (15-16 in Figure 1.13) also results in the formation of three or four-helix bundles (Figure 1.14). Consequently, most of the reviews\textsuperscript{7,23,130,131,132} that have been published on the topic of protein mimicking by imitating protein secondary structures all draw a similar conclusion: for the majority of protein binding sites, surface recognition will only be successful if it covers synthetic imitations of several structural resemblances over a large discontinuous surface area, because only in this way the required large number of weak interactions will be available to overcome the highly solvated character of the protein surface. The most critical feature will be a controlled balance of hydrophobic and charged regions of the discontinuous mimic to give high-affinity interactions with the target, without problems such as insolubility, aggregation, or nonspecific protein binding.
1.3.3 Level 3: Discontinuous Protein Binding Sites as Targets for Synthetic Protein Mimics

The increasing interest in the identification of synthetic molecules that bind to protein targets with similar affinity and specificity as antibodies has resulted in strategies that rely on mimicking the discontinuous properties of native proteins. Expressed in the nature and position of multiple peptide domains, the general aim in this field is to develop new ways to properly connect and orient multiple peptides on a single scaffold. In this respect, twenty years ago, Mutter and co-workers developed the so-called template-assembled synthetic protein (TASP) scaffold\(^1\) (17 in Figure 1.15). This scaffold allowed the introduction of different protecting groups on lysine residues, enabling the addition of five different peptide fragments to the scaffold and thereby covering a larger surface area of a protein discontinuous binding site. This templating effect of the TASP scaffold forced each element to adopt the three-dimensional structure of the native protein. Due to the possibility of selective functionalization of the lysine residues in the template, the TASP scaffold has also been called a regioselective addressable functionalized template (RAFT).

![Figure 1.15: Scaffolds for mimicry of protein tertiary structures.](image)

Although scaffold decoration requires a combination of cumbersome and complex solid- and solution-phase synthesis,\(^{133}\) the RAFT scaffold has been applied by the groups of Hauert and Haehnel to construct a mimetic of the von Willebrand factor (WWF) A1 domain\(^ {134}\) and a water-soluble cytochrome b model protein MOP\(^ {135}\)/Heme-Ru-MOP3,\(^ {136}\) respectively. Beyermann and co-workers used the RAFT template to mimic the extracellular binding domain (ECD) of G protein-coupled receptors (GPCR’s), by covalently connecting the N-terminal linear sequence (ECD1) and the three looped sequences (ECD2, ECD3 and ECD4) using a combination of recombinant, enzymatic and chemical synthesis strategies.\(^ {137}\) Further applications include sensors, carbohydrate clusters, cell and tissue targets, binding
loops and ion channels\textsuperscript{138} illustrating the versatility of applications of these promising types of scaffolds.

The introduction of the RAFT template was a trigger for further development in the field of protein mimicry and immunology. Fairlie and coworkers described a macrocyclic scaffold constrained by oxazoles and thiiazoles (18 in Figure 1.15) to support two peptide loops projecting orthogonally from the same face of the scaffold.\textsuperscript{139} This macrocyclic scaffold nicely mimicked the two inter-helical loops of cytochrome b\textsubscript{562}. Later, the group of Spadaro showed the use of a calix[4]arene scaffold (19 in Figure 1.15) for the synthesis of a novel fully synthetic cancer vaccine candidate in which by sequential amide coupling four S-Tn glycomimetic antigens and one P3CS immune-adjuvant moiety have been anchored at the two different faces (wide upper rim and narrow lower rim).\textsuperscript{140} The production of Tn specific IgG antibodies in mice was improved compared to an analogous monovalent construct, opening a route towards application in cancer immunotherapy. The group of Hamilton has used calix[4]arenes for protein surface recognition by preparing a series of unsymmetrical receptors in which two different loops are attached to the core calixarene,\textsuperscript{141} or by assembling four identical peptide loops mimicking the surface recognition of a platelet-derived growth factor (PDGF).\textsuperscript{142}

Other multifunctional scaffolds, like steroids (20 in Figure 1.15), were used as templates for the combinatorial addition of new receptor molecules.\textsuperscript{143} As recently reviewed by Jensen and Brask, rigid poly-functional carbohydrates (21 in Figure 1.15) were applied to combine multiple peptide fragments that fold into helix bundles.\textsuperscript{144} However, these constructs are relatively small and may have limitations when larger surface areas in proteins need to be covered. As a solution, Eichler developed other cyclic scaffolds such as 22 (Figure 1.15) that offer an additional benefit of enabling variation of scaffold ring size and, consequently, flexibility of the scaffold before attachment of two or three different peptide fragments.\textsuperscript{145} This resulted in the generation of new templated peptides presenting three HIV-gp120 protein fragments that comprise the primary contact residues for the discontinuous interaction with the CD4 receptor.\textsuperscript{146} Antibodies raised against this synthetic CD4-binding site mimetic were found to recognize the parent protein gp120, as well as to compete with the broadly neutralizing antibody mAb b12 whose binding site overlaps with the CD4-binding site for binding to gp120.\textsuperscript{147} The same epitope on gp120 was targeted by the group of Liskamp using the small Fmoc-o-NBS-Aloc protected triazacyclophane (TAC) scaffold\textsuperscript{148} 23 (Figure 1.15) for the mimicry of the HIV-gp120/CD4-receptor binding site.\textsuperscript{149} Very recently, this group also presented the use of the TAC scaffold for proper positioning of a $\beta$-hairpin loop together with an N- and a C-terminal peptide segment of cystatin B that are all involved in inhibiting the enzymatic activity of papain.\textsuperscript{150} The TAC scaffold was applied to make small triple-looped peptide libraries, forming several gp120 mimics by synthesizing all the possible combinations of three looped peptides on one scaffold.\textsuperscript{151}

Obviously, each of the discussed strategies above involves covalent constraining of the chosen peptide sequence through a series of synthetic modifications. Hamilton and coworkers described an alternative supramolecular approach using non-covalently stabilized macrocyclic scaffolds to which peptides were attached. Self-assembly of multiple adjacent peptide loops in order to orient and enforce their conformational restrictions gave rise to the formation of both homo- and hetero-combinations of two peptide loops on a surface.\textsuperscript{152} Due
to the non-covalent nature of this strategy, the formation of mixtures of homo- and hetero-dimers can never be prevented, which severely limits the supramolecular approach for therapeutic applications.

Until now, most of the examples of scaffolds used for discontinuous binding sites are based on a divergent synthetic strategy, using readily made central scaffolds bearing several reactive groups to attach peptides via a stepwise (either covalent or non-covalent) manner. Such approaches suffer from complex synthetic routes towards the final mimics due to the sequential protecting group manipulations and peptide couplings. Moreover, the peptides are sometimes limited in the use of (side-chain unprotected) amino acids or the peptide length, as well as the separation of hetero- from homo-dimers and the fact that the final mimics are often highly hydrophobic, thereby impeding purification.
1.4 STRATEGY AND OUTLINE OF THIS THESIS

The research described in this thesis aims for the mimicry of complex discontinuous protein binding sites by developing a straightforward route towards smaller synthetic mimics that can eventually be used as immunogens. For the development of discontinuous binding site mimics, attachment to a scaffold of several structurally constrained and naturally defined elements of a binding site is required, as well as a decent 3D-orientation of these fragments. In order to overcome the hurdles identified using a divergent approach, that is based on one central scaffold molecule to which multiple different peptides are being attached, which leads to multistep coupling procedures and difficult protection/deprotection cycles, we envisioned a novel convergent approach that fully relies on the orthogonal ligation of pre-functionalized and cyclized peptide fragments that simultaneously constitute the central scaffold. This novel methodology forms the basis of our synthetic strategy and circumvents troublesome deprotection and purification of partially or fully protected peptides with strongly hydrophobic nature. This strategy, that will eventually result in the design of successful epitope mimics, forms the central goal of this thesis and may be divided into two different objectives:

1) To develop a robust synthetic route towards double-CLIPS discontinuous binding site mimics, and
2) To investigate the minimal structural requirements needed for the successful mimicry of native protein binding sites in order to define a simple design of potential immunogens

As a result, this thesis is divided into two parts, to describe i) the development and optimization of a synthetic strategy towards the double-CLIPS constructs (chapters 2-3) and ii) the application of this technology to three model proteins (chapters 4-6) as shown in Figure 1.16.

**Figure 1.16:** Outline of this Thesis
The synthesis of double-CLIPS compounds is based on the use of functionalized scaffolds for cyclization and follow-up ligation relying on reactions that do not require protection of the peptide side-chain functionalities. Mutually reactive scaffold molecules that are suitable for the envisioned orthogonal ligation must be inert towards unprotected amino acid side chain functionalities and not interfere in the CLIPS cyclization reaction. In addition to this, the ligation reactions should proceed under physiological conditions and must be sufficiently fast to allow the use of stoichiometric amounts of (high molecular weight) peptides at relatively low concentrations (typically $\sim 1 \text{ mM}$). These boundary conditions prompted us to develop an easily accessible set of (water-soluble) scaffolds, the synthesis of which is described in chapter 2. Accordingly, chapter 3 presents the peptide synthesis and cyclization using the novel scaffolds, and also discusses the scope and application of four different orthogonal ligation methods to covalently connect two dissimilar CLIPS-cyclized peptides. These strategies include \(i\) the Cu-catalyzed, or \(ii\) the strain-promoted azide-alkyne cycloaddition reaction, \(iii\) the thiol-ene reaction, and finally \(iv\) the oxime formation reaction. Finally, the introduction of (a) disulfide bond(s) between the loops is realized to further constrain the mimic.

Once a straightforward synthesis of double-CLIPS peptides is at hand, the second objective of this thesis, i.e. to translate this into the synthesis of ‘real’ discontinuous protein mimics, is discussed in chapters 4-6. The minimal structural requirements needed to present the correct discontinuous 3D-structure is described in detail for a variety of discontinuous binding site mimics of human Follicle Stimulating Hormone (hFSH) in chapter 4. The hFSH protein is a member of the cys-knot protein family,\(^\text{153,154}\) that also includes the structurally related proteins \textit{human Chorionic Gonadotropin} (hCG) and \textit{human Vascular Endothelial Growth Factor} (hVEGF) (X-ray structures shown in Figure 1.17).

\textbf{Figure 1.17:} X-ray structures of hCG, hFSH and hVEGF. The alpha-subunit of hFSH and hCG is shown in green and their beta-subunit is shown in blue. For hVEGF, the two identical monomer units of this homodimer are colored green and blue. Cystein-knot units are marked in orange boxes.

It is likely to assume that an important epitope on these proteins consists of a combination of two looped entities (\(\beta_1\)- and \(\beta_3\)-loop) of the \(\beta\)-subunit (blue), thereby covering the highly discontinuous binding site of a number of anti-protein monoclonal antibodies. By combining several techniques (i.e. peptide arrays, binding ELISA, competition ELISA, CD measurements and sheep immunizations) we investigated numerous structural parameters
Chapter 1

that could influence epitope recognition of the mimics, and ultimately could generate cross-reactive antibodies. These parameters include the benefit of adding the β1-loop to the β3-loop, the best scaffold positions on the loops, the optimal lengths and residue relevance of the amino acids within the β1 loop, and finally the influence of (non)-native disulfide bonds to further constrain the loops and so enhance their 3D-topology. This information should eventually provide valuable knowledge for binding site mimic design. After exploring the boundaries of structural and topological requirements for hFSH-mimics, the new technology will be further applied and tested for targeting other discontinuous binding sites, that are either highly comparable to hFSH (i.e. hCG, chapter 5) or more challenging (i.e. hVEGF, chapter 6), in order to investigate the broad application of our novel technology. Finally, we present some preliminary results (Appendix to this thesis) on the synthesis of trifunctional scaffolds suitable for the production of triple-CLIPS mimics that allow the introduction of a third peptide fragment, and so enhance mimic recognition.
1.5 REFERENCES


