CMLS, Cell. mol. life sci. 53 (1997) 851–863
1420-682X/97/120851-13 $ 1.50 + 0.20/0
© Birkhäuser Verlag, Basel, 1997

CMLS Cellular and Molecular Life Sciences

Reviews

Non-native architectures in protein design and mimicry

M. Mutter and G. Tuchscherer*

Institute of Organic Chemistry, University of Lausanne, BCH-Dorigny, CH-1015 Lausanne (Switzerland), Fax +41 21 692 3955, e-mail: gabriele.tuchscherer@ico.unil.ch

Abstract. Protein design aims to mimic some of the structural and functional properties of native proteins [1–6]. The complexity of the folding mechanism, i.e. the pathway by which a linear polypeptide chain finds its unique 3D-structure, represents one of the most intriguing hurdles in this rapidly growing field. In order to bypass this well-known protein-folding problem [7–10], some years ago we proposed the construction of non-native chain architectures with a high propensity for folding [11–13]. According to this concept, termed TASP (template-assembled synthetic proteins), topological templates [14, 15] are used as a built-in device for directing covalently attached peptide blocks to a predetermined packing arrangement, resulting in branched chain architectures. Recent progress in the synthetic methodology for assembling peptides now allows us to access the full potential of the TASP concept. In this article, we discuss the state of the art of template-based protein de novo design, with special emphasis on progress in peptide synthesis and template design and show that some fundamental questions in protein assembly, structure and function can be approached by designing protein mimetics of reduced structural and functional complexity.

Keywords: Protein de novo design; template assembled synthetic proteins; locked-in folds; non-native architectures; topological templates; chemoselective ligation.

The template concept

The use of templates to direct organic synthesis has a longstanding history [16, 17]. More recently, topological templates have become a versatile tool in peptide mimicry (fig. 1), and their full potential is only now about to be recognized [18]. In view of the expanding areas of applications and functions, topological templates may be generally characterized as 'synthetic devices, that orient functional groups or structural units in well-defined spatial arrangements'. Typically, template molecules represent structural motifs such as constrained peptides, cyclodextrines or polycyclic systems disposing selectively addressable functional groups. The use of templates exhibiting a predetermined backbone conformation as host for the selective attachment of functional sites (e.g. amino acid side-chains or peptides) represents a conceptually new approach in molecular recognition studies and peptide mimicry. As depicted in figure 2, these topological templates disposing functional groups in spatially defined positions for interaction with an acceptor molecule, are ideal candidates to mimic bioactive conformations of peptide ligands [19–22] or protein surfaces, e.g. discontinuous epitopes, binding and catalytic sites. A number of most encouraging applications of this concept have been reported recently. For example, using β-D-glucose as a template ('scaffold'), Hirschmann et al. [23] were able to design a non-peptidic mimic of the somatostatin agonist SRIF, exhibiting a new functional profile. Covalent attachment of the amino acid side chains of Phe, Trp and Lys to D-glucose derivatives resulted in a molecule recognized by G protein-coupled receptors. The potential of this approach was demonstrated by the finding that subtle

* Corresponding author.

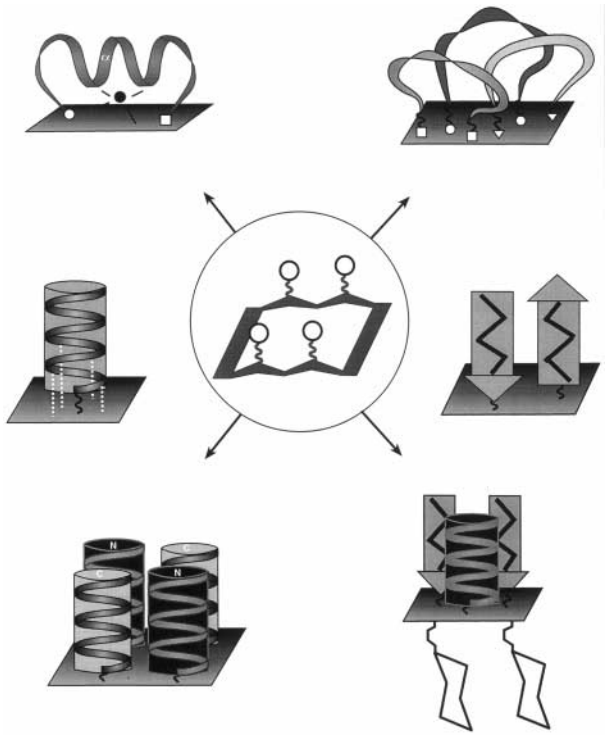


Figure 1. Template molecules are synthetic devices to induce and stabilize secondary structure formation, e.g. α-helices and β-sheets, and to enhance folding into a predetermined topology such as helical bundles, β-sheeted peptides or combinations thereof.

structural differences in the side chains attached to the template unexpectedly changed the biological profile of the individual peptidomimetics. In an alternative application of template molecules, Hruby and co-workers describe the use of tetrahydroisoquinoline derivatives as stable backbone templates ('topographical templates') for controlling the spatial orientation of amino acid side chains of bioactive peptides thought to interact with the receptor molecule [24]. By incorporating the template molecule at different positions of μ opioid receptor-specific octapeptides, the relation between conformation and dynamics to bioactivity could be delineated applying modern nuclear magnetic resonance (NMR) techniques.

We have recently proposed the use of topological templates for mimicking conformational epitopes of proteins ('surface mimetics') and bioactive conformations of peptides [19–21, 25]. As template molecules, constrained cyclic peptides (see fig. 3 and [25]) or stabilized secondary structure elements such as helices, β-sheets, loops or β turns proved to be versatile structural motifs for disposing selectively addressable groups. Alternative motifs such as cyclodextrins, glycosidic moieties or polycyclic systems might also be feasible templates for specific applications (fig. 2). Taking advantage of today's potential for chemoselective ligation methods [26–30] and protection techniques, topological templates carrying up to eight independently addressable sites are conceivable. The preparation of template-based peptidomimetics is achieved by a two-step approach:

1) The chemical synthesis of a topological template carrying orthogonally protected attachment sites,

2) the chemoselective ligation of amino acid side chain derivatives, functional groups or peptide blocks via covalent bond formation.

Template assembled synthetic proteins (TASP)

A most attractive approach with regard to their use in protein design is the induction and stabilization of short secondary structure blocks by synthetic devices (templates, fig. 4).

The use of conformationally constrained molecules as templates by geometrically fixing the first amino acid in the proper orientation for helix or β-sheet initiation is one way to bypass the entropically unfavourable nucleation step in secondary structure formation. The amphiphilic character of such stabilized helical or β-sheet peptide blocks is the prerequisite for self-association in solution and the major driving force for formation of more complex packing topologies characteristic of proteins.

Thus, DeGrado designed a membrane channel-forming α-helical peptide using only leucine (hydrophobic) and serine (hydrophilic) residues [31]; possibly the most consequent application of this 'amphiphilic principle' represents the design of polypeptide sequences with potential for 4α-helical bundle formation by use of a binary code [32] as a general design strategy.

Since the complex folding mechanism has yet to be unravelled, we proposed a conceptually different approach in protein de novo design to bypass the folding problem: the TASP concept [11–13]. As a key element, a topological template serves as a 'built-in' device to induce and reinforce intramolecular interaction of the covalently attached amphipathic peptide blocks, thus leading to well-defined packing topologies such as α-helical bundle or β-sheet TASP molecules.

Typically, template molecules represent structural motifs such as constrained peptides, cyclodextrines or polycyclic systems disposing selectively addressable functional groups. As prototype template molecules in the TASP approach, cyclic decapeptides derived from the antibiotic gramicidin S containing four lysines as attachment sites were used. As a second generation of this type of template, RAFT (regioselectively addressable functionalized template) molecules exhibit selectively addressable sites due to orthogonal protection techniques or unique chemical reactivity.

Various examples for template-induced tertiary structure formation have been reported in recent literature, e.g. the

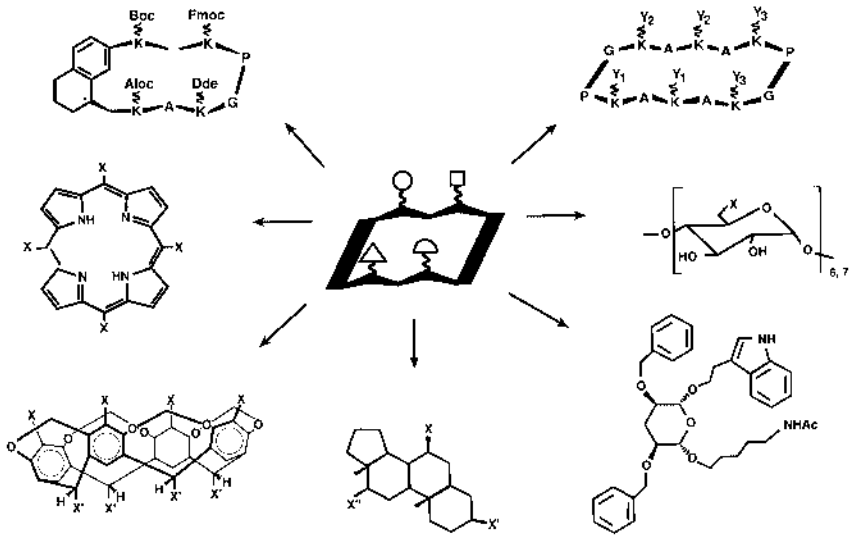


Figure 2. A variety of molecules can serve as templates, e.g. cyclic peptides, porphyrins, cyclodextrins, calixarenes etc.; the only requirement is an appropriate spatial orientation of the attachment sites.

use of the tetraphenyl porphyrine system as a template for the construction of a designed hemeprotein [31, 33, 34]. Another elegant approach to assemble helical bundles was followed by Ghadiri and Choi. using transition metals for the complexation of helices via N-terminal ligands. In a further step, a heterodinuclear three-helix bundle metalloprotein was synthesized with increased thermodynamic stability [35].

More recently, rigid organic macrocyclic scaffolds with an enforced cavity have been used for the orientation of a four-helix bundle protein; by varying the length of the spacer between peptide and template, the influence of flexibility/rigidity of the scaffold on stability can be investigated [36]. As a common feature, template molecules limit the degree of freedom of the attached α-helical or β-sheet peptides and thus promote intramolecular self-association to the desired folding topology.

Synthesis of TASP molecules

Two strategies have been applied for the synthesis of proteins: stepwise assembly from their constituent amino acids on a solid support (solid phase peptide synthesis, SPPS) [37], or convergent coupling of peptide segments in solution (fragment condensation) [38]. Although SPPS has been optimized to the extent that proteins of about 100 amino acids in length can be synthesized, accumulation of side products over the many coupling steps render purification of the target product laborious and time-consuming. Convergent strategies have the considerable advantage that synthesis and purification of peptide segments up to 30 residues in length is straightforward, but they are limited by the poor solubility of fully protected peptide segments and the tendency of α-carboxy-activated peptides to racemize.

Many of these difficulties can be circumvented by using recently introduced chemoselective ligation methods [26–30] which allow for the condensation of completely unprotected peptide fragments in aqueous medium (fig. 5).

Chemoselective ligation methods appear to be particularly useful in the TASP design. With respect to the elaborated protection chemistry in peptide synthesis, cyclic peptides as topological templates with up to four orthogonal protection groups for the lysine side chains as attachment sites are accessible. Selective cleavage allows for appropriate selective functionalization leading to RAFT [15] molecules as key compounds for the construction of TASP molecules of higher complexity. For example, TASP with up to four different helices using different chemoselective ligation procedures can be synthesized.

Although the utility of chemical ligation strategies was demonstrated for the synthesis of larger and more complex peptides/proteins [39, 40] and its sole limitation seemed to be the availability of larger, appropriately modified peptide segments, one fundamental problem has not yet been taken into account. The peptide fragments, designed to be amphiphilic with a high propensity for secondary structure formation, have a strong tendency in aqueous solution for self-association. Such high molecular weight aggregates are unfavourable in

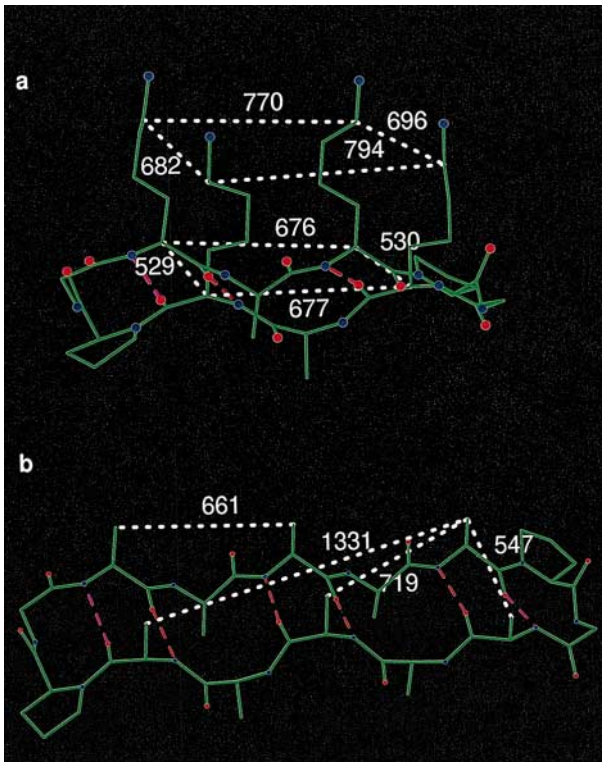


Figure 3. (*a*) Molecular modelling and NMR studies have shown that cyclic peptides (10-mer) as templates can adopt a low-energy conformation where the lysine side-chains as attachment sites are oriented toward the same face of the plane of the molecule, thus offering ideal distances for the attachment of e.g. four helices. (*b*) The ring size of cyclic peptides can easily be extended to a 14-mer, resulting in an increased number of attachment sites and range of accessible distances, e.g. to accommodate six peptide blocks. Depicted are the distances of the β C-atoms in pm.

the ligation process and result, for example, in extended reaction times or incomplete reaction. To prevent self-association of unprotected, secondary structure-forming peptide sequences during chain assembly, pseudo-prolines (ψPro) [41] have been introduced as a powerful tool to modify temporarily the intrinsic properties of peptides that are responsible for aggregation and secondary structure formation (fig. 6). Pseudo-prolines consist of serine- or threonine-derived oxazolidines and cysteine-derived thiazolidines and are obtained by reacting the free amino acid with aldehydes or ketones.

Due to the presence of a cyclic system (fixed ϕ-angle) in addition to the preference for a cis-amide bond [42] with the preceding residue, the incorporation of a ψPro moiety results in a kink conformation, thus preventing peptide aggregation, self-association or β-structure formation.

Consequently, pseudo-prolines fulfil two functions simultaneously: they serve (i) as temporary protection for Ser, Thr and Cys and (ii) as solubilizing building blocks to increase solvation and coupling rates during peptide synthesis and in subsequent chain assembly. Finally, ring opening of the ψPro by strong acid results in the completely deprotected peptide, restoring the regular Ser, Thr, Cys side-chains, and the molecule can adopt the designed topology.

Functional TASP molecules

Several authors have reported on selective membrane channel-forming TASP molecules using topological templates to define and orient membrane-spanning helical segments. For example, DeGrado et al. [31] synthesized a four-helix bundle proton channel using a tetraphenylporphyrin system as template, and Mutter et al. [43] designed and synthesized a template-assembled channel-forming protein derived from the bee venom melittin. The membrane channel-forming TASP molecules exhibit single channel conductance, ion selectivity and high thermodynamic stability as striking common features.

Numerous activities in the template design of functional molecules are being observed in the field of immunology. For example, Tam established the 'multiple antigenic peptide' (MAP) [44] approach using branched oligolysines as a template for the attachment of antigenic peptides. Here, the template acts merely as a support to increase immunogenicity rather than as a structure-inducing device. Similarly, Rose used multiple oxime ligations to make a totally synthetic macromolecule of controlled structure and molecular weight in the protein range (MW $\sim$ 19,000 D) [39]. Another chemoselective ligation method was used by Kent for the total synthesis of a four-helix TASP molecule [26] where the fragments to be joined had complementary reactivity (e.g. a thioacid and a bromoacetyl function).

As demonstrated for a number of model peptides, the combination of orthogonal protection techniques for resin cleavage and side chain protection with chemoselective ligation reactions represents a versatile tool for constructing TASP molecules of high structural complexity [26, 43]. Furthermore, with the introduction of secondary structure disrupting, solubilizing techniques [41, 42, 45–47], molecules which were previously inaccessible have become accessible.

We have synthesized a TASP with an antiparallel 4α-helical bundle topology [30] for the evaluation of the overall stability of a 4α-helical bundle in a parallel versus antiparallel arrangement. The modified helical segments derived from hen eggwhite lysozyme [48] have been covalently attached to a selectively addressable template by sequential ligation to the template via oxime bond formation. As followed by HPLC, the

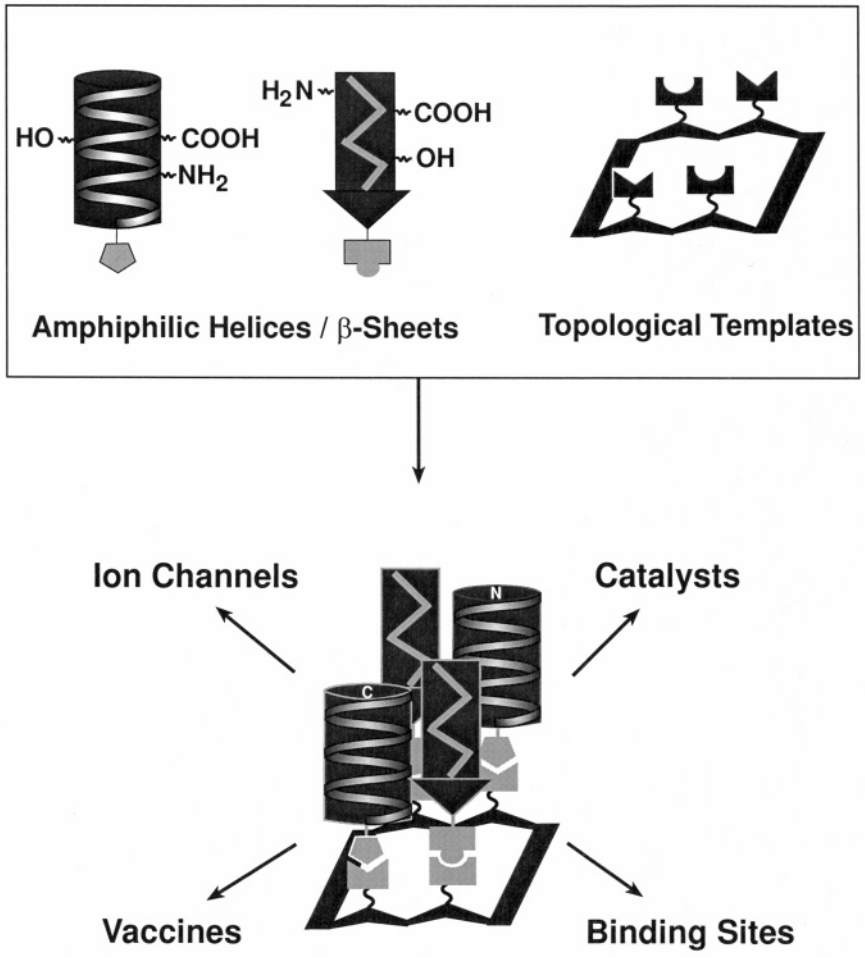


Figure 4. The concept of template-assembled synthetic proteins (TASP): topological templates induce folding of covalently attached peptide blocks into predetermined packing arrangements.

oxime bond formation proceeded under mild conditions to completion; interestingly, the ligation of the last two helical blocks did not significantly interfere with the prior attached two helices. Circular dichroism (CD) spectroscopic investigation revealed high helical content and a significant increase in secondary structure formation due to template-induced long-range interactions of the attached helices. Surprisingly, the parallel arrangement of the helices resulted in higher thermodynamic stability of the TASP compared with the antiparallel arrangement. Amphiphilic β-sheet-forming peptides of the type $(A\text{-}B)_n$; A = Ser, Thr; B = Ala, Leu, Val and their assembly to a topological template represent an even more challenging example of the solubilizing effect of pseudo-prolines and the versatility of chemoselective ligations [49, 50]. As prototypes for this interesting class of peptides, we have synthesized a series of peptides with the repetitive sequence $(Xaa\text{-}Ser)_n$. Applying standard protection chemistry, we were not able to achieve quantitative reactions after passing the critical chain length for β-sheet formation ($n \sim 3$–5). However, insertion of two ψPro-protected residues resulted in complete disruption of aggregates throughout the synthesis of the peptide, which was reflected in high coupling yields [49].

The ψPro-containing peptide was readily soluble in water and various organic solvents, and the infrared (IR) spectrum indicates that it adopts a random coil conformation (amide I band at $\sim$1650 cm^{-1}). After ring opening the peptide was insoluble in water and most organic solvents, and the IR spectrum shows the typical bands of a β-sheet (e.g. amide I band at $\sim$1620 cm^{-1}). The partially deprotected ψPro-containing peptide was chemoselectively ligated to a maleimide-functionalized topological template to give a 4β-bundle TASP. As followed by HPLC, the ligation reaction of the peptide to the template via

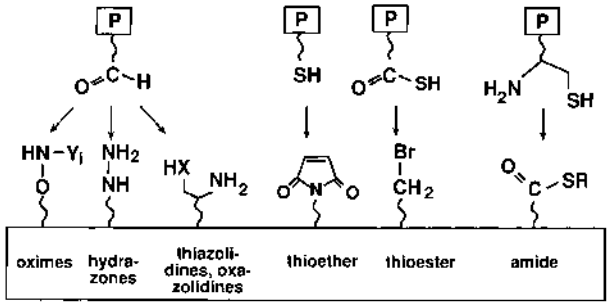


Figure 5. Recently introduced chemoselective ligation methods allow for the condensation of completely unprotected peptide fragments in aqueous solution.

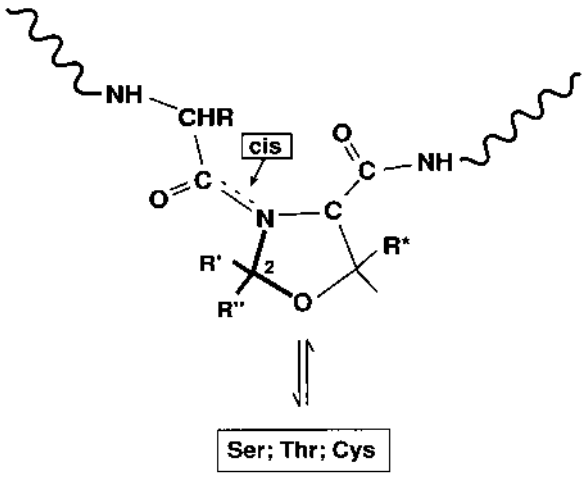


Figure 6. Ψ-prolines derived from serine (Ser), threonine (Thr) or cysteine (Cys) induce a kink in the peptide backbone due to the formation of a cis-amide bond with the preceding amino acid, thus disrupting any regular conformation resulting in an enhanced solvation of the peptide chain in the course of the synthesis or ligation process. The substituents at C2 (R′, R″) are derived from the aldehyde component. Ring opening with acid restores the regular amino acid.

thioether formation proceeded to completion within less than 4 h, indicating the absence of aggregation or β-sheet formation during the ligation process.

Similar effects were observed in the synthesis of a membrane channel-forming peptide. Here, the ψPro unit was inserted to induce a reversible kink in the helical peptide [50]. Despite the hydrophobic character of the peptide, the presence of the ψPro resulted in good solvation, and the subsequent coupling steps proceeded to completion as followed by HPLC. The peptide was cleaved from the Rink amide resin under acidic conditions, thus preserving the oxazolidine ring structure of the Thr($\psi^{H,H}$pro) residue. As indicated by CD and ATR-IR Attenuated Total Reflexion Infrared spectroscopy studies, the ψPro indeed distorts the helix to some extent. The helical transmembrane peptides are subject to chemoselective ligation to topological templates to access membrane-active TASP molecules with well-defined three-dimensional (3D) structures.

Separating structure and function: template-assembled binding loops

In conceptually separating structural and functional domains of native proteins, we have recently proposed the use of topological templates as scaffolds for the assembly of receptor binding loops (see fig. 7 and [51]). This extension of the TASP concept comprises two key elements: (i) peptide sequences (loops) containing C- and N-terminal functional groups (sticky ends) for chemoselective ligation and (ii) topological templates for the regioselective assembly of the loops.

Choosing from a variety of orthogonal protecting groups for peptide synthesis, we have developed a set of cyclic peptides as templates for the selective attachment of loop sequences [14, 15]. In tailoring the flexibility of the backbone and side chains, this template motif may dispose any desired number and geometry of attachment sites. The stepwise condensation of peptide sequences (loops, Li) onto these topological templates Ti (table 1) is achieved by amide, oxime and thioether formation [26, 28, 49, 52] or combinations thereof.

In strategy A (scheme 1), the C-terminally activated (step 1) fully protected loop Li is fixed to the template T, which contains one single reactive site after removal of Yi (step 2, amide 1). Selective deprotection of the N-terminal amino group of Li (step 3) and activation of the carboxylic group on the template (step 4) completes the cyclization of the loop (amide 2). Alternatively, the selective attachment of the N-terminal end may proceed via the formation of an oxime bond after the aminooxy group on the template has been deprotected (step 2 in strategy B). The condensation of peptides containing the same functional groups at both chain ends, e.g. carboxyl or aldehyde groups (strategies A′ and D, respectively), proceeds in one step and results in a mixture of two isomers with differently oriented loops. These strategies are very suitable for preparing TASP compound libraries for functional screening. Chemoselective ligation procedures allow for the condensation of completely unprotected peptides in aqueous solution or solvant mixtures (strategies C and D). Here, the regioselective attachment of a loop peptide to the template in one single condensation step (strategy C) can be achieved by the combination of two orthogonal ligation techniques (e.g. oxime and thioether formation).

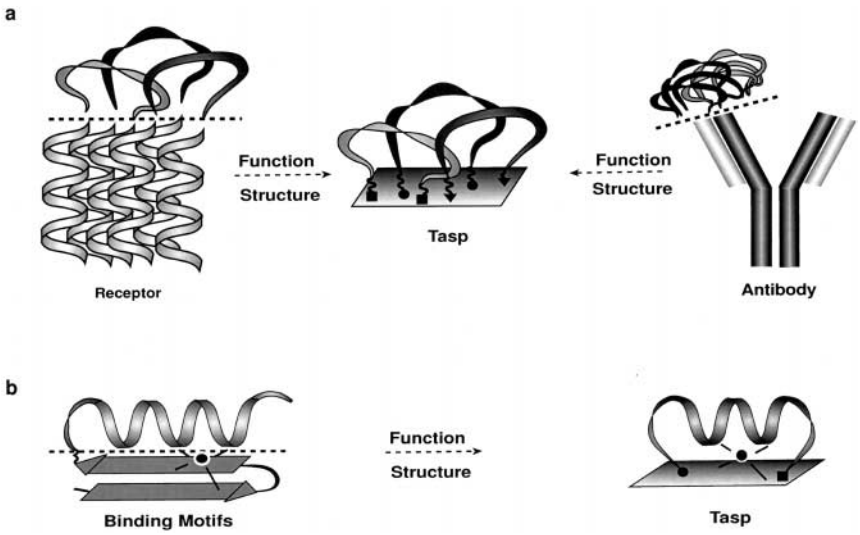


Figure 7. Separating structure from function: (*a*) The functional part of a protein, e.g. binding loops of a receptor or an antibody, is attached to a synthetic template molecule substituting for the structural domain of a native protein. (*b*) The helix of the characteristic $\beta\beta\alpha$-folding motif of a zinc finger is covalently attached via both chain ends ('locked-in') to an underlying β-sheet; thus the conformational space is reduced considerably.

Starting from a pool of n orthogonal amino or carboxyl protecting groups, up to n different loops can be selectively fixed according to the individual strategies depicted in scheme 1. By extending the palette of orthogonal protecting groups and combining various ligation techniques, TASP compounds of higher structural and functional complexity (including TASP compound libraries) become accessible.

As a first example of a potentially ligand-binding TASP compound (metal, substrate, antigen, transition state analogue) peptides derived from the CDR (complementarity-determining region) loop sequences of the phosphorylcholine-binding monoclonal antibody McPC603 [53] were covalently attached to a topological template (fig. 3, Tasp I in scheme 2). Starting from a cyclic peptide with pairs of selectively addressable reactive sites, the stepwise condensation of three peptides by amide bond formation according to strategy A′ provided Tasp I in high yield and purity. The CD spectra show the characteristic features of a loop cluster conformation [14, 53, 54].

The regioselective fixation of loops (strategy B) for TASP compounds with defined chain topology was illustrated for the design and synthesis of a potentially metal-binding two-loop TASP (Tasp II in scheme 2).

These prototype TASP compounds proved to be readily soluble in aqueous buffer solutions and polar organic solvents, and thus allowed a variety of biochemical investigations of the structure. Preliminary investigations of the conformational and binding properties support the hypothetical structures of Tasp I and II as suggested by molecular modelling studies (fig. 8).

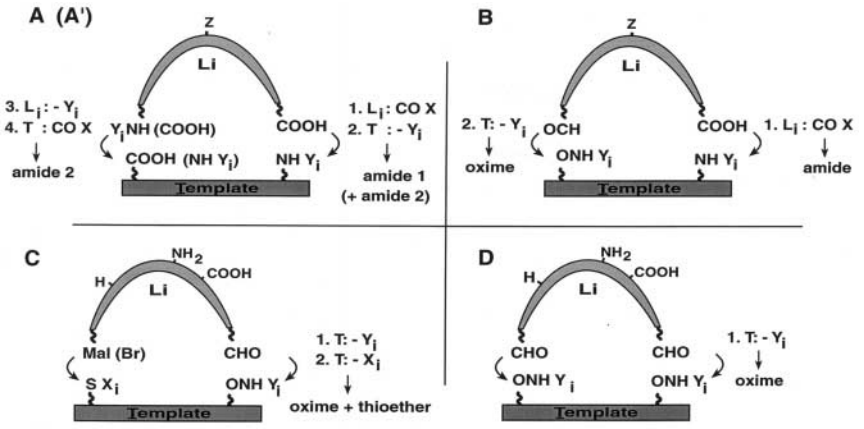
The strategies presented provide a synthetic entry to a new generation of TASP compounds, thereby extending the template concept [5] in protein design and mimicry. Most notably, the stepwise condensation of peptide loops to regioselectively addressable templates may be performed in aqueous solution and proceeds to completion within hours. The envisaged reattachment of the topological templates to the solid support will further simplify the multistep synthesis of differential loop

Table 1. Building blocks for the synthesis of Tasp molecules[a].

Templates T
T1: c[K(Y_1)C(Acm)K(Y_1)PGK(Y_2)AK(Y_2)AK(Y_3)PGK(Y_3)A]
T2: c[K(Y_4)PGK(Y_1)AK(Y_2)PGK(Y_3)A]
T3: c[K(Y_5)PGK(Y_6)AK(Y_2)PGK(Y_3)A]
T4: c[K(Y_6)PGCDRKK(Y_6)PGFACA]

Loop sequences L
L1: Suc-K(Boc)GY(^{t}Bu)NG-OH
L2: Suc-FGLY(^{t}Bu)G-OH
L3: Suc-E(^{t}Bu)LGR(Pmc)G-OH
L4: Boc-S(^{t}Bu)H(Trt)AGH(Trt)G-OH
L5: Aloc-S(^{t}Bu)H(Trt)AGH(Trt)G-OH
L6: X_1-HPGHK(X_2)G-NH_2
L7: X_2-FSRSDELTRHIRIHTGK(X_2)G-OH

[a]Abbreviations: Y_1: Boc; Y_2: Dde; Y_3: Aloc; Y_4: Fmoc; Y_5: $COCH_2CH_2S$-Trt; Y6: H_2NOCH_2CO; Suc: $HOOC(CH_2)_2CO$; X_1: maleoyl-β-Ala; X_2: OCHCO; A: Ala; C: Cys; D: Asp; E: Glu; F: Phe; G: Gly; H: His; I: Ile; K: Lys; L: Leu; N: Asn; P: Pro; R: Arg; S: Ser; Y: Tyr; Acm: acetamidomethyl; Trt: Trityl; Pmc: 2,2,5,7,8-pentamethylchroman-6-sulfonyl.

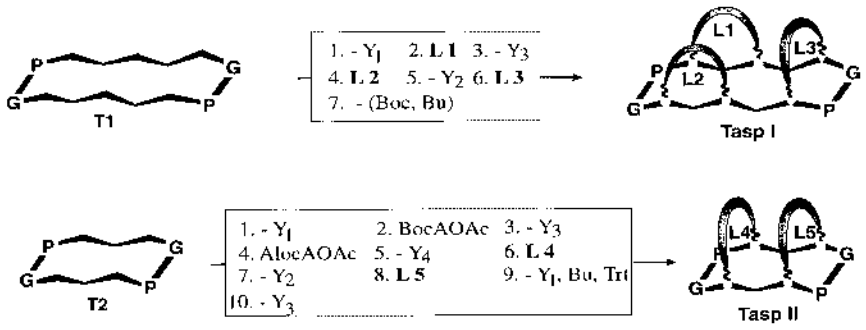


Scheme 1. Strategies *A–D* for the attachment of peptide loop sequences (**Li**) to topological templates (**T**); $\mathbf{Y}_i$ and $\mathbf{Y}_j$ amino protecting groups of the attachment sites; Z: side-chain protecting groups of Li; A/A′: **L1–L3: T1;** B: **L4**, **L5**, **T2**; C: **L6**, **T3**; D: **L7**, **T4** (see table 1).

cyclizations for generating TASP libraries. The state of the art in chemoselective ligation and orthogonal protecting techniques determines the number of selectively attachable loops and thus the complexity of the resulting TASP compounds. However, by exploiting the immense repertoire of synthetic organic chemistry for the incorporation of functional groups (sticky ends) into peptide chains, the scope of the present approach will be expanded rapidly. For example, ligand-directed assembly of helices, β-sheets and loops onto tailor-made templates represents a new tool for studying supramolecular assembly and molecular recognition processes.

Locked in tertiary folds as protein mimetics

The construction of protein-like folding motifs as structurally stable scaffolds for the introduction of 'function' now represents a major goal in protein design. As shown above, the use of topological templates allows us to bypass the notorious folding problem of linear polypeptides and offers a way to mimic native packing topologies by the template-directed self-assembly of helical and/or β-sheeted peptide blocks. The most consequent application of these concepts would be the construction of protein-like motifs featuring a fully branched chain architecture. These synthetic constructs, termed 'locked-in tertiary folds' (LIF), rely on recent progress in the methodology of peptide assembly and exhibit some unique conformational properties. As depicted in figure 9, the folding to a predetermined 3D structure is achieved by covalently linking the constituent peptide blocks via templates or spacer molecules to multibridged tertiary folds according to the principles of a molecular kit. The large number of alternative folding pathways



Scheme 2. Synthetic strategies for the assembly of Tasp I (fig. 2d) and Tasp II (fig. 2c). Tasp I: L1-L3: T1; Tasp II: L4: L5: T2 (see table 1).

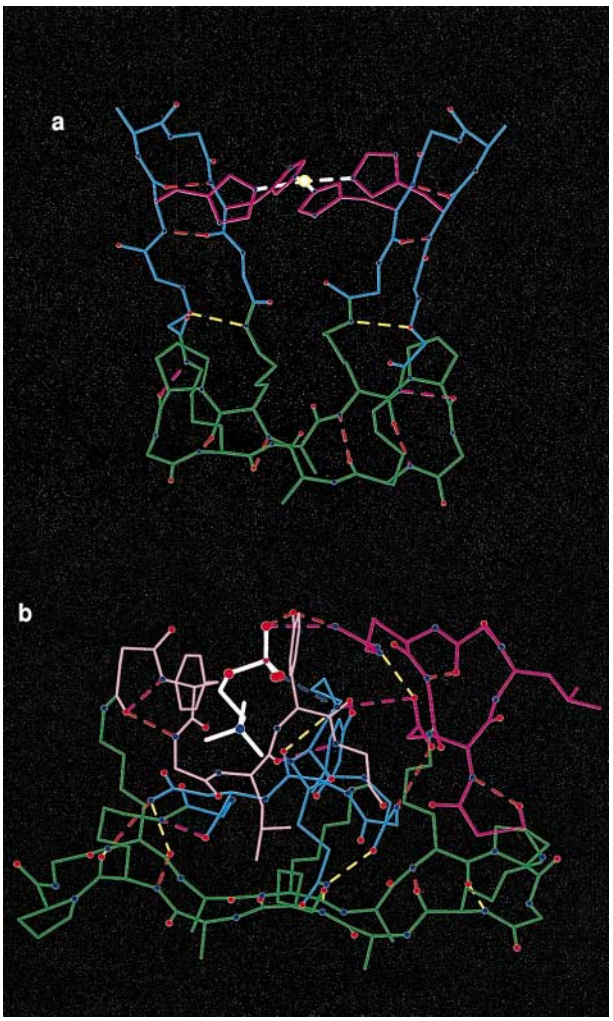


Figure 8. (*a*) Two-loop Tasp II (scheme 2) featuring a square planar metal-binding site formed by two identical HisAlaGlyHis-Gly sequences assembled on the ten-membered cyclic β-sheet template (**T2**). (*b*) One out of eight possible directional isomers of the three-loop Tasp I (scheme 2) approaching the antigen-binding site of the phosphorylcholine-specific antibody McPC603 [53] with the binding cavity in suitable shape to host the amphiphilic zwitterionic species.

encountered in the folding process of linear polypeptides is drastically reduced due to the constrained conformational space. The concept has been evaluated for the example of a LIF derived from the zinc finger (Zif) motif [55]. The DNA binding zinc finger proteins contain one of the most interesting structural motifs for studying molecular recognition processes and for modulating the DNA-binding specificity [56, 57]. The consensus sequence

Xaa_3-Cys-Xaa_{2-4}-Cys-Xaa_{12}His-Xaa_{3-4}-His-Xaa_4

(Xaa representing any amino acid) has been shown to fold in the presence of Zn^{2+} into a $\beta\beta\alpha$-folding unit which recognizes by its helical face a three base pair subsite in the DNA. Multimeric zinc finger proteins, consisting of several Zif modules, are considered to play a key role in controlling gene expression. The construction of Zif modules with novel DNA binding properties represents a major goal in protein design dealing with this motif. One way to achieve this is the systematic modulation of the DNA binding residues in preserving the consensus residues needed for Zn^{2+} complexation and folding. Surprisingly, this approach tolerates the substitution of a large range of residues within the linear polypeptide without losing its DNA binding capacity [58–60].

In applying the general concept of LIF (fig. 7), the framework of the Zif motif (fig. 7b) immediately suggests the use of a molecular kit system consisting of a cyclic β-sheet peptide template as a mimetic of the β-strand-turn-β-strand motif and a helical block featuring the corresponding attachment sites ('sticky ends'). For assembling these building blocks to a LIF (LIF 1), the helix is covalently attached at both chain ends to the cyclic 14-mer template via non peptidic linker groups. Super position of the X-ray structure of DNA-bound Zif 268 and the conformationally relaxed LIF 1 reveals a nearly perfect match of the two molecules (fig. 10), pointing to the general validity of our approach [55, 61]. Most notably, by proper design of the template, the zinc complexation site could be readily accommodated in the LIF 1 molecule.

The chemical synthesis of LIF 1 was achieved by preparing the individual building blocks, i.e. the template molecule and the helical segment by stepwise solid phase synthesis [36]. After cleavage from the resin and HPLC purification, the side chain deprotected peptides (containing chemoselectively addressable groups as attachment sites) were reacted in aqueous solution by a two-step condensation of the helical 18-mer peptide to the cyclic template via oxime bond formation. Most notably, the final chemoselective ligation step (C-terminal attachment of the helix) proceeded very fast and in high yields as monitored by analytical HPLC.

As suggested by molecular modelling studies (fig. 10), LIF 1 preserves the essential structural and functional features of the native Zif molecule, as documented by the absorption and CD spectra, and by the Zn^{2+}-complexation properties (fig. 11). Contrary to the native linear polypeptide [56, 57], LIF 1 retains its secondary structure elements even in the absence of the zinc complex [55].

As seen from the CD curve, Zn^{II} complexation results only in a slight increase in helicity. It is interesting to note that Zn^{2+}-independent folding has been described for a Zif-derived linear $\beta\beta\alpha$-motif after iterative redesign of the native sequence [62]. In the case of LIF 1, the onset of tertiary interactions between helical block and template induced by C- and N-capping of the helix results in helix stabilization and in a higher overall

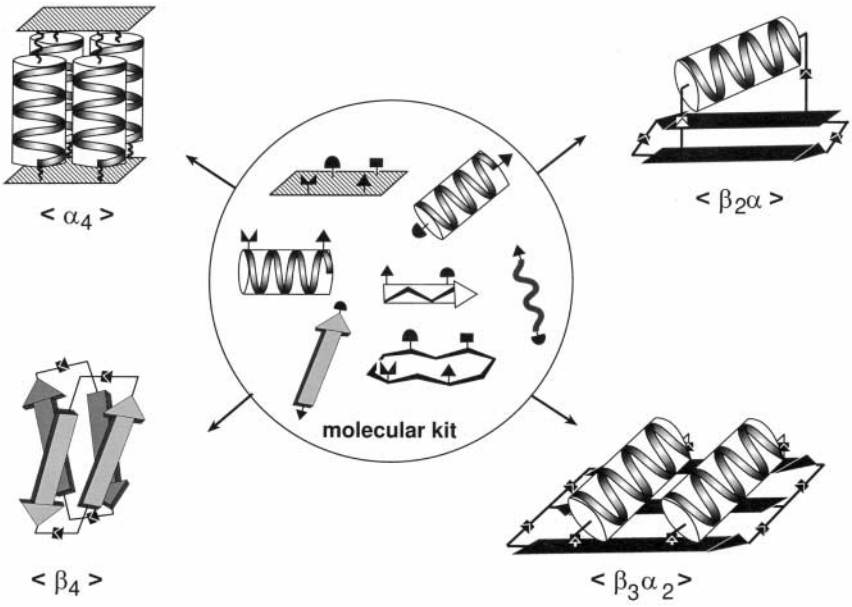


Figure 9. The construction of locked-in tertiary folds is based on the principles of a molecular kit. Secondary structure elements with 'sticky' ends are covalently attached via both chain ends to appropriately functionalized templates to result in, for example, locked-in four helical bundles, $\beta\beta\alpha$- or more complex folding topologies.

thermodynamic stability of the Zif motif as documented by the denaturation behaviour [55].

According to our general strategy of sequentially reducing the structural complexity of native proteins, we focus on downsizing the Zif motif to its minimal structural and functional elements applying the locked-in strategy. To this end, we have designed a second generation of zinc finger mimetic in reducing the length

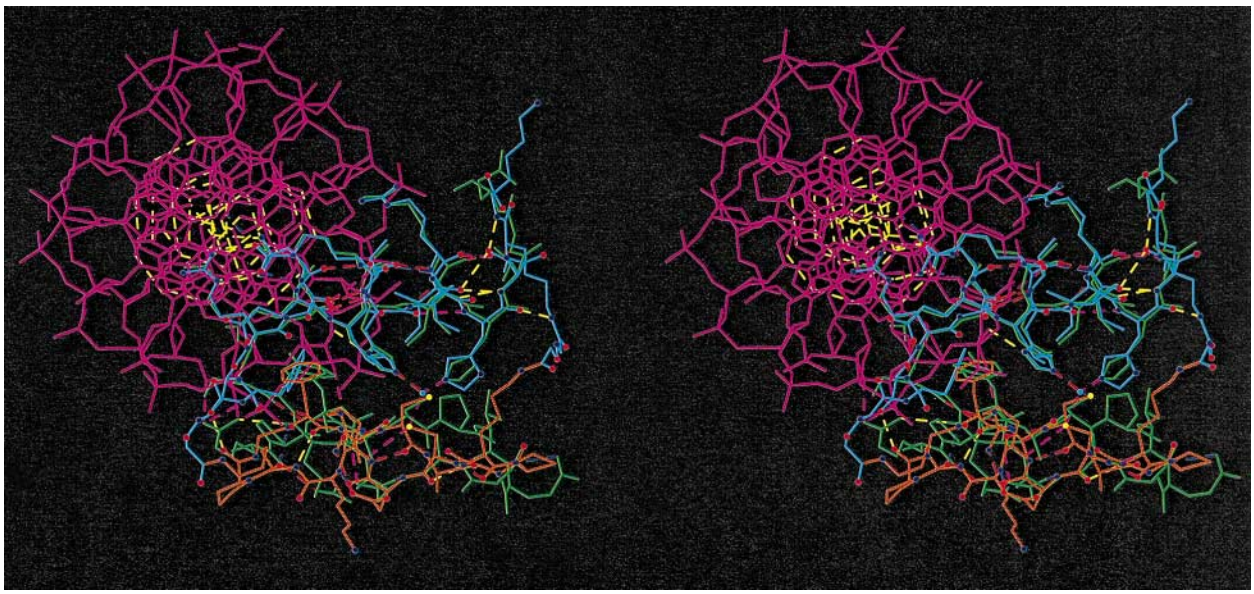


Figure 10. Superimposition of a locked-in $\beta\beta\alpha$ zinc finger mimetic (LIF 1 template in orange, helix in blue) with the X-ray structure of finger 1 (green) of Zif 268 in the DNA (magenta) bound state reveals a nearly perfect match.

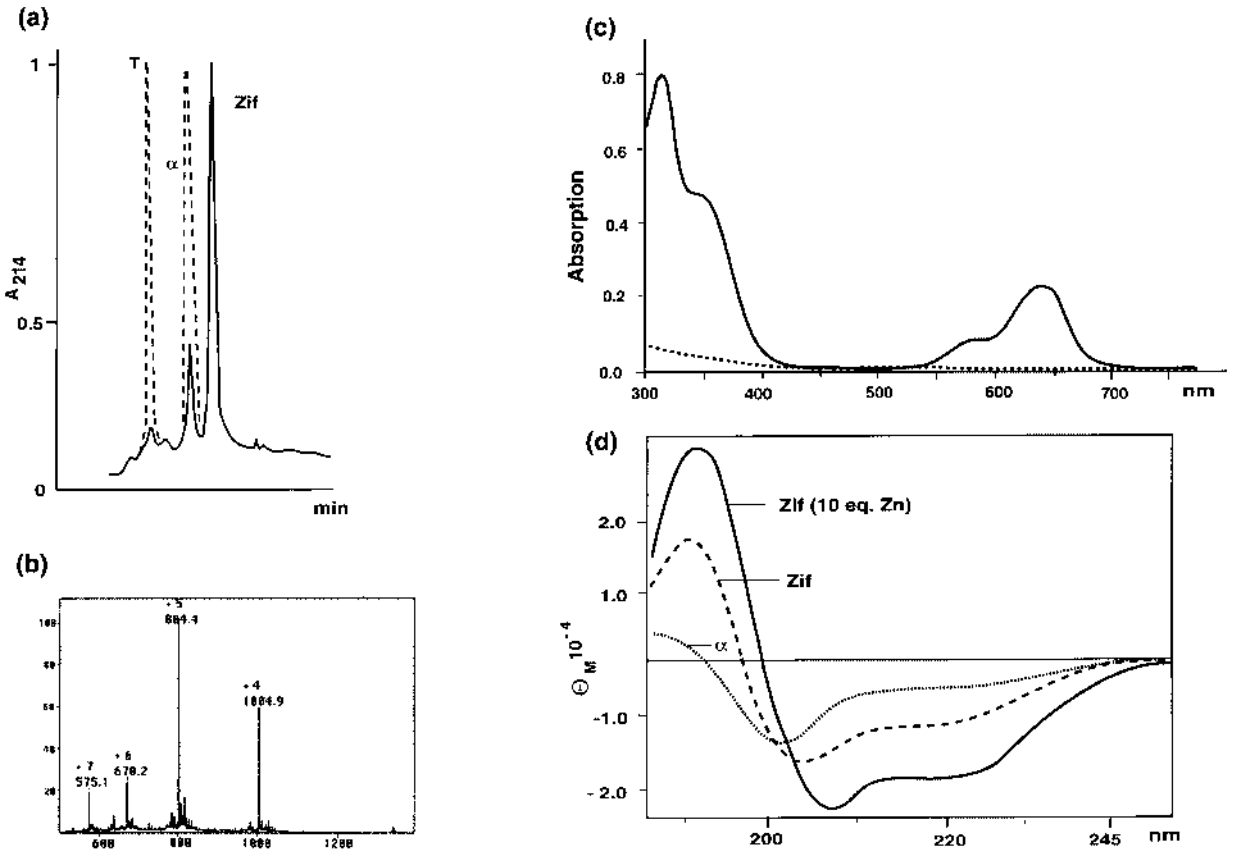


Figure 11. (*a*) The condensation of the helix to the template was followed by RP-HPLC (the dashed lines indicate the starting material; the helix was used in a slight excess) and (*b*) the chemical integrity of the resulting Zif mimetic was confirmed by ES-MS. (*c*) The UV spectrum obtained upon complexation with Co^{2+} shows the characteristic absorption bands of natural zinc finger molecules (solid line). Addition of Zn^{2+} readily displaces Co^{2+} from the complex (dashed line). (*d*) CD studies have shown that the 'locked-in' Zif mimetic retains significant helicity even in the absence of the zinc complex (dashed line). The spectra of the the Zn-complexed Zif mimetic (solid line) and the helix sequence alone (dotted line) are shown for comparison.

of the helical block (functional part) as well as its underlying β-strand-turn-β-strand template mimetic (structural part). This minimalistic locked-in Zif retains its metal-binding properties as monitored by the UV spectra (fig. 12). Due to the weak DNA binding capacity of single Zif molecules, the construction of multimeric locked-in Zif molecules applying chemoselective ligation methods is presently a major objective in our laboratory. The present approach offers some unique features:

(1) The conformationally constrained template molecule serves as a stable scaffold mimicking the β-strand-turn-β-strand motif of Zif.

(2) The variable helical block is readily accessible by solid-phase peptide synthesis, and the incorporation of non-proteinogenic building blocks offers a wide range of structural and functional modifications.

(3) The assembly to tertiary LIFs is achieved by a single condensation step, enabling easy access of a large number of analogues, including LIF libraries by stepwise SPPS!

(4) The synthesis of multimeric LIF modules exhibiting new DNA binding specificities should become a powerful tool for studying DNA-protein interactions.

Conclusions

The further elaboration of today's methodologies for peptide synthesis will rapidly expand the scope of template-based protein design. Using these synthetic tools, the protein folding problem no longer represents a substantial hurdle for accessing protein-like packing topologies. In stepwise reducing the complexity of native proteins, this hierarchical approach opens the way for rational design of protein mimetics with a pivotal role in drug design.

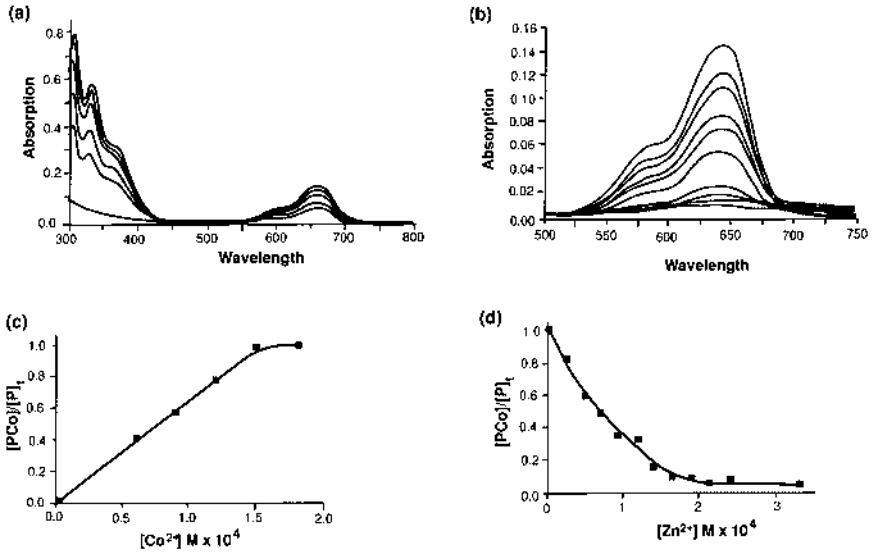


Figure 12. Complexation of locked-in Zif molecule. (*a*) UV spectra upon titration with increasing amounts of Co^{2+} in the range 300–800 nm. (b) Back titration with Zn^{2+} in the range 500–750 nm. (*c*) Concentration dependence of complex formation as a function of Co^{2+} concentration. (*d*) Zn^{2+} concentrations for the displacement of Co^{2+} from the complex.

1 Regan L. and DeGrado W. F. (1988) Charactrization of a helical protein from first principles. Science **241:** 976–978
2 DeGrado W. F., Wassermann Z. R. and Lear J. D. (1989) Protein design, a minimalist approach. Science **243:** 622–628
3 Mutter M. (1985) The construction of new proteins and enzymes: A prospect for the future? Angew. Chem., Int. Ed. Engl. **24:** 639–653.
4 Hecht M. H., Richardson J. S., Richardson D. C. and Ogden R. C. (1990) De novo design, expression, and characterization of four-helix bundle protein of native-like structure. Science **249:** 884–891
5 Mutter M. and Vuilleumier S. (1989) A chemical approach to protein design: template assembled synthetic proteins (TASP). Angew. Chem. (Engl.) **28:** 535–554
6 Gutte B., Daeumigen M. and Wittschieder E. (1979) Design, synthesis and characterization of a 34-residue polypeptide that interacts with nucleic acids. Nature **281:** 650–655
7 Creighton T. E. (1993) Proteins: Structures and Molecular Properties, 2nd edn., Freeman, New York
8 Jaenicke R. (1987) Folding and association of proteins. Prog. Biophys. Mol. Biol. **49:** 117–237
9 Dill K. A. (1990) Dominant forces in protein folding. Biochem. **29:** 7133–7155
10 Lecomte J. T. J. and Matthews C. R. (1993) Unraveling the mechanism of protein folding: new tricks for an old problem. Protein Eng. **6:** 1–10
11 Mutter M. (1988) Synthetic proteins with a new three-dimensional architecture, in peptides: Chemistry and Biology, pp. 349–354, Marshall G. R. (ed.), ESCOM, Leiden
12 Mutter M. (1988) Nature's rules and chemist's tools: a way of creating novel proteins. Trends Biochem. Sci. **13:** 261–265
13 Mutter M. and Tuchscherer G. (1988) The construction of new proteins. Template assembled synthetic 4-helix bundle protein. Makromol. Chem. Rapid Commun. **9:** 437–443
14 Dumy P., Eggleston I. M., Esposito G., Nicula S. and Mutter M. (1996) Solution structure of regioselective addressable functionalized template (RAFT): an NMR and restrained molecular dynamics investigation. Biopolymers **39:** 297–308
15 Dumy P., Eggleston I. M., Cervigni S., Sila U., Sun X. and Mutter M. (1995) A convenient synthesis of cyclic peptides as regioselectively addressable functionalized templates (RAFT). Tetrahedron Lett. **36:** 1255–1258
16 Thompson M. C. and Bush D. H. J. (1964) Reactions of coordinated ligands. VI. Metal ion central in the synthesis of planar nickel (II) complexes of a-diaceto-bis-mercaptoimines. J. Am. Chem. Soc. **86:** 213–217
17 Cram D. J. and Lam P. Y.-S. (1986) Host guest complexation-37. Synthesis and binding properties of a transacylase partial mimic with imidazole and benzyl alcohol in place. Tetrahedron **42:** 1607–1615
18 Anderson S., Anderson H. L. and Sanders J. K. M. (1993) Expanding roles for template synthesis. Acc. Chem. Res. **26:** 469–475
19 Tuchscherer G., Dörner B., Sila U., Kamber B. and Mutter M. (1993) The TASP concept: mimetics of peptide ligands, protein surfaces and folding units. Tetrahedron **49:** 3559–3575
20 Mutter M., Bianco A., Dörner B., Flögel R., Nyanguile O., Razaname A. et al. (1992) Redesigning peptide structures: side chain assembly on topological templates. In: Peptides, pp. 595–596, Eberle A. and Schneider C. (eds.), ESCOM, Leiden
21 Mutter M., Tuchscherer G., Miller C., Altmann K.-H., Carey R. I., Wyss D. F. et al. (1992) Template-assembled synthetic proteins with four helix bundle topology. Total chemical synthesis and conformational studies. J. Am. Chem. Soc. **114:** 1463–1470
22 Tuchscherer G., Servis C., Rivier J., Blum U., Corradin G. and Mutter M. (1992) Chemical synthesis, characterization and immunological properties of a MHC Class I model using the TASP concept for protein de novo design. Protein Sci. **1:** 1377–1386
23 Hirschmann R., Nicolaou K. C., Pietranico S., Salvino J., Leahy E. M., Sprengeler P. A. et al. (1992) Nonpeptidal peptidomimetics with a β-D-glucose scaffolding. A partial somatostatin agonist bearing a close structural realtionship to a potent, selective substance P antagonist. J. Am. Chem. Soc. **114:** 9217–9218
24 Kazmierski W. M., Yamamura H. I. and Hruby V. J. (1991) Topographic design of neurotransmitters and hormones on stable backbone templates: Relation of conformation and dynamics to bioactivity. J. Am. Chem. Soc. **113:** 2275–2283

25 Flögel R. and Mutter M. (1992) Molecular dynamics conformational search of six cyclic peptides used in the TASP approach for protein de novo design. Biopolymers **32:** 1283–1310

26 Dawson P. E. and Kent S. B. H. (1993) Convenient total synthesis of a 4-helix TASP molecule by chemoselective ligation. J. Am. Chem. Soc. **115:** 7263–7266

27 Kemp D. S. and Carey R. I. (1993) Synthesis of a 39-peptide and a 25-peptide by thiol capture ligations: observation of a 40-fold rate acceleration of the intramolecular O, N-acyl-transfer reaction between peptide fragments bearing only cystein protective groups. J. Org. Chem. **58:** 2216–2222

28 Gaertner H. F., Rose K., Cotton R., Timms D., Camble R. and Offord R. E. (1992) Construction of protein analogues by site-specific condensation of unprotected fragments. Bioconjugate Chem. **3:** 362–368

29 Liu C.-F. and Tam J. P. (1994) Chemical ligation approach to form a peptide bond between unprotected segments. J. Am. Chem. Soc. **116:** 4149–4153

30 Nyanguile O., Mutter M. and Tuchscherer G. (1994) Synthesis of antiparallel 4α-helix bundle TASP by chemoselective ligation. Lett. Peptide Sci. **1:** 9–16

31 Ackerfeldt K. S., Kim R. M., Camac D., Groves J. T., Lear J. D. and DeGrado W. F. (1992) Tetraphilin: a four-helix bundle proton channel built on a tetraphenylporphyrin framework. J. Am. Chem. Soc. **114:** 9656–9657

32 West M. W. and Hecht M. H. (1995) Binary patterning of polar and nonpolar amino acids in the sequences and structures of native proteins. Protein Sci. **4:** 2032–2039

33 Sasaki T. and Kaiser ET. (1990) Synthesis and structural stability of helichrome as an artificial hemeprotein. Biopolymers **29:** 79–88

34 Grove A., Mutter M., Rivier J. E. and Montal M. (1993) Template-assembled synthetic proteins designed to adopt a globular four helix bundle conformation form ionic channels in lipid bilayers. J. Am. Chem. Soc. **115:** 1204–1208

35 Ghadiri M. R. and Choi C. (1992) Secondary structure nucleation in peptides. Transition metal ion stabilized α-helices. J. Am. Chem. Soc. **112:** 1630–1632

36 Gibb B. C., Mezo A. R. and Sherman J. C. (1995) Synthesis and characterization of a de novo four-helix bundle. In: Peptides: Chemistry, Structure and Biology, pp. 585–586, Kaumaya P. T. P. and Hodges R. S. (eds.), Mayflower Scientific Ltd, Kingswinford

37 Stewart J. M. and Young J. D. (1984) Solid Phase Peptide Synthesis, 2nd edn., Pierce Chemical, Rockford, IL

38 Lloyd-Williams P., Albericio F. and Giralt E. (1991) Convergent solid-phase peptide synthesis. Int. J. Pep. Protein Res. **37:** 58–60

39 Rose K. (1994) Facile synthesis of homogeneous artificial proteins. J. Am. Chem. Soc. **116:** 30–33

40 Dawson P. E., Muir T. W., Clark-Lewis I. and Kent S. B. H. (1994) Synthesis of proteins by native chemical ligation. Science **266:** 776–779

41 Wöhr T., Wahl F., Nefzi A., Rohwedder B., Sato T., Sun X. et al. (1996) Pseudo-prolines as a solubilizing, structure-disrupting protection technique in peptide synthesis. J. Am. Chem. Soc. **118:** 9218–9227

42 Dumy P., Keller M., Ryan D. E., Rohwedder B., Wöhr T. and Mutter M. (1997) Pseudo-prolines as a molecular hinge: reversible induction of cis amide bonds into peptide backbones. J. Am. Chem. Soc. **119:** 918–925

43 Pawlak M., Meseth U., Dhanapal B., Mutter M. and Vogel H. (1994) Template-assembled melittin: structural and functional characteristics of a designed, synthetic channel-forming protein. Protein Sci. **3:** 1788–1805

44 Tam J. P. (1988) Synthetic peptide vaccine design: synthesis and properties of a high- density multiple antigenic peptide system. Proc. Natl. Acad. Sci. USA **85:** 5409–5413

45 Haack T. and Mutter M. (1992) Serine derived oxazolidines as secondary structure disrupting, solubilizing building blocks in peptide synthesis. Tetrahedron Lett. **33:** 1589–1592

46 Sheppard R. C. (1994) A cycle of solid phase synthesis, in Peptides 1994, pp. 3–15, Maia H. L. S. (ed.), ESCOM, Leiden

47 Nefzi A., Schenk K. and Mutter M. (1994) Pseudo-prolines: temporary insertion of a cis- amide bond into a peptide backbone. Protein and Peptide Lett. **1:** 66–69

48 Satow Y., Cohen G. H., Padlan E. A. and Davies D. R. (1986). Phosphocholine binding immunoglobulin Fab McPC 603. An X-ray diffraction study at 2.7 Å. J. Mol. Biol. **593:** 190–196

49 Nefzi A., Sun X. and Mutter M. (1994) Chemoselective ligation of multifunctional peptides to topological templates by thioether formation according to the TASP concept. Tetrahedron Lett. **36:** 229–230

50 Mutter M., Nefzi A., Sato T., Sun X., Wahl F. and Wöhr T. (1995) ΨProlines (ΨPro) for accessing "inaccessible" peptides. Pep. Res. **8:** 145–153

51 Mutter M., Dumy P., Garrouste P., Lehmann C., Mathieu M., Peggion C. et al. (1996) Template assembled synthetic peptides (TASP) as functional mimetics of proteins. Angew. Chem. (Engl) **35:** 1482–1485

52 Tuchscherer G. (1993) Template assembled synthetic proteins: condensation of a multifunctional peptide to a topological template via chemoselective ligation. Tetrahedron Lett. **34:** 8419–8422

53 Bloomberg G. B., Askin D., Gargaro A. R. and Tanner M. J. A. (1993) Synthesis of a branched cyclic peptide using a strategy employing Fmoc chemistry and two additional orthogonal protecting groups. Tetrahedron Lett. **34:** 4709–4713

54 Manning M. C. and Woody R. W. (1987), Theoretical determination of the CD of proteins containing closely packed antiparallel β-sheets. Biopolymers **26:** 1731–1752

55 Mathieu M., Lehmnann C., Razaname A. and Tuchscherer G. (1997) Engineering of zinc finger and MHC motifs to locked-in tertiary folds. Lett. Peptide Sci. (LIPS) **4:** 95–100

56 Elrod-Erickson M., Rould M. A., Nekludova L. and Pabo C. O. (1996) Zif 268 protein- DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. Structure **4:** 1171–1180

57 Klug A. and Schwabe J. W. (1995) Protein motifs 5. Zinc fingers. FASEB J. **9:** 597–604

58 Rebar E. J., Greisman H. A. and Pabo C. O. (1996) Phage display methods for selecting zinc finger proteins with novel DNA-binding specificites. Methods Enzmol. **267:** 129– 149

59 Desjarlais J. R. and Berg J. M. (1993) Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding properties. Proc. Natl. Acad. Sci. USA **90:** 2256–2260

60 Choo Y., Sanchez-Garcia I. and Klug A. (1994) In vivo repression by a site-specific DNA binding protein designed against an oncogenic sequence. Nature **15:** 642–645

61 Gerber P. R. (1992) Peptide mechanics: a force field for peptide and proteins working with entire residues as smallest units. Biopolymers **32:** 1003–1017

62 Cheng R. P. and Imperiali B. (1996) Design of a monomeric 23-residue polypeptide with defined tertiary structure. Science **271:** 342–345