CHIMIA 50 (1996) Nr. 12 (Dezember)

Chimia 50 (1996) 644–648

ISSN 0009–4293

Protein *de novo* Design

Gabriele Tuchscherer, Pascal Dumy, and **Manfred Mutter***

Abstract. The ultimate goal in protein *de novo* design is the construction of artificial proteins exhibiting tailor-made structural and functional properties. To create native-like macromolecules in copying nature's way has proven to be difficult because the mechanism of folding in its complexity has yet to be unraveled. In order to bypass the well-known folding problem, we have developed the concept of template assembled synthetic proteins (TASP); this meanwhile widely accepted strategy uses topological templates as 'built-in' devices for the induction of well-defined folding topologies. Progress in synthetic strategies, *e.g.*, chemoselective ligation methods and orthogonal protection techniques open the way for the design of more complex TASP molecules featuring functional properties such as membrane channels, vaccines, catalysts, receptors, or ligands.

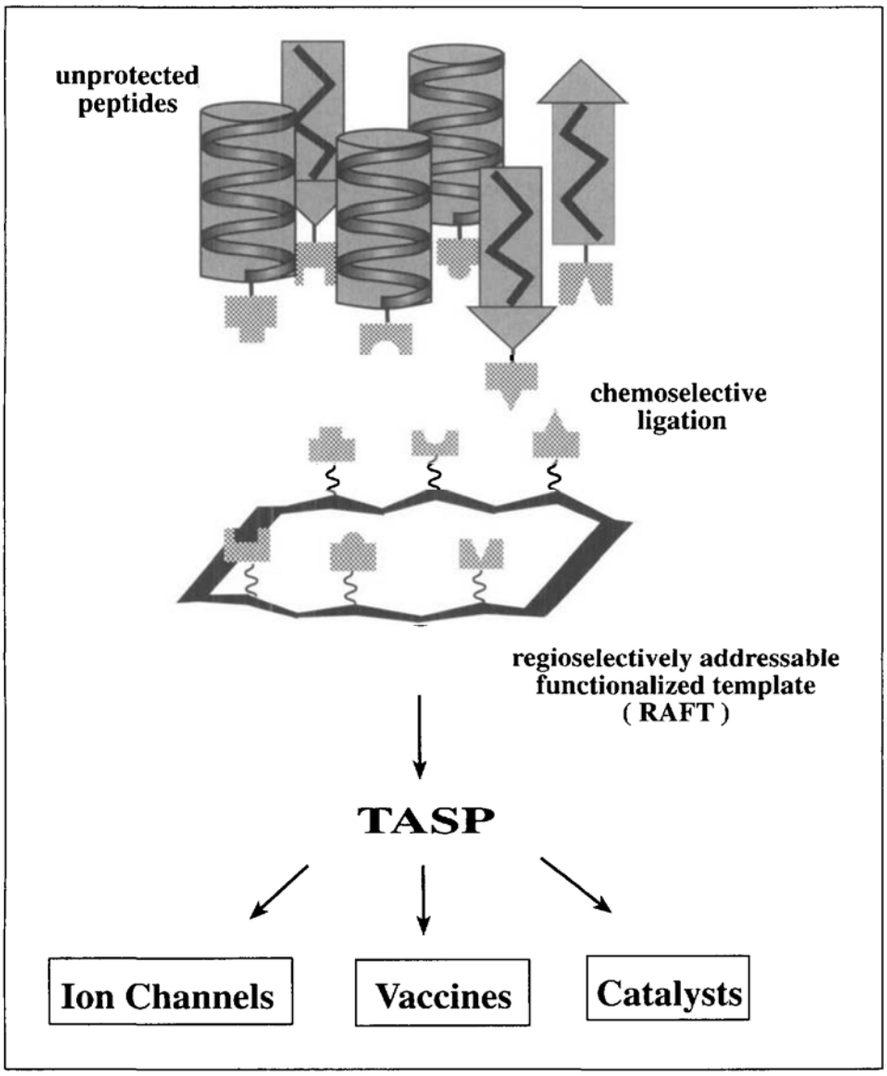


Fig. 1. *Topological templates as folding devices have been used for the induction and stabilization of protein folds such as α-helical or β-sheet TASP molecules for mimicking some properties of natural proteins*

Introduction

Constructing novel proteins with functional properties similar to natural proteins is the ultimate and one of the most challenging goals in biomimetic chemistry [1–3].

In the absence of a detailed knowledge of the folding mechanism of natural proteins, a general design strategy of polypeptide sequences with a high propensity to fold in a predetermined three-dimensional structure appears to be still out of reach. The major obstacle in the construction of artificial proteins rests in the complexity of the folding pathway of a linear polypeptide chain to its unique 3D structure. Among the surprisingly small number of recurring structural motifs [4][5] the formation of a specific globular fold is not confined to one specific amino-acid sequence. This degeneracy of the folding code ('designability') is thought to represent a fundamental selection principle in the evolution of native proteins [6].

The present hypotheses on the mechanism of protein folding [7] do not allow to derive generally valid principles for protein *de novo* design; however, the synthetic chemist is not constrained to nature's machinery for polypeptide synthesis, *i.e.*, he may create nonnatural chain topologies with an increased potential for intramolecular folding.

The Template Concept

Since the complex folding mechanism has yet to be unraveled, we proposed some years ago a conceptually different approach in protein *de novo* design to bypass the folding problem: the template assembled synthetic proteins (TASP) concept [8–14] (*Fig. 1*). 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.

Topological templates may be generally characterized as synthetic devices that orient functional groups or structural units in well-defined spatial arrangements [15]. Typically, template molecules represent

**Correspondence*: Prof. Dr. M. Mutter
Institute of Organic Chemistry
University of Lausanne
BCH-Dorigny
CH–1015 Lausanne

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 [16–19]. As a second generation of this type of templates, RAFT (regioselectively addressable functionalized templates) molecules exhibit selectively addressable sites due to orthogonal protection techniques or unique chemical reactivity [20–22].

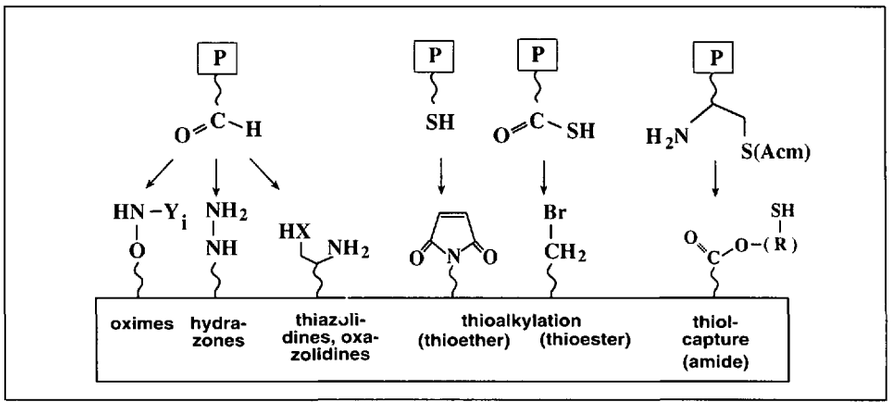


Fig. 2. *Chemoselective ligation methods for protein design*

Synthetic Aspects in Protein Design

Molecular biology has revolutionized the study of protein structure and function. Not only has the microbial production of enzymes and other proteins in useful amounts become routine in recent years but with systematic alteration of protein sequence by site-specific mutagenesis virtually any protein sequence is accessible. Recent developments for ligating peptides, however, promises to make chemical synthesis of large proteins an attractive alternative to biosynthesis, particularly for the construction of novel molecules containing nonnatural amino acids or other structural modifications.

The well established difficulties of solid-phase peptide synthesis [23] and fragment condensation [24] can be circumvented by using the recently introduced chemoselective ligation methods [25–28] which allow for the condensation of completely unprotected peptide fragments in aqueous medium (*Fig. 2*).

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 regioselectively addressable functionalized template [21] molecules as key compounds for the construction of TASP molecules of higher complexity. As shown in *Fig. 3*, a prototype TASP with up to four different helices using different chemoselective ligation procedures can be synthesized.

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 unfavorable in the

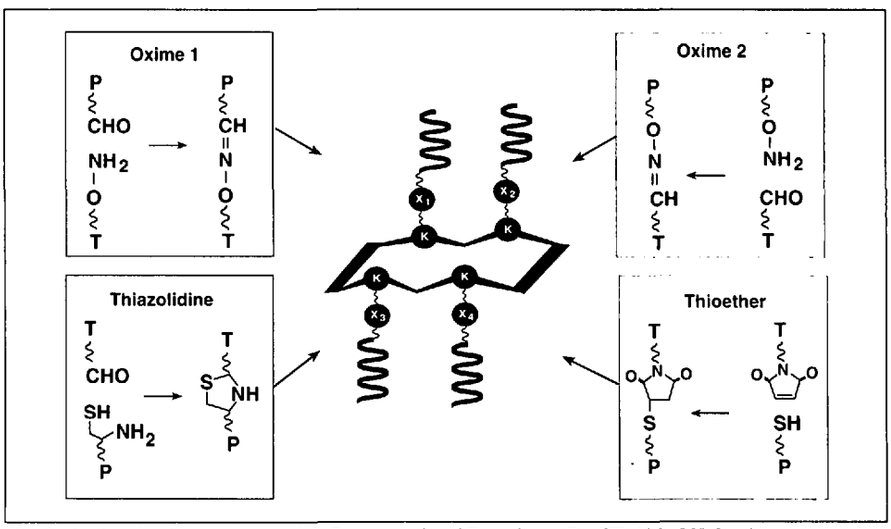


Fig. 3. *Regioselectively Addressable Functionalized Templates* (RAFT) [20–22] *for the construction of TASP molecules by chemoselective ligation procedures* (see text)

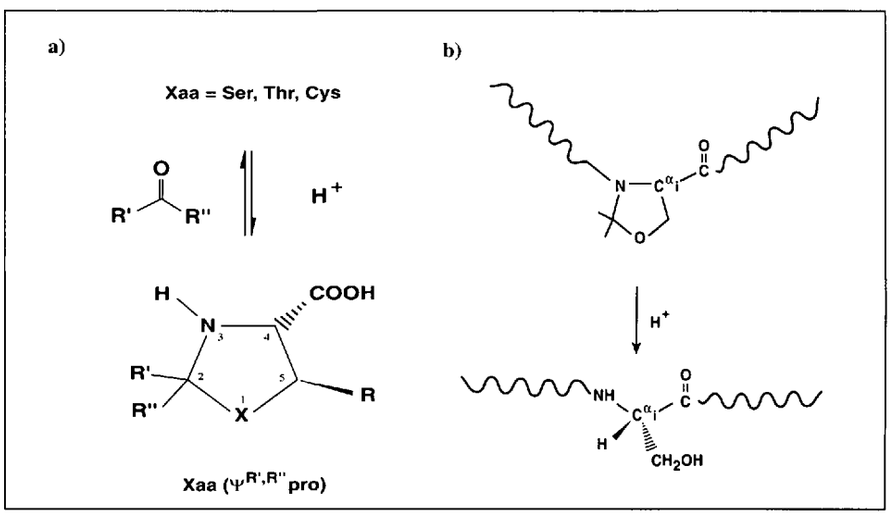


Fig. 4. a) *Ser-, Thr-, Cys-derived* pseudo-*prolines* (ψPro); b) *effect of ψPro upon peptide backbone by induction of a* cis-*amide bond* [29–34]

ligation process which results, *e.g.*, in extended reaction times or incomplete reactions. To prevent self-association of unprotected, secondary-structure forming peptide sequences during chain assembly, *pseudo*-prolines (ψPro) have been introduced [29–34] as a powerful tool to modify temporarily the intrinsic properties of peptides that are responsible for aggregation and secondary-structure formation (*Fig. 4*). *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 (*Fig. 4, a*).

Due to the presence of a cyclic system (fixed ϕ-angle) in addition to the preference for a *cis*-amide bond [30][33] with the preceding residue, the incorporation of a ψPro moiety results in a kink conformation (*Fig. 4, b*), 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.

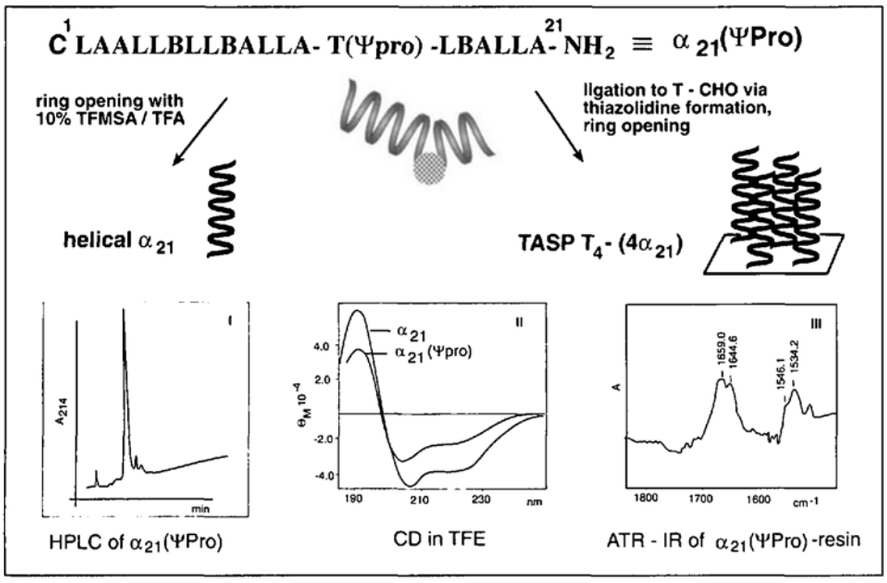


Fig. 5. *Synthesis scheme for a membrane spanning hydrophobic peptide using Thr(ψPro); I: HPLC of the crude, ψPro containing peptide α_{21}; II: CD of α_{21} and α_{21}(ψPro) in TFE; III: ATR-IR of the resin-bound α_{21}(ψPro)*

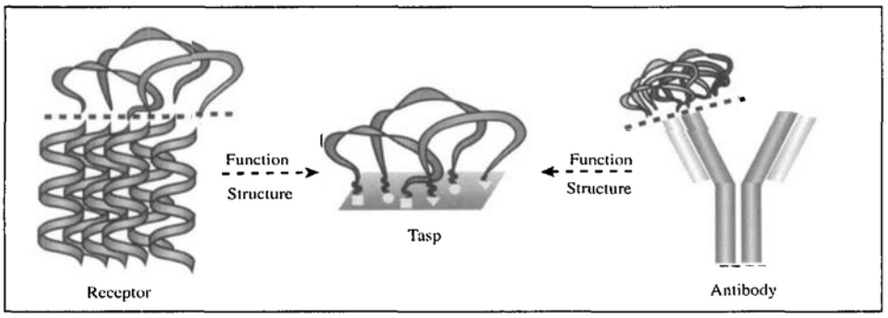
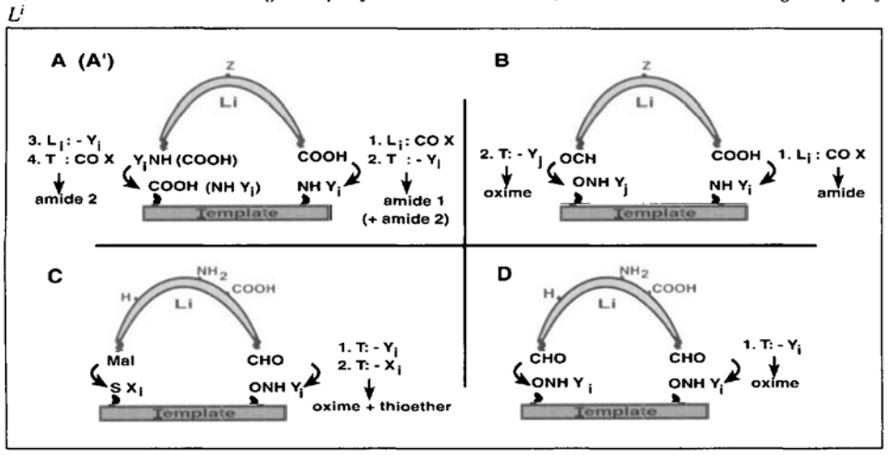


Fig. 6. *Concept of template assembled synthetic proteins as mimetics of the binding region of receptors and antibodies*

Scheme. *Strategies A–D for the Attachment of Peptide Loop Sequences* (L^i) *to Topological Templates* T^j; Y^i *and* Y^j *Amino-Protecting Groups of the Attachment Sites; Z: Side-Chain-Protecting Groups of* L^i



Template Assembled Synthetic Proteins (TASP)

Several authors have reported on selective membrane channel forming TASP molecules using topological templates to define and orient membrane spanning helical segments [35][36]. For example, *Pawlak et al.* designed and synthesized a template assembled channel-forming protein derived from the bee venom melittin [37].

As a striking common feature, the membrane channel forming TASP molecules exhibit single channel conductance, ion selectivity and high thermodynamic stability.

As a new design element, a ψPro unit was inserted to induce a reversible kink in the helical transmembrane peptide [31][32] [34]. 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 mild acidic conditions, thus preserving the oxazolidine ring structure of the Thr($\psi^{H,H}$Pro) residue. As indicated by CD and ATR-IR studies [38], the ψPro indeed distorts the helix to some extent (*Fig. 5*). These helical transmembrane peptides are presently subject to chemoselective ligation to topological templates to access membrane active TASP molecules with well-defined three-dimensional structures.

Amphiphilic β-sheet forming peptides and their assembly to a topological template represent an even more challenging example for the solubilizing effect of *pseudo*-prolines and the versatility of chemoselective ligations [31][38].

Applying similar synthetic strategies, template assembled bioactive peptides derived from angiotensin II and neuropeptide Y resulted in interesting modifications of their physicochemical and pharmacokinetic properties [39][40]. Moreover, TASP libraries mimicking protein surfaces represent a first step towards functional mimetics of proteins according to the template approach [16][41].

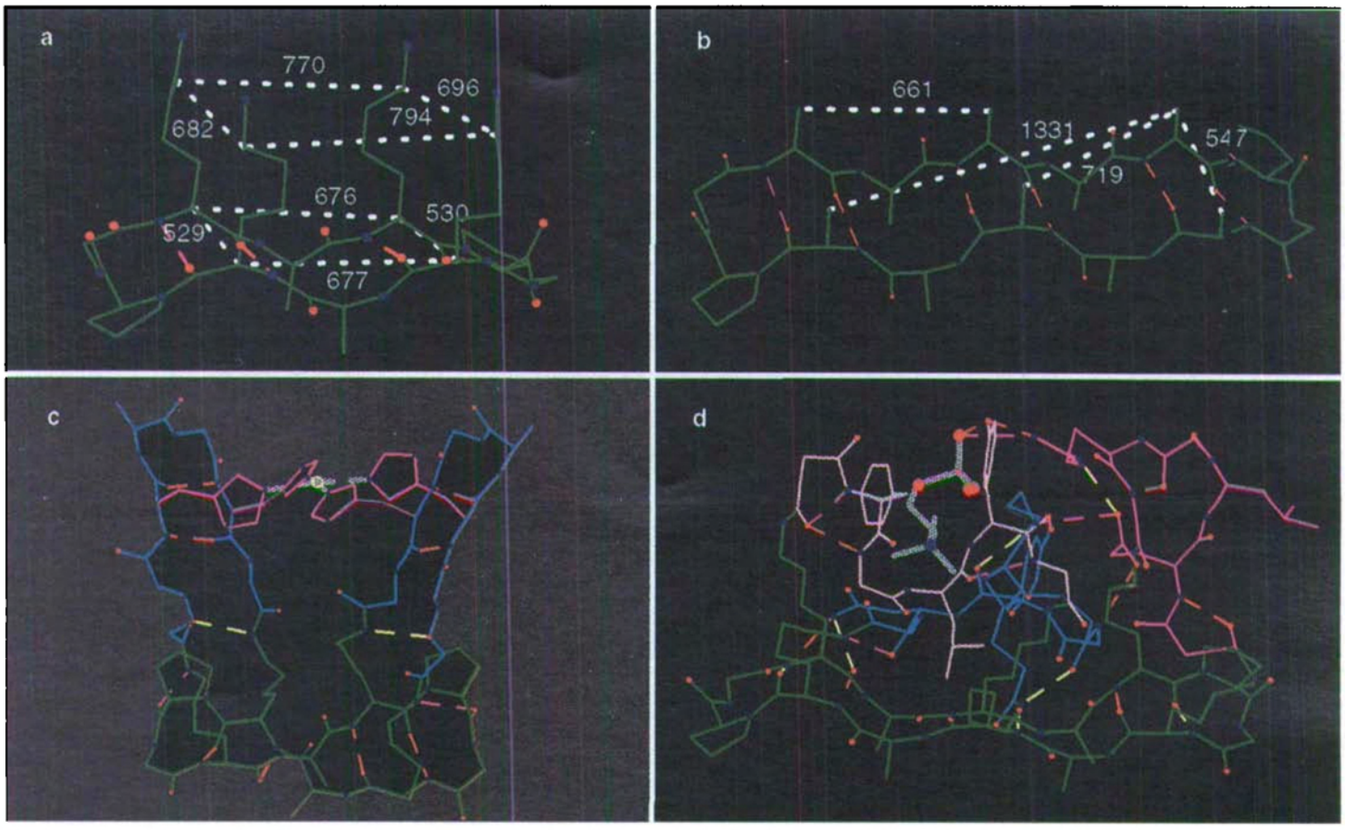


Fig. 7. *Design of TASP molecules as functional protein mimetics* [42]. *a*) A cyclic (ProGlyLysAlaLys)$_2$ β-sheet mimetic obtained as a full structure from a regularized protein Ca-trace has been built by means of the program MOLOC [43]. *b*) Extending the chain length of the cyclopeptide to the 14-mer cyclic peptide (ProGlyLysAlaLysAlaLys)$_2$ yields an increasing number of attachment sites and range of accessible distances. Thus larger distances become accessible to span functional loops which even may mimic an enzymatic binding cavity. *c*) Two-loop TASP featuring a square planar metal-binding site formed by two identical HisAlaGlyHisGly sequences assembled on the 10-membered cyclic β-sheet template. *d*) One out of 8 possible directional isomers of a 3-loop TASP approaching the antigen-binding site of the phosphorylcholine-specific antibody McPC603.

TASP as Mimetics of Proteins and Receptors

In separating structure and function of a protein in its constituent elements we have recently proposed the use of topological templates as mimetic of the structural part [42].

The functional part (*e.g.* the binding loops of a receptor or antibody) is detached from its supporting structural framework and assembled on a topological template by selective chemical ligation reactions (*Fig. 6*). So far, limitations in the synthetic strategies prevented this novel generation of functional TASP molecules from being realized.

Antibody Mimetics: Attaching Binding Loops to Templates

The binding of an antigen by the hypervariable loops of its corresponding antibody molecule represents a general principle for molecular recognition in biological systems. Because of the distinct separation into a 'functional' (complementarity determining region, CDR) and 'structural' (immunoglobuline fold, Ig) domain, this class of proteins seems to be an excellent candidate for probing the TASP concept (*Fig. 6*).

For realizing the synthetic key steps in this approach, *i.e.*, the sequential condensation of peptide loops of different chain length and sequence to a topological template, we have elaborated several strategies in combining orthogonal protection techniques and chemoselective ligation methods (*Scheme*). Starting from a pool of *n* orthogonal amino- or carboxylic protection groups, up to *n* different loops can be selectively fixed by the individual strategies depicted in the *Scheme* [42].

Metal Binding TASP Molecules

The design and synthesis of a potentially metal-binding TASP molecule was achieved by connecting two peptides (containing two histidine residues each acting as metal-complexing sites) *via* linker groups onto a cyclic template. The binding site is formed by two identical HPGHGG sequences with model conformations determined by the same folding principle as the one operating in the template. After conformational relaxation, the TASP molecule adopts a favorable spatial arrangement of the ligands for accommodating a square planar metal-binding site. Interestingly, intramolecular H-bonds and the adoption of a β-turn type II conformation of the loop sequences point to an additional stabilization of the metal complexing structure (*Fig. 7*).

The chemical synthesis of this two loop TASP molecule was achieved according to strategy B (*Scheme*). To this end, the cyclic template containing four orthogonal amino-protecting groups (Aloc, Dde, Boc, Fmoc) was transformed into a chemoselectively addressable template, allowing for the subsequent regioselective condensation of the loops *via* amide and oxime bond formation. A more efficient solid-phase strategy for this type of TASP molecules is presently elaborated in our laboratory.

Conclusions

In summary, the presented strategies provide a synthetic entry to protein mim-

icry and artificial proteins. 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 molecules including TASP libraries for functional screening. However, the exploitation of the immense repertoire of synthetic organic chemistry for the incorporation of functional groups ('sticky ends') into peptide chains will rapidly expand the scope of the present approach. For example, ligand-directed assembly of helices, β-sheets, and loops onto tailor-made templates represents a powerful tool for studying supramolecular assembly and molecular recognition processes. Furthermore, the concepts described above open the way for a new generation of functional protein mimetics with a pivotal role in drug design.

This work was supported by the *Swiss National Science Foundation.*

Received: September 30, 1996

[1] M. Mutter, *Angew. Chem. Int. Ed.* **1985**, *24*, 639.
[2] J.S. Richardson, D.C. Richardson, *Trends Biochem. Sci.* **1989**, *14*, 304.
[3] J.W. Bryson, S.F. Betz, H.S. Lu, D.J. Suich, H.X. Zhou, K.T. O'Neil, W.F. DeGrado, *Science* **1995**, *270*, 935.
[4] J.S. Richardson, *Adv. Prot. Chem.* **1981**, *34*, 167.
[5] S.R. Presnell, F.E. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 6592.
[6] H. Li, R. Helling, C. Tang, N. Wingreen, *Science* **1966**, *273*, 666.
[7] S. Borman, *Chem. Eng. News*, May **1996**, *27*, 29.
[8] M. Mutter, G. Tuchscherer, *Makromol. Chem. Rapid Commun.* **1988**, *9 (6)*, 437.
[9] M. Mutter, R. Hersperger, K. Gubernator, K. Mueller, *Proteins* **1989**, *5*, 13.
[10] M. Mutter, E. Altmann, K.-H. Altmann, R. Hersperger, P. Koziej, K. Nebel, G. Tuchscherer, *Helv. Chim. Acta* **1989**, *71*, 835.
[11] M. Mutter, S. Vuilleumier, *Angew. Chem. Int. Ed.* **1989**, *28*, 535.
[12] G. Tuchscherer, M. Mutter, *J. Peptide Sci.* **1995**, *1*, 3.
[13] G. Tuchscherer, M. Mutter, *J. Biotechnol.* **1995**, *41*, 197.
[14] M. Mutter, G. Tuchscherer, C. Miller, K.-H. Altmann, R.I. Carey, D. Wyss, A.M. Labhardt, J. Rivier, *J. Am Chem. Soc.* **1992**, *114*, 1463.
[15] S. Anderson, H.L. Anderson, J.K.M. Sanders, *Acc. Chem. Res.* **1993**, *26*, 469.
[16] U. Sila, M. Mutter, *J. Mol. Recognition* **1995**, *8*, 29.
[17] M. Mutter, *Trends Biochem. Sci.* **1989**, *13*, 260.
[18] R. Flögel, M. Mutter, *Biopolymers* **1992**, *32*, 1283.
[19] I. Ernest, J. Kalvoda, C. Sigel, G. Rihs, H. Fritz, M.J.J. Blommers, F. Raschdorf, E. Francotte, M. Mutter, *Helv. Chim. Acta* **1993**, *76*, 1539.
[20] P. Dumy, I.M. Eggleston, S.E. Cervigni, U. Sila, X. Sun, M. Mutter, *Tetrahedron Lett.* **1995**, *36*, 1255.
[21] P. Dumy, I.M. Eggleston, G. Esposito, S. Nicula, M. Mutter, *Biopolymers*, in press.
[22] I.M. Eggleston, M. Mutter, *Macromol. Symp.* **1996**, *101*, 397.
[23] J.M. Stewart, J.D. Young, 'Solid Phase Peptide Synthesis', Pierce Chemical, 1984.
[24] P. Lloyd-Williams, F. Albericio, F. Giralt, *Int. J. Pept. Protein Res.* **1991**, *37*, 58.
[25] L.A. Vilaseca, K. Rose, R. Werlen, A. Meunir, R.E. Offord, C.L. Nichols, W.L. Scott, *Bioconj. Chem.* **1993**, *4*, 515.
[26] C.-F. Liu, J.P. Tam, *J. Am. Chem. Soc.* **1994**, *116*, 4149.
[27] P.E. Dawson, S.B.H. Kent, *J. Am. Chem. Soc.* **1993**, *115*, 7263.
[28] G. Tuchscherer, *Tetrahedron Lett.* **1993**, *34*, 8419.
[29] T. Haack, M. Mutter, *Tetrahedron Lett.* **1992**, *33*, 1589.
[30] A. Nefzi, K. Schenk, M. Mutter, *Protein Peptide Lett.* **1994**, *1*, 66.
[31] M. Mutter, A. Nefzi, T. Sato, X. Sun, F. Wahl, T. Wöhr, *Pept. Res.* **1995**, *8*, 145.
[32] T. Wöhr, A. Nefzi, B. Rohwedder, T. Sato, X. Sun, F. Wahl, M. Mutter, *J. Am. Chem. Soc.*, in press.
[33] P. Dumy, M. Keller, D.E. Ryan, B. Rohwedder, T. Wöhr, M. Mutter, *J. Am. Chem. Soc.*, in press.
[34] A. Nefzi, X. Sun, M. Mutter, *Tetrahedron Lett.* **1995**, *36*, 229.
[35] K.S. Ackerfeldt, R.M. Kim, D. Camac, J.D. Groves, J.D. Lear, W.F. DeGrado, *J. Am. Chem. Soc.* **1992**, *114*, 9656.
[36] A. Grove, M. Mutter, J.E. Rivier, M. Montal, *J. Am. Chem. Soc.* **1993**, *115*, 1872.
[37] M. Pawlak, U. Meseth, B. Dhanapal, M. Mutter, H. Vogel, *Protein Sci.* **1994**, *3*, 1788.
[38] H. Vogel, M. Mutter *et al.*, in preparation.
[39] E. Grouzmann, D. Felix, H. Imboden, A. Razaname, M. Mutter, *Eur. J. Biochem.* **1995**, *234*, 44.
[40] E. Grouzmann, T. Buclin, M. Martire, B. Dörner, A. Razaname, M. Mutter, *Eur. J. Biochem.*, in press.
[41] G. Tuchscherer, B. Dörner, U. Sila, B. Kamber, M. Mutter, *Tetrahedron* **1993**, *49*, 3559.
[42] M. Mutter, P. Dumy, P. Garrouste, C. Lehmann, M. Mathieu, C. Peggion, S. Peluso, A. Razaname, G. Tuchscherer, *Angew. Chem. Int. Ed. E* **1996**, *35*, 1482.
[43] P.G. Gerber, K. Müller, *J. Comput. Aided Mol. Des.* **1995**, *9*, 251.

Chimia 50 (1996) 648–649

ISSN 0009–4293

Azasugar Glycosidase Inhibitors: Useful Tools for Glycobiologists

Sylviane Picasso*

Abstract. Glycosidase inhibitors are moving increasingly as potential agents for the treatment of diseases by altering glycosylation or catabolism of glycoconjugates. The azasugars, a potent class of inhibitors have received much attention as tools for therapeutic investigation and in understanding of certain biological processes such as glycoprotein processing and enzymatic mechanisms. In this paper, we present some new synthetic azasugar compounds which show interesting activity as glycosidase inhibitors.

Introduction

Glycoproteins have many important biological functions: receptors, carriers, structural proteins, enzymes, hormones, lectins, and antibodies. Inhibitors of glycosidases, key enzymes in glycoprotein processing, are thus important as potential antiviral, antiadhesive, antitumoral, antibacterial, antihyperglycemic, or immunostimulatory agents [1]. Thus, an intense interest in the chemistry, biochemistry and pharmacology of glycosidase inhibitors has grown during the last decade and has

*Correspondence: Dr. S. Picasso
Institut de chimie organique
Université de Lausanne
CH–1015 Lausanne