Journal of Computer-Aided Molecular Design, **12:** 1–6, 1998.
KLUWER/ESCOM

Perspective

Protein secondary structure templates derived from bioactive natural products

Combinatorial chemistry meets structure-based design

Gerhard Müller[a] and Henry Giera[b]
[a]*IM-FA, Computational Chemistry,* [b]*Central Research ZF-SF O2, Synthesis Research, Combinatorial Chemistry, Bayer AG, Building Q18, D-51368 Leverkusen, Germany*

Received 21 August 1997; Accepted 10 November 1997

Key words: combinatorial chemistry, cyclo-isodityrosine, lead finding, natural products, peptidomimetics, secondary structure-templated libraries, spatial screening

Summary

Lead finding strategies in pharmaceutical research comprise structure-based drug design as well as screening efforts of natural product pools or large chemical libraries. In this context we propose a combined approach by utilizing natural product-derived structure information on receptor- or enzyme-complementarity for designing unique core structures that can be employed as privileged template molecules underlying combinatorial libraries. A set of rules for the transformation of molecular frameworks from natural products to structurally defined peptidomimetics is introduced. Special emphasis is laid on the correspondence in the orientational properties and functionalization patterns between natural products and regular protein secondary structures.

Molecular recognition is one of the key principles underlying all biochemical processes. Thus, the structural investigation of such phenomena by biophysical and computational means, especially those which include the interaction of low-molecular weight compounds with highly complex biopolymers, has evolved as a central topic in pharmaceutical research [1, 2].

Thorough structural analyses of high-resolution structures of biomedically interesting target-effector complexes comprising peptide and protein ligands revealed a predominant role of regular secondary structure elements, such as α helices, β strands and reverse turns, as structure determinants within numerous enzyme-substrate/inhibitor and receptor-agonist/antagonist complexes. The α helix is a prominent structural motif in, e.g., protein-DNA interactions, various protease inhibitors fold into β strand conformations, while peptide ligands targeted to, e.g., peptide hormone receptors frequently utilize reverse turn structures for molecular recognition [3]. Consequently, approaches to a rational design of protein secondary structure mimetics have attracted the attention of medicinal chemists over the last decade [4, 5].

On the other hand, several non-peptidic natural products of non-mammalian origin with promising pharmacological profiles are described as potent interaction partners for many peptide- or protein-binding macromolecular targets of biomedical interest [6].

In addition, the drug discovery process is currently being revolutionized by a set of high-throughput chemistry methods covered by the term combinatorial chemistry, which allow the identification and optimization of new drug leads in a resource- and time-efficient manner [7, 8]. The potential of high-speed chemical synthesis based on parallel and automated solution phase, solid phase and hybrid techniques is exploited for the production of large ensembles of molecules, commonly termed libraries, composed of single compounds or even mixtures of compounds.

In this context, we comment on a new synergistic approach combining structure-based peptidomimetic design based on molecular frameworks of natural products with the concept of combinatorial chemistry,

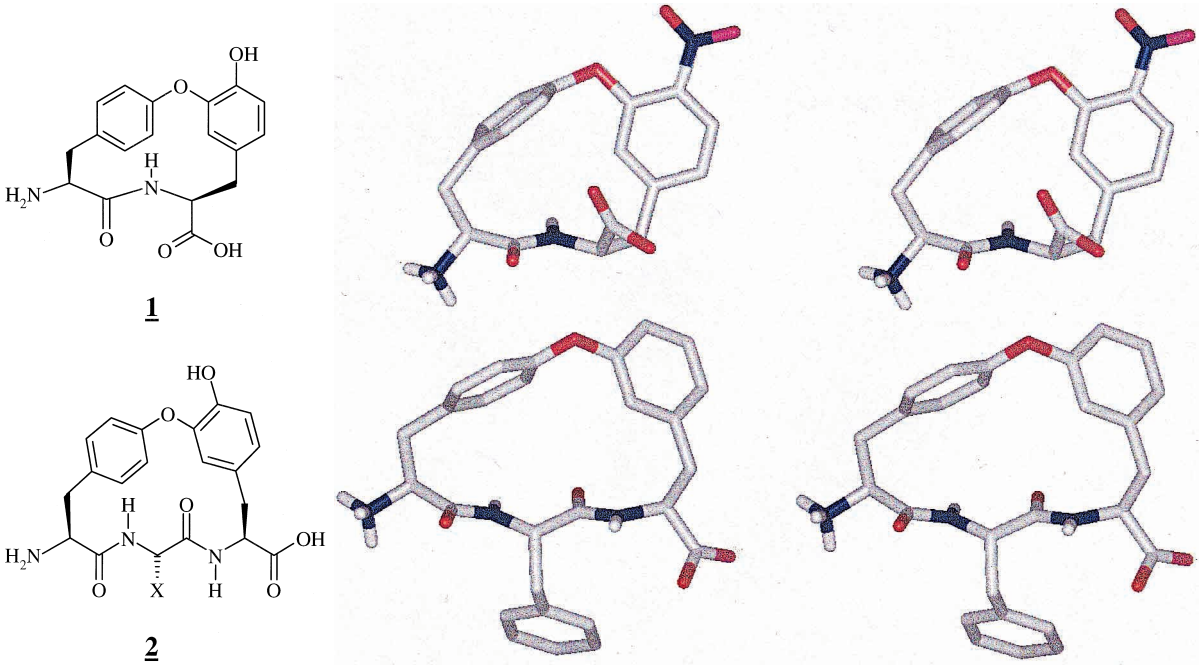


Figure 1. Left: Chemical formulas of the cyclic isodityrosine dipeptide (**1**) and tripeptide (**2**, X denotes the side-chain of any amino acid residue). Right: Side-by-side stereo presentations of the crystallographically determined structures of the cyclic compounds. For reasons of clarity only heavy atoms and polar hydrogen atoms are depicted. The coordinates of the derivative of **1** (top) were extracted from the Cambridge Structural Database (CSD entry code: TECXUU), while the structural data of the derivative of **2** (bottom) were kindly provided by Dan Rich.

exemplified with side-chain-bridged cycloisodityrosine di- [9–11] and tripeptides [12–14] (**1**,**2**), independently described in the literature only recently (Figure 1).

This pair of biphenylether-bridged macrocycles deserves our special attention, since we envision these molecules as structural switches between two prominent regular secondary structure mimics, namely the β turn (**1**) and the β strand (**2**) conformation, by utilizing the identical endocyclic biaryl ether linkage tethering either a dipeptide (**1**) or a tripeptide (**2**), respectively. In the following we want to elaborate the potential of both conformationally restricted structural motifs as core structures for secondary structure-templated libraries, with special emphasis laid on the fact that these molecules emerged from the natural product pool.

The 14-membered biaryl ether core of **1** is the characteristic structural motif of a large class of naturally occurring bicyclic hexapeptides isolated from *Rubia akane* [15]. These so-called RA peptides possess potent antitumor activity by inhibiting protein synthesis through eukaryotic 80S ribosomal binding. Within all experimentally determined structures of RA hexapeptides, the side-chain-bridged Tyr-Tyr dipeptide occupies the $i + 1$-$i + 2$ position of regular and distorted β turns, thus being the predominant structure-inducing element of the overall peptide backbone conformation. Depending on the Tyr-Tyr peptide bond configuration, the adopted turns are of type II (*trans*), or type VI (*cis*), respectively. Boger's group at the Scripps Research Institute published the synthesis and crystal structure of a cyclic Tyr-Tyr dipeptide derivative of **1** [9].

We stress the conformational compatibility of the restricted dipeptide mimic **1** with native β turn conformations by reversing the structural function of the dipeptide moiety within the RA hexapeptides. Instead of representing the $i + 1$-$i + 2$ side-chain positions, we assign a different structural correspondence to the biphenyl ether moiety with the β turn-stabilizing $NH^{i+3} \rightarrow O{=}C^{i}$ hydrogen bond (Figure 2). This particular structural reversal of the dipeptide yields a hydrogen bond-fixed β turn mimetic with potential attachment sites that perfectly correspond to the $i + 1$, $i + 2$, and $i + 3$ side-chain-mounting positions within regular β turns, respectively [16].

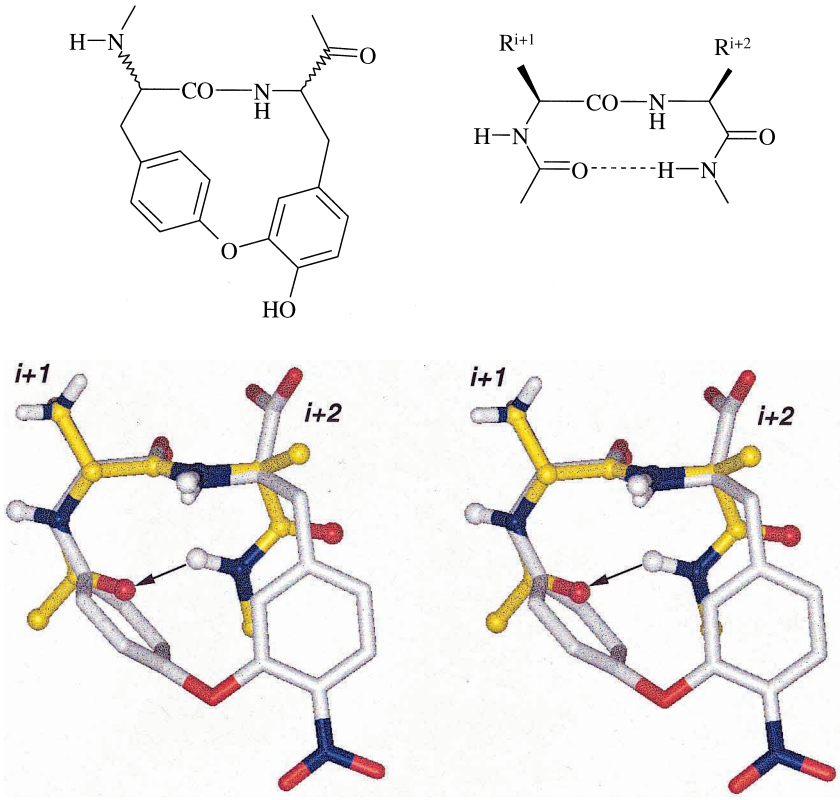


Figure 2. Top: The structural correspondence of the cycloisodityrosine dipeptide (left) with a native β turn topology is based on the assumption that the biaryl ether linkage mimics the turn-stabilizing hydrogen bond. Bottom: Side-by-side stereoview of the structural superposition of the crystallographically determined structure of a derivative of **1** with a β II′ turn structure with residue $i + 1$ in D configuration. The model turn is shown in a ball-and-stick mode, the C atoms colored in yellow. With respect to Figure 1 the structure of compound **1** is reoriented by 180 ° in order to mimic the β turn structure.

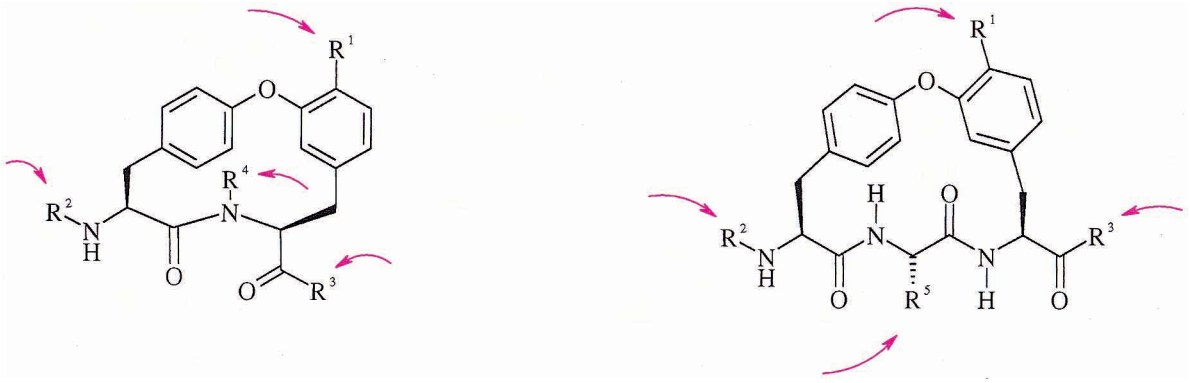


Figure 3. Schematic presentation of the decoration pattern displayed by the natural product-derived template molecules **1** and **2** for combinatorial chemistry approaches. Arrows indicate potential functional groups that allow to increase the dimensionality of the library architecture, thus increasing the diversity of the resulting compound set.

Only recently, Janetka and Rich reported on the structure-based design of a β strand mimic targeted against the HIV 1-protease [14], based on the finding that isodityrosine-derived cyclic tripeptides are naturally occurring inhibitors of metalloproteases, the most prominent representatives from the natural product pool being OF4949-III and K-13 [17, 18].

As shown by crystallography, an analogue of the 17-membered macrocycle **2** exhibits an extended β strand conformation in the peptide backbone portion, thus meeting exactly the complementarity requirements for efficient binding to various types of proteases [13]. By attachment of a side-chain-bearing aspartyl protease-specific hydroxyethylamine dipeptide mimic onto **2** (X = Ile), the authors succeeded in deriving a potent HIV 1-protease inhibitor with nanomolar activity [14].

From a structural comparison of compounds **1** and **2**, an *a priori* unexpected but valuable option arises for the structure-based design of natural product-derived core structures for combinatorial libraries, namely the secondary structure-switch between a β turn mimic (**1**) and a β sheet mimic (**2**) by utilizing the identical isodityrosine moiety as structural clamp.

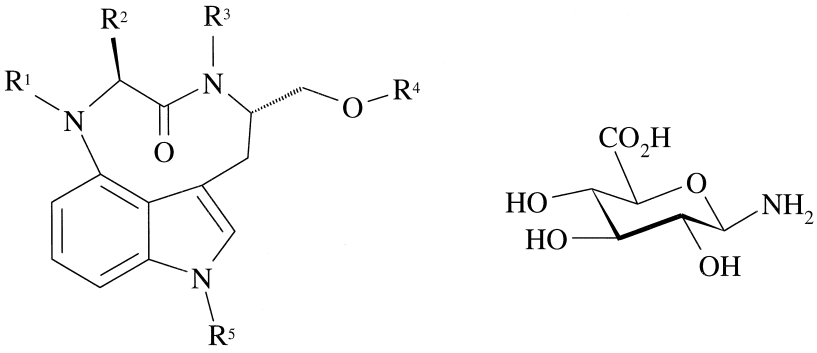
From the viewpoint of chemical feasibility, several groups have established efficient synthetic schemes for generating the heterodetic cyclic peptides with endocyclic biaryl ether linkages. Thus, the studies of, e.g., Boger et al. [9–11, 19–22], Janetka et al. [12–14], Zhu et al. [23, 24], and Burgess et al. [25] opened the route towards a combinatorial exploitation of the secondary structure mimics described in this perspective. The synthetic problems related to the biaryl ether macrocycles were overcome by a synthesis strategy based on intramolecular aromatic nucleophilic (SN_{Ar}) reactions for macrocylization which is superior to numerous alternative attempts including macrolactamization, intramolecular Ullmann reactions or oxidative phenol couplings. The crucial step of mild cyclization was achieved by Rich et al. [12] using $RuCp^{+}\mu$ complexes of tripeptides as electrophilic component, whereas Boger et al. [9, 21, 22] utilize the nucleophilic substitution of an activated fluorine functionality in a 3-nitro-4-fluoro-phenylalanine building block.

In the context of privileged template structures, the 14- and 17-membered macrocycles display a variety of functionalities at multiple sites in the molecule which can be elaborated by established combinatorial techniques on the solid phase (Figure 3).

Indeed, Burgess et al. [25] demonstrated the facile construction of the core structure on a solid support, which can be employed in a combinatorial way by combination of a diverse set of readily available building blocks, i.e., different amino acid residues for the R^5 position and capping reagents for the amino group (R^2 position) as shown in Figure 3. Furthermore, the R^1 substituent in the biphenyl ether bridge can be addressed either as amino or hydroxyl functionality, easily derived from the nitro group following the nitro-fluoro-phenylalanine strategy. Both can be derivatized following alkylation or acylation synthetic schemes. Additionally, the linkage to the solid support using the terminal carboxylic acid group in the usual acid based linker strategy offers the introduction of further diversity by linking the functional group with an additional building block R^3, e.g., a further amino acid, to the solid support. The decoration concept of the cyclic peptides is further diversified when different orthogonally protected core structures are attached to a resin via one of the functional groups, while the reactive groups are transformed during the course of the synthetic sequence. The complexity of these secondary structure based libraries is further increased when the amino acids R^5 contain functionalities which can be elaborated with additional sets of building blocks. In the light of these examples, we intend to emphasize the importance of the natural product pool as a rich source of unique molecular frameworks, some of which can be employed as peptide structure mimetic molecular cores. By taking advantage of the orientational properties of a natural product for the design of structurally defined peptidomimetics, the following requirements must be fulfilled:

(i) 3D structural correspondence of the natural product core with regular secondary structure elements of peptides;
(ii) the pattern of functional groups allows to mount amino acid side-chain-related substituents in an iso-structural manner compared to a native parent peptide; and
(iii) chemical feasibility to set up a multi-dimensional combinatorial library applying an orthogonal synthetic scheme.

Summarizing, we propose to envision natural products that mimic sterically the bioactive conformation of peptide or protein ligands as privileged molecular frameworks for the design of secondary structure-templated libraries. This conceptual procedure, exemplified in this contribution with the turn and sheet mimics **1** and **2**, respectively, undoubtly enriches the newly established combination of structure-based design with combinatorial chemistry in conjunction



3 **4**

Figure 4. Chemical formulas of **3** and **4**.

with high-throughput screening technology [26, 27] by a further facet, namely the natural product-encoded diversity and uniqueness of privileged core structures. The structural theme of side-chain-cyclized di- and tripeptides discussed above can be developed further by e.g. conformationally restricted analogues of the indolactam type (**3**), which represent the common core structure of the naturally occurring tumor-promoting teleocidins [28].

The indolactam core structure **3** displays a decoration pattern which allows to set up a multidimensional combinatorial library, based on a natural product-derived, privileged template with unique spatial presentation properties.

As a last example for the structure-based use of natural products as peptide secondary structure mimetics, we refer to the work of Kessler et al., employing carbohydrate core structures (**4**) for β turn mimicry [29]. Compounds of type **4** meet all the requirements defined in this perspective for discovering protein secondary structure templates from natural products with the aim to design secondary structure-templated combinatorial libraries for lead finding and optimization within pharmaceutical research.

Concluding, we strongly believe that it is this combination of speed and sampling power of combinatorial chemistry with the natural product-derived structural information on receptor- and enzyme-complementarity that will aid the drug discovery process in the near future, in that the probability of finding new leads with unique core structures is increased. Whether the outlined ideas will evolve to a generic template-finding strategy for focused libraries will depend on the interplay between medicinal chemistry, natural product chemistry and molecular modeling. Within such a multidisciplinary approach, computer-aided molecular design will play a key role in uncovering appropriate natural products, in evaluating structural correspondence with secondary structure elements, and, if necessary, in modifying complex natural product frameworks in order to meet the above defined requirements.

References

1. Verlinde, C.L.M.J. and Hol, W.G.J., Structure, 2 (1994) 577.
2. Petsko, G.A., Nature Supplement (intelligent drug design) 384 (1996) 7.
3. See for example: Branden, C. and Tooze, J. (Eds.) Introduction to Protein Structure, Garland Publishing, Inc., New York, NY, 1991.
4. Horwell, D.C., Bioorg. Med. Chem., 4 (1996) 1573.
5. Adang, A.E.P., Hermkens, P.H.H., Linders, J.T.M, Ottenheijm, H.C.J. and van Staveren, C.J., Recl. Trav. Chim. Pays-Bas, 113 (1994) 63.
6. Wiley, R.A. and Rich, D.H., Med. Res. Rev., 13 (1993) 327.
7. Balkenhohl, F., Bussche-Hünnefeld, C. v.d., Lansky, A. and Zechel, C., Angew. Chem. Int. Ed. Engl., 35 (1996) 2288.
8. Thompson, L.A. and Ellman, J.A., Chem. Rev., 96 (1996) 555.
9. Boger, D.L., Zhou, J., Borzilleri, R.M. and Nukui, S., Bioorg. Med. Chem. Lett., 6 (1996) 1089.
10. Boger, D.L. and Zhou, J., Bioorg. Med. Chem., 4 (1996) 1597.
11. Boger, D.L., Zhou, J., Borzilleri, R.M., Nukui, S. and Castle, S.L., J. Org. Chem., 62 (1997) 2054.
12. Janetka, J.W. and Rich, D.H., J. Am. Chem. Soc., 117 (1995) 10585.
13. Janetka, J.W., Satyshur, K.A. and Rich, D.H., Acta Crystallogr., Sect. C (1996) 3112.
14. Janetka, J.W., Raman, P., Satyshur, K.A., Flentke, G.R. and Rich, D.H., J. Am. Chem. Soc., 119 (1997) 441.
15. Itokawa, H. and Takeya, K., Heterocycles, 35 (1993) 1467.
16. See for example: Müller, G., Gurrath, M., Kurz, M. and Kessler, H., Proteins Struct. Funct. Genet., 15 (1993) 235.

17. Evans, D.A. and Ellman, J.A., J. Am. Chem. Soc., 111 (1989) 1063.
18. Boger, D.L. and Yohannes, D., J. Org. Chem., 55 (1990) 6000.
19. Boger, D.L. and Zhou, J., J. Am. Chem. Soc., 115 (1993) 11426.
20. Boger, D.L., Patane, M.A. and Zhou, J., J. Am. Chem. Soc., 116 (1994) 8544.
21. Boger, D.L. and Borzilleri, R.M., Bioorg. Med. Chem. Lett., 5 (1995) 1187.
22. Boger, D.L., Borzilleri, R.M., Nukui, S. and Beresis, R.T., J. Org. Chem., 62 (1997) 4721.
23. Beugelmans, R., Bigot, A., Bios-Choussy, M. and Zhu, J., J. Org. Chem., 61 (1996) 771.
24. Zhu, J., Synlett, 2 (1997) 133.
25. Burgess, K., Lim, D., Bois-Choussy, M. and Zhu, J., Tetrahedron Lett., 38 (1997) 3345.
26. Qabar, M., Urban, J., Sia, C., Klein, M. and Kahn, M., In Chaiken, I.M. and Janda, K.D. (Eds.), Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery, American Chemical Society, Washington, DC, 1996, p. 2.
27. Graybill, T.L., Agrafiotis, D.K., Bone, R., Illig, C.R., Jaeger, E.P., Locke, K.T., Lu, T., Salvino, J.M., Soll, R.M., Spurlino, J.C., Subasinghe, N., Tomczuk, B.E. and Salemme, F.R., In Chaiken, I.M. and Janda, K.D. (Eds.), Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery, American Chemical Society, Washington, DC, 1996, p. 16.
28. Irie, K., Isaka, T., Iwata, Y., Yanai, Y., Nakamura, Y., Koizumi, F., Ohigashi, H., Wender, P.A., Satomi, Y. and Nishino, H., J. Am. Chem. Soc., 118 (1996) 10733.
29. Graf von Roedern, E. and Kessler, H., Angew. Chem. Int. Ed. Engl., 33 (1994) 687.