Molecular Diversity, 2 (1996) 46–56
ESCOM

MOLDIV 036

Design considerations and computer modeling related to the development of molecular scaffolds and peptide mimetics for combinatorial chemistry

Victor J. Hruby[a],*, Mark Shenderovich[a], Kit S. Lam[b] and Michal Lebl[c],**

[a]*Department of Chemistry, University of Arizona, Tucson, AZ 85721, U.S.A.*
[b]*Arizona Cancer Center, University of Arizona, Tucson, AZ 85724, U.S.A.*
[c]*Selectide Corporation, a subsidiary of Hoechst Marion Roussel, 1580 E. Hanley Boulevard, Tucson, AZ 85737, U.S.A.*

Received 13 February 1996
Accepted 20 June 1996

Keywords: Scaffold design; Backbone conformation; Topography; Chi space; Topographical design

Summary

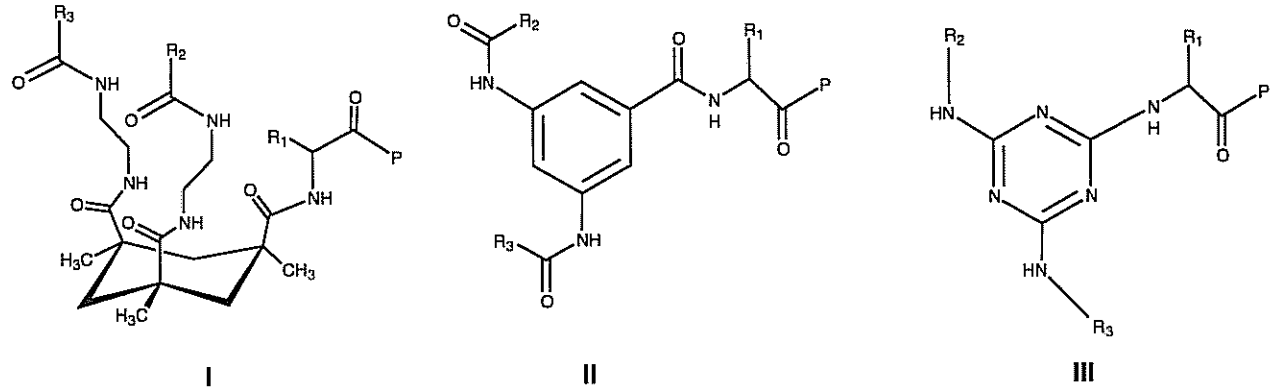
A critical issue in drug discovery utilizing combinatorial chemistry as part of the discovery process is the choice of scaffolds to be used for a proper presentation, in a three-dimensional space, of the critical elements of structure necessary for molecular recognition (binding) and information transfer (agonist/antagonist). In the case of polypeptide ligands, considerations related to the properties of various backbone structures (α-helix, β-sheets, etc.; ϕ, ψ space) and those related to three-dimensional presentation of side-chain moieties (topography; χ (chi) space) must be addressed, although they often present quite different elements in the molecular recognition puzzle. We have addressed aspects of this problem by examining the three-dimensional structures of chemically different scaffolds at various distances from the scaffold to evaluate their putative diversity. We find that chemically diverse scaffolds can readily become topographically similar. We suggest a topographical approach involving design in chi space to deal with these problems.

Introduction

Peptides, linear polypeptide fragments from proteins, and discontinuous peptide fragments from larger proteins often constitute the targets of modern drug-discovery efforts. Thus, the desire to develop peptidomimetic, and non-peptide mimics of peptides has become a goal of many drug-discovery efforts. A central issue that has arisen in such studies is the choice of scaffolds for displaying 'diverse' libraries. This is particularly important in peptide and non-peptide mimetic design, since the molecular recognition process involves both backbone structure and side-chain moieties of specific amino acid residues in the molecular recognition process, and for information transfer (transduction) as well. It is critical that the scaffolds used will provide structures that will lead to the proper presentation, in three-dimensional (3D) space, of the key elements of structure necessary for molecular recognition and information transduction. This issue is further complicated depending on whether an agonist or antagonist biological activity is desired, since the stereostructural requirements for agonists and antagonists generally are different.

In the case of polypeptide-mimetic ligands, two structural problems must be considered in the design process. The first is the requirement for an appropriate secondary structure (α-helix, β-turn, β-sheets, etc.). The second is the appropriate presentation, in 3D topographical space (chi space), of the side-chain moieties involved in molecular recognition, and then in signal transduction if an agonist activity is desired. This brings many challenges, and to date very few de novo peptidomimetic designs have considered both of these aspects (e.g. Refs. 1 and 2; however, see Ref. 3), especially in terms of proper scaffold design. The purpose of this paper is to discuss some of the problems and approaches to rational scaffold design,

*To whom correspondence should be addressed.
**Present address: Houghten Pharmaceuticals Inc., 3550 General Atomics Court, San Diego, CA 92121, U.S.A.



I II III

Fig. 1. Three chemically different scaffolds used in these studies: I: Kemp's triacid derivatives; II: 1,3,5-trisubstituted aromatic derivatives; and III: 2,4,6-trisubstituted triazine derivatives.

with particular emphasis on the interplay of structural considerations, computational methods, and careful analysis of structure–biological activity relationships.

Materials and Methods

Synthesis of specialized amino acids, peptides, scaffolds and libraries

The specialized α-amino acids utilized in this paper were prepared by methods developed in the Hruby laboratory and have been published previously (e.g. Refs. 4–6), or will be published elsewhere. Peptides were prepared by standard methods of solid-phase peptide synthesis utilizing either an N^{α}-Boc/benzyl protection strategy or an N^{α}-Fmoc/*t*-butyl strategy. The purity of the peptides after synthesis and purification was assessed by amino acid analysis, thin layer chromatography (TLC) in three or four different solvent systems, high-performance liquid chromatography (HPLC) in two different systems, by mass spectrometry, and by nuclear magnetic resonance (NMR) spectroscopy. Biological assays were performed by literature procedures.

All of the organic acids, amines, aromatics, and other organic compounds used to construct combinatorial libraries are commercially available, or precursors can be purchased commercially and converted to these simple organic molecules by well-established chemical methods. The syntheses of specially modified scaffolds were done at Selectide Corporation (Tucson, AZ, U.S.A.) using well-precedented organic synthetic methods, and the methodology will be reported elsewhere. The construction of the chemical libraries reported here were mostly accomplished at Selectide Corporation using the Selectide method of one-bead–one-peptide construction [7], and the evaluation of binding to Streptavidin utilized methods previously reported [7–9].

Computational methods used to evaluate scaffold conformations, to compare structures, and to evaluate chi space in unusual amino acids

Molecular modeling in this work has been performed on the Silicon Graphics workstations using the MacroModel molecular modeling system [10,11], v. 4.0 and 4.5.

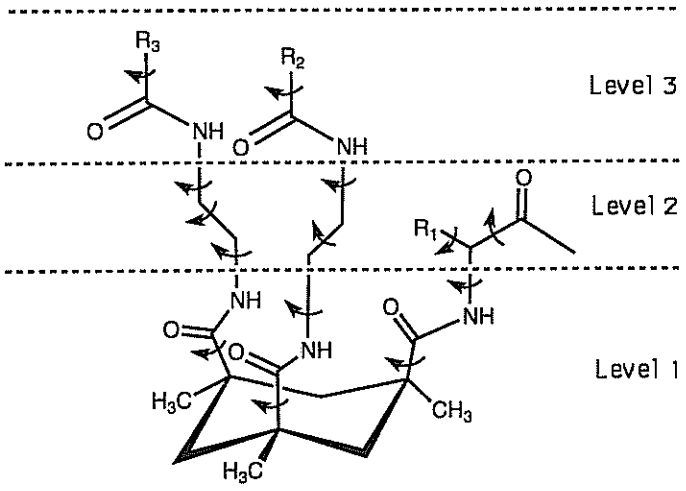


Fig. 2. The level approach to molecular modeling of Kemp's triacid scaffold I. Arrows indicate variable torsional angles included into systematic conformational searches at each level.

TABLE 1
CHEMICAL/STEREOELECTRONIC DIVERSITY OF PEPTIDES AND PROTEINS

Chemical functionality

Side-chain and backbone groups contain acids, bases, nucleophiles, electrophiles, H-bond donors, H-bond acceptors, aromatics, aliphatics, alcohols, thiols, heteroaromatics (plus whatever we want to add: metals, organic ligands, etc.)

Structures (different possible sequences with L-amino acids)

1. Dipeptides: $20^2 - 4 \times 10^2$
2. Decapeptides: $20^{10} - \sim 1 \times 10^{13}$
3. A 100-peptide: $20^{100} - \sim 10^{130}$

The mass necessary to produce 1 mg of all 100-peptide sequences: $\sim 10^{124}$ kg
The mass of the earth: 6×10^{24} kg

Conformation/topology

A. Chirality
At least one chiral center in all common amino acids but glycine
B. Phi/psi space
1. Very conservative: 10 low-energy structures
10^6 different conformations for a hexapeptide
C. Chi space
χ_1 and χ_2: six different low-energy conformations; others accessible

Molecular modeling and search for low-energy conformers of scaffold compounds For the modeling of the ring-based scaffolds I, II and III (Fig. 1), a build-up procedure was developed which first considered the rings with directly attached carboxamide substituents (Level 1), and then gradually extended the substituent chains (Levels 2, 3, etc.) until entire scaffolds used in synthetic combinatorial libraries were constructed. This approach is illustrated in Fig. 2 with the example of Kemp's triacid scaffold I. Variable R groups of the scaffolds were in most cases mimicked by methyl groups. For the cyclohexane-based scaffold I the ideal chair conformation of the ring was taken as the initial approximation, while both axial and equatorial orientations of each substituent were considered.

Conformational searches at each level of the build-up scheme were performed using the MULTIC procedure [12] of the MacroModel software [10,11], which allows one to collect sterically feasible conformers by means of a grid search, energy-minimize them, and then select a set of geometrically different low-energy conformers. Conformers selected at lower levels of the build-up scheme were used as starting structures at higher levels. At the first level of the build-up procedure, energy minimization was performed using the MM3 force-field [13] which had been developed for aliphatic and aromatic hydrocarbons. The all-atom AMBER* force-field [14,15] implemented in the MacroModel program was employed for the ring-based scaffolds with longer substituent chains, which was supposed to be more adequate for compounds rich in amide bonds. The distance-dependent dielectric constant $\varepsilon = 4.0\, r_{ij}$, where r_{ij} are interatomic distances, was used with the AMBER* force-field in all calculations of this work. Energy minimization was performed using the Polak–Ribiere conjugated gradient (PRCG) algorithm [16] with a maximum number of iterations of 1000 at earlier stages of the build-up procedure and of 5000 at final stages.

For the Kemp's triacid and 1,3,5-(*S*)-triazine compounds (Fig. 6), which were shown (see Results and Discussion) to possess considerable binding affinity to Streptavidin, molecular dynamics (MD) simulations at 500 K were performed with the AMBER* force-field. Starting conformations for the MD simulations were selected among the low-energy conformer of scaffolds I and III after addition of the terminal groups R_1, R_2, and R_3, and complete energy minimization with the full matrix Newton–Raphson (FMNR) algorithm [17]. The SHAKE [18] procedure was applied during MD simulations in order to maintain correct bond lengths to hydrogen atoms. Sample conformers were stored each 1 ps of the 100 ps MD trajectories and energy-minimized using the PRCG optimizer [16]. Geometrically different low-energy conformers ($E - E_{min} \leq 50$ kJ/mol) were selected for comparison and classification.

Comparison of conformers and classification of scaffolds Comparison of low-energy conformers of the same scaffold or of different scaffolds was performed using the best-fit matching [19] of atoms belonging either to the common elements of scaffolds (rings, etc.) or to the end groups of side chains. Rms deviations upon matching of end groups were considered as a measure of similarity of conformers. At lower levels of the build-up scheme described above, terminal methyl groups were used to substitute for more distant fragments of side chains. Distances between end groups and distances from the center of a scaffold ring to end groups, were chosen for classification of scaffolds. The minimum, maximum and average end-to-end distances calculated with the FILTER module of the MacroModel program were shown to be distinct for different types of scaffolds, as well as for the axial and equatorial orientation of substituents. These distances were used for the preliminary classification of scaffolds.

The (χ_1,χ_2) energy maps The energy maps were calculated using the Drive module of MacroModel v. 4.5 [11]. The maps were constructed with 20° steps from −180° to 180° both in the χ_1 and χ_2 directions. At each (χ_1,χ_2) grid point, the energy was minimized over all other degrees of freedom using the united-atom AMBER* force-field [14] with a distance-dependent dielectric constant $\varepsilon = 4.0\ r_{ij}$. Extended initial backbone conformations with $\phi,\psi = 180°$ were used for energy minimization at each grid point. Contours of equal relative energies $E - E_{min}$ were drawn with the 1.0 or 2.0 kcal/mol steps using the Plt2D module of MacroModel 4.5. The (χ_1,χ_2) maps were calculated for *N'*-methyl derivatives of tyrosine and TMT residues, while a free amino group in a neutral form was chosen to mimic the incorporation of these residues into the N-terminal position of peptides.

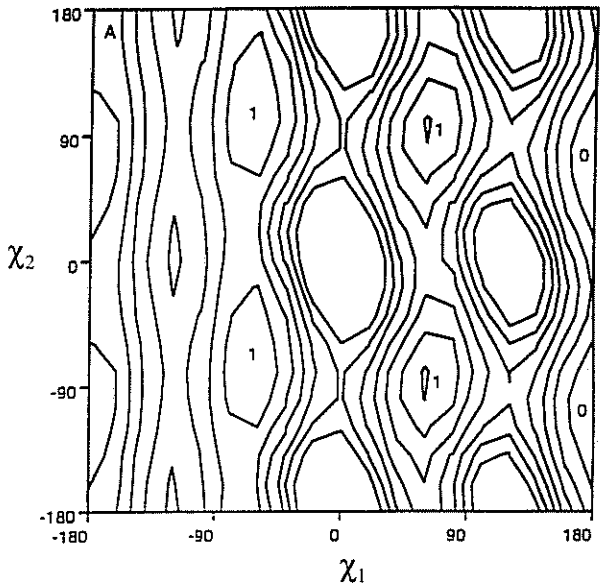


Fig. 3. χ_1 versus χ_2 energy map for tyrosine. The cutoff value for energies is at the 5 kcal/mole level.

Results and Discussion

The difficulty of the problem of scaffold design and its use for obtaining peptidomimetic ligands is often overlooked by medicinal chemists. For example, it has been suggested or implied that polypeptides are not sufficiently diverse. In fact, they are extremely diverse. For example, as shown in Table 1, there is not sufficient matter in our galaxy to examine the primary structural diversity of even a 100-amino-acid peptide much less its conformational diversity. Furthermore, natural peptides and proteins possess all of the structural moieties necessary for binding and reactivity, and additional structural elements such as metal ions, small organic molecules, etc., can be readily accommodated into these structures. Thus, peptides and proteins have been chosen by living systems to be its catalysts, structural motifs, information transducers, surveillance systems, and so forth. A simplification and proper focusing of the astronomical number of possibilities for peptidomimetic design is needed. As discussed in the introduction, therefore, insight into the peptide backbone structure (α-helix, β-turn, etc.) and topographical structure (chi space) is central to peptidomimetic scaffold design. With regard to the backbone conformation, the seminal insight of Ramachandran [20] that the low-energy structural features of a peptide backbone depended primarily on the ϕ and ψ angles, and that the α-helix, β-sheet, β-turn and extended conformations were the low-energy backbone conformations, is critical. It seemed to us reasonable to suppose, therefore, that most polypeptide ligands would assume one of these conformations [21] for bioactivity. Indeed, we have made this assumption as part of our design from the beginning, and suggested that conformational constraints that led to stabilization of these structures could provide insight into the bioactive conformations [3,22]. Within this context, the disposition of the side-chain group of amino acids in 3D space becomes an important focus. These will not be the same for different secondary structures of the same sequence. Moreover, different receptors for the same sequence (for example, different opiate receptors) might chose different low-energy conformations of the ligand (e.g., enkephalins) on molecular interaction. This realization can be utilized for de novo peptide design (e.g. Refs. 3, 22), and has begun to be considered in peptidomimetic design [23]. The chi space available for a typical natural amino acid is

TABLE 2
LIGANDS THAT BIND TO THE BIOTIN BINDING SITE IN STREPTAVIDIN[a]

1	Biotin
2	2-(4'-Hydroxyphenylazo)benzoic acid (HABA)
3	XHPQX, HPQXX, XXHPQ, HPMXX
4	XxWpH, XwFpH, XwYpH, YxIfP
5	wyque, wyhea, wyfya, wydya, wyefa
6	H-Cys-XHPM-Cys-OH H-Cys-HPQXX-Cys-OH H-Cys-HPQXXX-Cys-OH H-Cys-XHPQX-Cys-OH Ac-Cys-HPQXXX-Cys-NH_2

[a] See text for references which report the binding of these ligands.

illustrated in Fig. 3. Here we depict a topographical energy profile for tyrosine (plot of χ_1 versus χ_2). The figure illustrates that tyrosine has a preference in chi space for χ_1 of about –60°, 60° and ±180° (*gauche (–)*, *gauche (+)* and *trans* rotamers, respectively), and for χ_2 of about ± 90°. Though other conformations are accessible at physiological temperatures, it would be anticipated that in intramolecular and intermolecular interactions, these low-energy states would be the preferred, and indeed distributions of χ_1 in peptide and protein structures are consistent with this assumption. From these ideas, it follows that conformational constraints in side-chain moieties consistent with these favorable structures in chi space should provide ligands with a high degree of potency, receptor selectivity, and other useful biological properties, such as stability against biodegradation. An approach we have developed towards these goals is discussed below.

The Streptavidin–biotin complex provides an excellent test system for examining the relationships between scaffold design and side-chain topology. On the one hand, an X-ray structure of the Streptavidin–biotin complex is available [24]. At the same time, it has been shown, using both biological methods (phage display [25]) and chemical methods (the Selectide process [7]), that Streptavidin recognizes a consensus peptide sequence, the linear tripeptide HPQ (Table 2). We therefore wanted to know whether a cyclic peptide could be used as a scaffold for binding to the biotin binding pocket. We examined cyclic disulfide hexapeptide, cyclic heptapeptide and cyclic octapeptide libraries [8], that is, disulfide-containing peptides with respectively 4, 5 and 6 residues between the two cysteine residues, and found (Table 2) [8] numerous peptides that bound to the biotin binding site. Again, the HPQ sequence was found among the hits, as well as another tripeptide sequence, HPM. More recently, the cyclic octapeptide system Ac-CXXXXXXC-NH_2 (where X is an amino acid residue) also was explored by Katz [26], who found the same HPQ sequence (Table 2). In the meantime, other peptide motifs have been investigated to determine whether changes in the configuration would result in a change in the structural motif that is recognized by the biotin binding site in Streptavidin. For example, we have explored alternating L,D,L,D,L-pentapeptides in a highly diverse library format using the Selectide process [8]. As shown in Table 2, new consensus sequences emerged, including the XpH motif (X = W, F or Y). In addition, we have examined diverse, all D-amino acid peptide libraries with very interesting results [8]. Essentially entirely new motifs were discovered of which the sequence wyqxx was common (Table 2). In fact, the tripeptide sequences wyx (where x = q, h, d, e, f) all bind to Streptavidin. Some insights are already available in this regard. X-Ray crystal structures have been reported for complexes of Streptavidin with biotin [24] and 2-(4'-hydroxyphenylazo)benzoic acid (HABA) [27,28], for a complex with an HPQ-containing peptide, Phe-Ser-His-Pro-Gln-Asn-Thr [26,29], and for cyclic disulfide [26] and thioether [30] peptides containing the HPQ sequence. The binding of all of these ligands to Streptavidin apparently results in substantial changes in the conformation of the protein. There is no evidence for cooperativity of binding from either the X-ray studies [24] or other biophysical studies [31]. The various ligands, however, do bind somewhat differently, as briefly outlined in Table 3. In the case of biotin, the ligand is essentially completely enveloped by the protein, and extensive H-bonding (see Table 3) and van der Waals interactions between the ligand and Streptavidin further stabilize the binding interaction. The H-bonding patterns to the HABA analogues and HPQ peptides differ from the biotin one considerably, and from each other to some extent. In the peptide cases, though, the His-Pro-Gln portion of the ligand shows a similar bonding interaction. The rest of the peptide (the 'wings') remain more flexible, and a bound water molecule is found in the complex for both the HABA and peptide sequences (Table 3). It will be interesting to see whether the all-D-amino-acid-containing ligand binds in yet another way to the biotin binding pocket in Streptavidin. In collaboration with Dr. P.C. Weber, a crystal for one of these peptides bound to Streptavidin has been obtained, and its structure is being examined by X-ray crystallography.

The different binding modes of ligands to the same binding site of a protein as exemplified in the Streptavidin case, raises issues regarding the use of different scaffolds for probing the conformational space of ligands which can be used for molecular recognition. Clearly, the implications seem to be that a binding site of a protein can utilize a variety of different 'presentations' of structure for molecular recognition, but on the other hand, suggest that for optimal binding, perhaps only a single presentation may allow an optimal interaction and provide selec-

TABLE 3
BINDING MODES TO STREPTAVIDIN

I **Biotin**
- Flexible wing 47–51 of Streptavidin: becomes ordered
- Biotin H-bonding to Streptavidin of Asp^{128}, Thr^{90}, Ser^{27}:
  - to ureido-ring (Asp, Ser) and biotin-S (Thr);
  - binds as negatively charged species

II **2-(4'-Hydroxyphenylazo)benzoic acid**
- H-bonding of Asn^{23}, Tyr^{43}, Ser^{27}, Ser^{45}: all to carboxylate
  - Ligand–H_2O–Tyr^{54}: H-bonded

III **HPQ – Phe-Ser-His-Pro-Gln-Asn-Thr**
- Flexible wing 47–51: stays flexible
- H-bonding - Asp^{128}, Thr^{90}, Ser^{27} + H_2O
  - to His - Thr^{90} and Asp^{128} (via H_2O)
  - to Gln - Ser^{27}
  - H_2O - between His and Gln side chain
- Pro, tetrahydrothiophane ring and valeric acid aliphatic side chain: same location in hydrophobic pocket

IV **Cyclic HPQ – Ac-c[CHPQGPPC]-NH_2 or valeramide analogue** binds about the same as HPQ: involving Asp^{128} and Thr^{90}

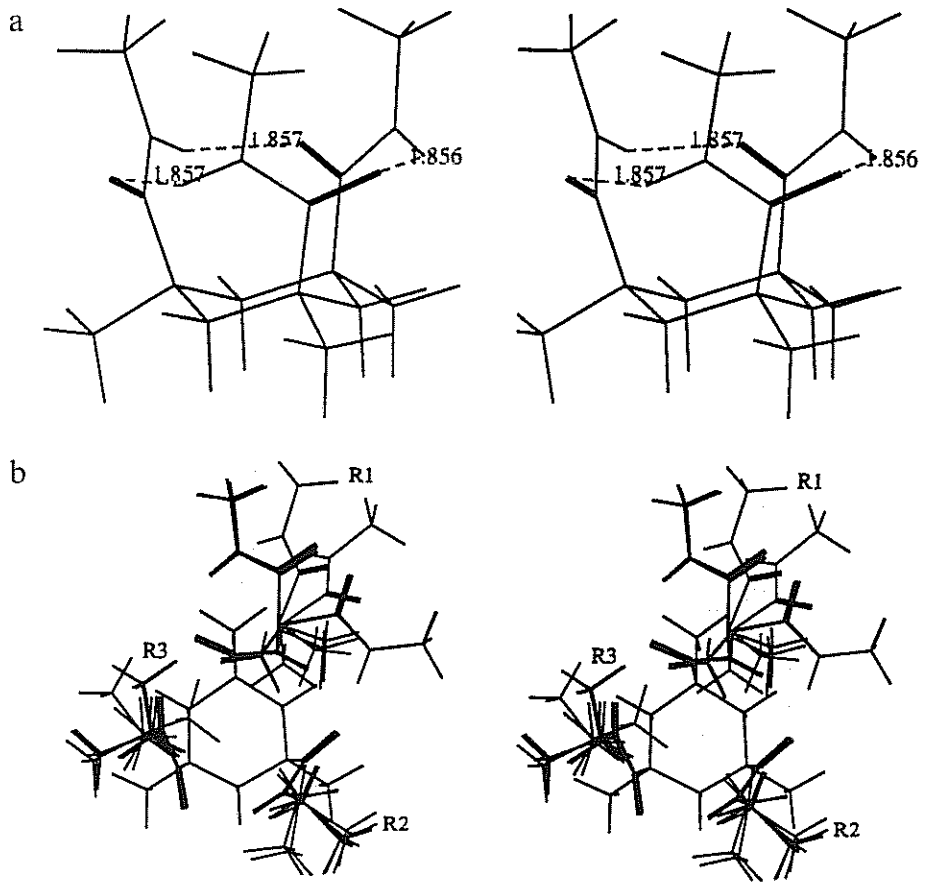


Fig. 4. (a) Stereoview of the lowest-energy conformation of scaffold I, Level 1; (b) stereoview of scaffold I, Level 2: superposition of low-energy conformers.

tivity for closely related analogous structures. This raises two issues, which we have begun to examine. First, will 'chemically diverse' scaffolds lead to diverse binding interactions when similar 'side-chain' groups are present on the scaffold? Second, can side-chain group constraints, applied to specific torsional angles in topographical space, provide a reliable approach for developing a preferred binding conformer, which would provide an accurate 3D model of the pharmacophore for a particular ligand interacting with a particular binding site in a macromolecular receptor/acceptor? These questions need to be critically addressed for future success in de novo design of ligands for target biological macromolecules. In terms of drug discovery using combinatorial methods, a further question is whether it is necessary to examine a highly diverse group of scaffolds on which to build side-chain pharmacophores, or whether a few well-chosen scaffolds will suffice? In other words, how should the conformational diversity of scaffolds, i.e. their ability to present chemically diverse functional groups in a variety of different posi-

TABLE 4
DISTANCES BETWEEN END GROUPS IN SCAFFOLDS I, II AND III AT LEVEL 1[a]

Scaffold	Isomer	Distance (Å)	Average (SD)	Minimum	Maximum
I	1,3,5-axial	R_{ax}–R_{ax}	5.1 (0.9)	3.7	5.8
	1-equatorial, 3,5-axial	R_{ax}–R_{ax}	5.5 (1.0)	3.9	6.8
		R_{ax}–R_{eq}	8.5 (0.8)	7.1	9.4
	1-axial, 3,5-equatorial	R_{ax}–R_{eq}	8.6 (0.3)	7.4	9.5
		R_{eq}–R_{eq}	8.2 (0.9)	6.9	9.8
	1,3,5-equatorial	R_{eq}–R_{eq}	8.2 (0.9)	6.9	9.8
II		R_1–R_2–R_3	8.8 (0.6)	7.8	9.7
III		R_1–R_2–R_3	6.3 (0.9)	4.9	7.3

[a] Distances between the carbon atoms of terminal methyl groups were calculated and averaged for all low-energy conformers of each scaffold ($E - E_{min} \leq 25$ kJ/mole). R_{ax} and R_{eq} represent axial and equatorial substituents, respectively, in scaffold I. The R_1, R_2 and R_3 substituents in scaffolds II and III at Level 1 are identical.

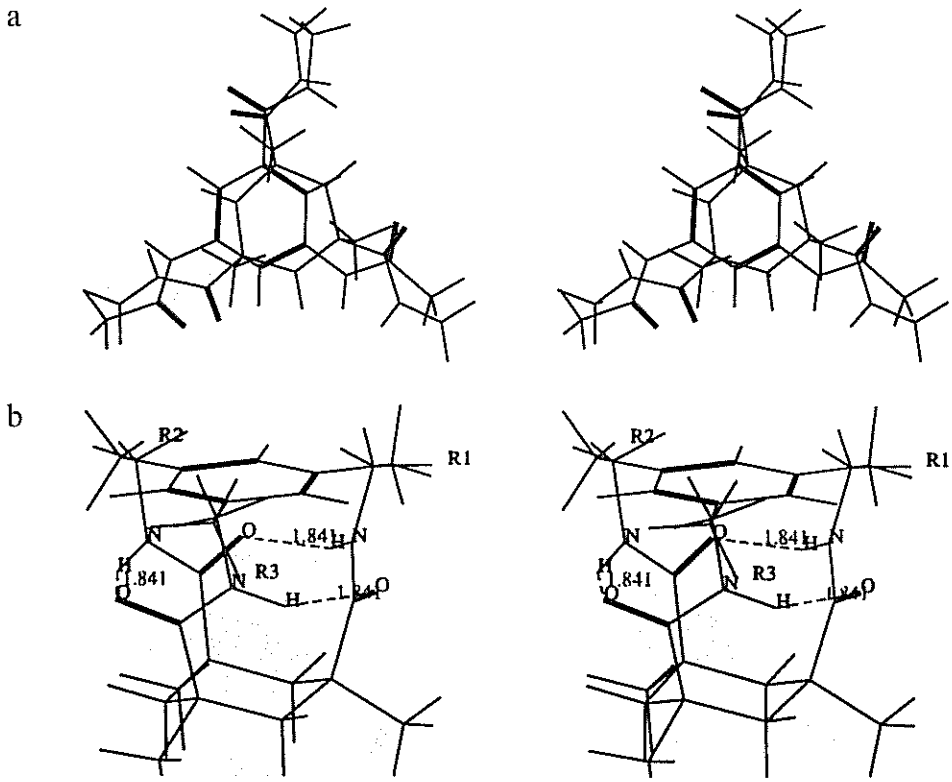


Fig. 5. (a) Stereoview of scaffold I (equatorial substituents) and scaffold II: superposition of the lowest-energy conformers at Level 1; (b) stereoview of scaffold I (axial substituents), Level 1: matching by a 1,3,5-R_3-benzene derivative.

tions and orientations in 3D space, be considered in the rational design of combinatorial libraries? We have addressed this point by examining the conformational space available to several chemically different structures, three of which are shown in Fig. 1, at various topographical levels (see Fig. 2) from the scaffold.

Low-energy conformations found for the Kemp's triacid scaffold I at Levels 1 and 2 are shown in Fig. 4. Although scaffold I with all axial substituents has a clearly defined lowest-energy conformation at Level 1, which is stabilized by dipole–dipole interactions and hydrogen bonding (Fig. 4a), considerable 3D space is covered already by terminal side-chain groups at Level 2 (Fig. 4b), and the topographies available for the functional R-groups further removed from the scaffold will soon encompass all the 3D space at one side of the cyclohexyl ring. Furthermore, each side chain of scaffold I may be attached to the ring either in an axial or in an equatorial configuration. Therefore, one should consider several stereoisomers of this scaffold with all possible combinations of axial and equatorial substituents. Even with our symmetrical presentation of all terminal R-groups by methyl groups (see Materials and Methods), it results in four different isomers of scaffold I listed in Table 4. This stereochemical diversity, together with the relative flexibility of side chains, results in the high conformational diversity of scaffold I.

Table 4 presents average and limit values of distances between end groups (represented by methyl carbons) in low-energy conformers ($E-E_{min} \leq 25$ kJ/mole) obtained by a systematic search for four isomers of scaffold I and for scaffolds II and III at Level 1. Note that distances between two axial substituents of scaffold I (from less than 4 to about 7 Å) do not overlap the range of distances available to axial–equatorial and equatorial–equatorial pairs (7 to 10 Å). Therefore, the all-axial and all-equatorial isomers exemplify two topographically distinct versions of scaffold I, which are not expected to place their functional R-groups in the same regions of 3D space. On the other hand, all four isomers considered together comprise the highly conformationally diverse scaffold, in which the R-groups are able to cover the entire 3D space around the central ring, limited only by the lengths of the side chains. Of course, the stereoisomers of scaffold I may not be equally populated. Our estimates with the MM3 force field in vacuo showed that, at Level 1, the all-axial isomer has the lowest potential energy due to the favorable electrostatic interactions between substituents (see Fig. 4a), and transition of one substituent to an equatorial configuration costs about 10 kJ/mole. Therefore, the all-axial isomer is expected to be considerably more populated than the all-equatorial isomer. However, solvation effects which may decrease electrostatic interactions and entropy effects which favor structures with a higher conformatio-

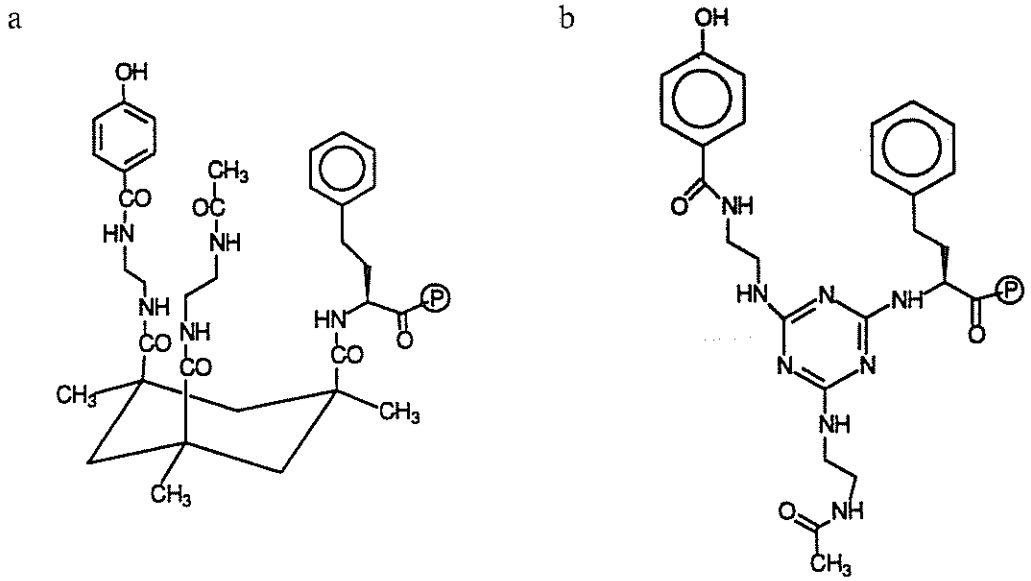


Fig. 6. Structures of non-peptide ligands that bind to Streptavidin. (a) The Kemp's triacid derivative; (b) the triazine derivative designed from (a).

nal mobility may result in a more uniform population of isomers at higher levels of scaffold I. Therefore, the stereochemical diversity should be taken into account for rational design of combinatorial libraries based on Kemp's triacids and similar scaffolds.

Distances between end groups of scaffold II at Level 1 (Table 4) overlap exactly with the range of distances available for a pair of equatorial substituents of scaffold I. Therefore we expect that almost the same 3D space will be covered by R-groups of scaffold II and of the all-equatorial isomer of scaffold I. This is illustrated in Fig. 5a, which shows a very close overlap of the lowest-energy conformers found for these two scaffolds at Level 1. However, because the all-equatorial isomer represents only one possible (and, perhaps, not the most favorable) topography of scaffold I, conformational diversity of scaffold II seems to be considerably limited in comparison with scaffold I. Distances between end groups of scaffold III overlap in part the ranges available for two axial and for two equatorial substituents. Therefore it is expected that the R-groups attached to scaffold III will cover part of, but not the entire 3D space available for R-groups attached to scaffold I. It is noteworthy that the aromatic-ring-based scaffolds II and III cannot match well the all-

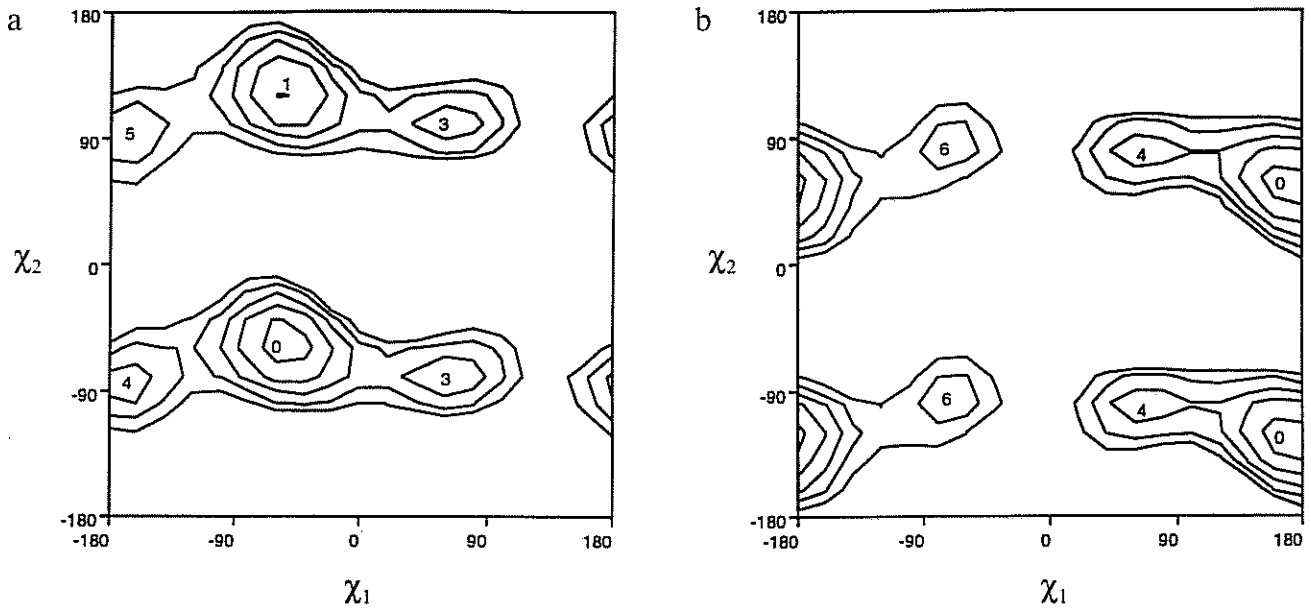


Fig. 7. χ_1 versus χ_2 energy maps for: (a) (*2S,3S*)-β-methyl-2',6'-dimethyltyrosine; and (b) (*2S,3R*)-β-methyl-2',6'-dimethyltyrosine. The cutoff value for energies is at 9 kcal/mole.

axial isomer of scaffold I. We found, however, that a scaffold that mimics this isomer may be designed based on a 1,3,5-methyl-substituted aromatic derivative which methyl groups match exactly the end groups of the all-axial isomer at Level 1 (see Fig. 5b).

The implications of these observations are far-reaching. They demonstrate that in considering template design, if one wishes to have differences in the conformational space addressed by a scaffold, it is important that the topographical space be examined carefully as part of the design process. In this regard, in terms of topography and depending on stereochemistry, scaffolds I, II and III can have many similarities, and are not diverse enough. To examine this hypothesis, we have utilized a structure that binds to Streptavidin which was found in a Kemp's tri-acid library (Fig. 6a), and have prepared a 1,3,5-(*S*)-triazine structure (Fig. 6b) that is related to it with similar side-chain groups. Structure b binds to Streptavidin. Results of preliminary molecular dynamic simulations (see Materials and Methods) indicate that these diverse compounds have a limited overlap in 3D space and hence may have different modes of binding to Streptavidin. Therefore, one must carefully consider the topographical as well as structural differences in scaffold design. Thus, it is critical that methods be developed for designing structural moieties into amino acids, dipeptides, amino acid mimetics and peptide mimetics that provide real chemical diversity in 3D space. For this purpose, we have been developing methods for the design and synthesis of novel amino acid derivatives and mimetics which constrain side-chain moieties in 3D space to specific χ angles (chi space).

A good example of an amino acid designed for this purpose is β-methyl-2',6'-dimethyltyrosine (TMT) which has four different isomers: (*2S,3S*), (*2S,3R*), (*2R,3S*), and (*2R,3R*). These amino acids were designed to constrain the tyrosine residue at both χ_1 and χ_2 side-chain torsional angles and to examine the effect of constraints in chi space on potency, receptor selectivity, and biological activity of bioactive peptides. A further goal has been to evaluate the importance of chi space in de novo design of peptidomimetics. The (χ_1/χ_2) energy maps for (*2S,3S*)- and (*2S,3R*)-β-methyl-2',6'-dimethyltyrosine are shown in Figs. 7a and 7b. Three χ_1 rotamers are still available for these amino acids: *gauche(−)* (−60°), *gauche(+)* (+60°) and *trans* (±180°). However, there is a considerable preference for one of these χ_1 conformers, and barriers for rotation of χ_1 are much higher than in tyrosine (compare with Fig. 3 where the energy cut-off value is 5 kcal/mole). The χ_2 torsional angle is constrained in regions around ± 90°, and, in agreement with our earlier dynamic NMR investigations [32], barriers of 15–20 kcal/mole were observed for rotation about the χ_2 angle (not shown in Fig. 7 where the energy cut-off value is 9 kcal/mole). Most interesting, as can be seen in Fig. 7, there is a good deal of chi space (about 70%) that is not available for the TMT residues at the 9 kcal/mole cut-off value. Thus, the initial goal to design amino acid residues constrained in both χ_1 and χ_2 space, which at the same time would retain the same general topographical preferences as the natural amino acid (in this case tyrosine) has been attained. Of particular importance is that clear conformational preferences in χ_1 and χ_2 space should be obtained that manifest themselves in providing a greater selectivity for molecular recognition. Incorporation of the TMT isomers into cyclic enkephalin and deltorphin resulted in analogues with high potency and selectivity [33,34]. A discussion of the specific findings are outside the scope of this paper. However, a few important points can be made with regard to the de novo design of peptides, peptidomimetic ligands, and non-peptide scaffolds which can mimic the bioactive topographies of peptides. First, it is clear that specific changes at one or two chi angles of one critical side-chain group important to molecular recognition at a receptor, can have profound effects on the binding potency, selectivity for receptor types and subtypes, and transduction (agonist versus antagonist) bioactivity. Second, specific chi constraints can convert an agonist analogue to an antagonist analogue. Clearly, considerations of chi space not only are a powerful tool for investigating conformation–activity relationships in bioactive peptides, but will be essential considerations in the de novo design of peptidomimetics. Furthermore, it is likely that a proper design in chi space can lead to peptides and peptidomimetics that are stable against enzymatic degradation, and that can lead to ligands with enhanced ability to cross membrane barriers and enhanced bioavailability [35,36].

Conclusions

The design of scaffolds for the discovery of bioactive molecules using combinatorial chemistry, requires a careful consideration of several stereostructural factors. A key issue is whether the scaffold can properly place the particular groups important for a particular pharmacophore in 3D space. Though much progress has been made [1–3], there is still much to learn, especially when specific details of configuration and conformational similarity become important, and when specific side-chain groups must be correctly located in 3D space. Within this context, it becomes important to consider the chi space. The results of these studies indicate that the design of scaffolds that can explore different aspects of topochemical space, will require a careful consideration of chemical differences, but equally important, different topographical properties in chi space. These aspects of scaffold design have generally been ignored, but it would seem highly beneficial to take them into critical consideration in scaffold design, especially when a high potency, selectivity and stability of bioactive compounds are important goals.

Acknowledgements

The support of the U.S. Public Health Service, National Cancer Institute, CA27502 and the National Institute of Drug Abuse DA06284 is gratefully acknowledged. We thank Magda Stankova and Ronald Ferguson for their collaboration on aspects of this research. We also thank Dr. Guigen Li, Dr. K.C. Russell, Dr. Xinhua Qian, and Dr. Ernesto Nicolás for preparing the specialized amino acids reported here.

References

1 Sawyer, T.K., *Peptidomimetic design and chemical approaches to peptide metabolism*, In Taylor, M.D. and Amidon, G.L. (Eds.) Peptide-Based Drug Design, American Chemical Society, Washington, DC, U.S.A., 1995, pp. 387–422.

2 Giannis, A. and Kolter, T., *Peptidomimetics for receptor ligands – discovery, development and medical perspectives*, Angew. Chem. Int. Ed. Engl., 32 (1993) 1244–1267.

3 Hruby, V.J., Al-Obeidi, F. and Kazmierski, W.M., *Emerging approaches in the molecular design of receptor-selective peptide ligands: Conformational, topographical and dynamic considerations*, Biochem. J., 268 (1990) 249–262.

4 Nicolás, E., Russell, K.C. and Hruby, V.J., *Asymmetric 1,4-addition of organocuprates to chiral α,β-unsaturated N-acyl-4-phenyl-2-oxazolidinones: A new approach to the synthesis of chiral β-branched carboxylic acids*, J. Org. Chem., 58 (1993) 766–770.

5 Li, G., Jarosinski, M.A. and Hruby, V.J., *Diastereospecific tandem Michael-like additionlelectrophilic bromination: A one-pot tandem asymmetric synthesis of precursors of unusual amino acids*, Tetrahedron Lett., 34 (1993) 2561–2564.

6 Qian, X., Russell, K.C., Boteju, L.W. and Hruby, V.J., *Stereoselective total synthesis of topographically constrained designer amino acids: 2',6'-Dimethyl-β-methyltyrosines*, Tetrahedron, 51 (1995) 1033–1054.

7 Lam, K.S., Salmon, S.E., Hersh, E.M., Hruby, V.J., Kazmierski, W.M. and Knapp, R.J., *A new type of synthetic peptide library for identifying ligand-binding activity*, Nature, 354 (1991) 82–84.

8 Lam, K.S., Lebl, M., Wade, S., Stierandova, A., Khattri, P., Collins, N. and Hruby, V.J., *Streptavidin–peptide interaction as a model system for molecular recognition*, In Hodges, R.S. and Smith, J.A. (Eds.) Peptides: Chemistry, Structure and Biology (Proceedings of the 13th American Peptide Symposium), ESCOM, Leiden, The Netherlands, 1994, pp. 1005–1006.

9 Lam, K.S. and Lebl, M., *Streptavidin and Avidin recognize peptide ligands with different motifs*, Immunomethods, 1 (1992) 11–15.

10 Mohamadi, F., Richards, N.G.J., Guida, W.C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T. and Still, W.C., *MacroModel – An integrated software system for modeling organic and bioorganic molecules using molecular mechanics*, J. Comput. Chem., 11 (1990) 440–467.

11 MacroModel, Interactive Molecular Modeling System, v. 4.5, Department of Chemistry, Columbia University, New York, NY, U.S.A., 1994.

12 Lipton, M. and Still, W.C., *The multiple minimum problem in molecular modeling. Tree searching internal coordinate conformational space*, J. Comput. Chem., 9 (1988) 343–355.

13 Allinger, N.L., Yuh, Y.H. and Lii, J.-H., *Molecular Mechanics. The MM3 force field for hydrocarbons 1*, J. Am. Chem. Soc., 111 (1989) 8551–8566.

14 Weiner, S.J., Kollman, P.A., Case, D.A., Chandra Singh, U., Ghio, C., Alagona, G., Profeta, S. and Weiner, P.J., *A new force field for molecular mechanical simulations of nucleic acids and proteins*, J. Am. Chem. Soc., 106 (1984) 765–787.

15 Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A., *An all-atom force field for simulations of proteins and nucleic acids*, J. Comput. Chem., 7 (1986) 230–252.

16 Polak, E. and Ribiere, G., *Rev. Franc. Inf. Rech. Oper.*, 16 (1969) 35; quoted from Ref. 10.

17 Burkert, U. and Allinger, N.L., Molecular Mechanics, ACS Monograph 177, American Chemical Society, Washington, DC, U.S.A., 1982, pp. 67–72.

18 Rykaert, J.-P., Ciccotti, G. and Berendsen, H.J.C., *Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes*, J. Comput. Phys., 23 (1977) 327–341.

19 Kabach, W., *A solution for the best rotation to relate two sets of vectors*, Acta Crystallogr., A32 (1976) 922–923.

20 Balaram, P. and Ramaseshan, S. (Eds.) Molecular Conformation and Biological Interactions, Indian Academy of Science, Bangalore, India, 1991.

21 Hruby, V.J. and Nikiforovich, G.V., *The Ramachandran plot and beyond: Conformational and topographical considerations in the design of peptides and proteins*, In Balaram, P. and Ramasheshan, S. (Eds.) Molecular Conformation and Biological Interactions, Indian Academy of Science, Bangalore, India, 1991, pp. 429–445.

22 Hruby, V.J., *Conformational restrictions of biologically active peptides via amino acid side-chain groups*, Life Sci., 31 (1982) 189–199.

23 Nicolaou, K.C., Solvino, J.M., Raynor, K., Pietranico, S., Reisine, T., Freidinger, R.M. and Hirschmann, R., *Design of synthesis of a peptidomimetic employing β-D-glucose for scaffolding*, J. Am. Chem. Soc., 115 (1993) 12550–12586.

24 Weber, P.C., Ohlendrof, D.H., Wendoloski, J.J. and Salemme, F.R., *Structural origins of high-affinity biotin binding to Streptavidin*, Science, 243 (1989) 85–88.

25 Devlin, J.J., Panganiban, L.C. and Devlin, P.E., *Random peptide libraries: A source of specific protein-binding molecules*, Science, 249 (1990) 404–406.

26 a. Katz, B.A., *Binding of protein targets of peptidic leads discovered by phage display: Crystal structures of Streptavidin-bound linear and cyclic peptide ligands containing the HPQ sequence*, Biochemistry, 34 (1995) 5421–5429.

b. Giebel, L.B., Cass, R.T., Milligan, D.L., Young, D.C., Arze, R. and Johnson, C.R., *Screening of cyclic peptide phage libraries identifies ligands that bind Streptavidin with high affinities*, Biochemistry, 34 (1995) 15430–15435.

27 Weber, P.C., Wendoloski, J.J., Pantoliano, M.W. and Salemme, F.R., *Crystallographic and thermodynamic comparison of natural and synthetic ligands bound to Streptavidin*, J. Am. Chem. Soc., 114 (1992) 3197–3200.

28 Weber, P.C., Pantoliano, M.W., Simons, D.M. and Salemme, F.R., *Structure-based design of synthetic azalienzine ligands for Streptavidin*, J. Am. Chem. Soc., 116 (1994) 2717–2727.

29 Weber, P.C., Pantoliano, M.W. and Thompson, L.D., *Crystal structure and ligand-binding studies of a screened peptide complexed with Streptavidin*, Biochemistry, 31 (1992) 9350–9354.

30 Katz, B.A., Johnson, C.R. and Cass, R.T., *Structure-based design*

of high-affinity Streptavidin-binding cyclic peptide ligands containing thioether cross-links, J. Am. Chem. Soc., 117 (1995) 8541–8547.

31 Jones, M.L. and Kurzban, G.P., *Noncooperativity of biotin binding of tetrameric Streptavidin*, Biochemistry, 34 (1995) 11750–11756.

32 Jiao, D., Russell, K.C. and Hruby, V.J., *Locally constrained tyrosine analogues with restricted side-chain dynamics*, Tetrahedron, 49 (1993) 3511–3520.

33 Qian, X., Kövér, K.E., Shenderovich, M.D., Misicka, A., Zalewska, T., Horvath, R., Davis, P., Porreca, F., Yamamura, H.I. and Hruby, V.J., *Newly discovered stereochemical requirements in side-chain conformation of δ-opioid agonists for recognizing opioid δ-receptors*, J. Med. Chem., 37 (1994) 1746–1757.

34 Qian, X., Shenderovich, M.D., Kövér, K.E., Bilsky, E.J., Horváth, R., Davis, P., Yamamura, H.I., Porreca, F. and Hruby, V.J., *Probing the stereochemical requirements for recognizing opioid δ-receptors through topographical design of the message domain of δ-opioid agonists*, J. Am. Chem. Soc., 118 (1996) 7280–7290.

35 Hruby, V.J., Davis, T.P., Polt, R., Bartosz-Bechowski, H., Misicka, A., Lipkowski, A., Sharma, S.D., Li, G., Bonner, G., Meyer, J.-P., Patel, D., Yamamura, H.I., Porreca, F. and O'Brien, D.F., *A systematic investigation of factors that enhance penetration of peptides across the blood–brain barrier*, In Kaumaya, P.T.P. and Hodges, R.S. (Eds.) Peptides: Chemistry, Structure and Biology (Proceedings of the 14th American Peptide Symposium), Mayflower Scientific Ltd., Kingswinford, U.K., 1996, pp. 154–156.

36 Hruby, V.J., Davis, T.P., Polt, R., Porreca, F., O'Brien, D., Yamamura, H.I., Bartosz, H., Szabo, L., Gillespie, T.J., Misicka, A., Lipkowski, A.W., Qian, X., Li, G., Patel, D. and Bonner, G., *Design and synthesis of peptide ligands with unique biochemical and biological profiles at opioid receptors that cross the blood–brain barrier*, Analgesia, 1 (1995) 469–472.