# Sequence-Selective Evaluation of Peptide Side-Chain Interaction. New Artificial Receptors for Selective Recognition in Water

Md. Alamgir Hossain and Hans-Jörg Schneider*

*FR Organische Chemie der Universität des Saarlandes*
*D 66041 Saarbrücken, Germany*

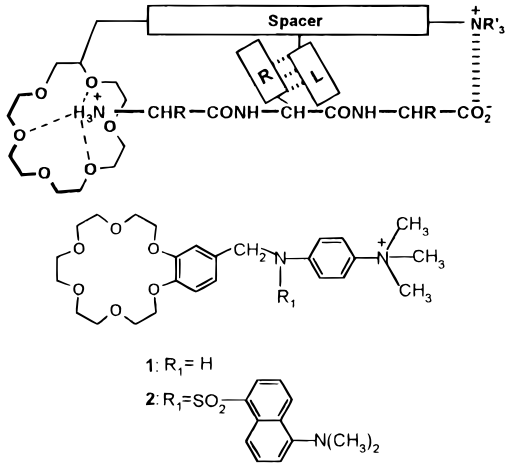
*Received July 10, 1998*

Reading and transmission of information in peptides relies exclusively on interactions with the different amino acid side chains, in contrast to the mechanism in nucleic acids where hydrogen bonds provide the corresponding codes. Recognition of amino acid sequences, in particular by synthetic receptors, is furthermore hampered by the lack of conformational order, again in contrast to that in the conformationally more stable nucleic acids. A truly biomimetic peptide recognition system should provide for the directional alignment of a peptide chain, with differentiation of the N and C terminus and between different side chains, allowing, for instance, the differentiation of a Phe-Gly from a Gly-Phe sequence. We describe here a strategy which also allows the evaluation in water of the energetic contributions of side-chain substituents, which should help our understanding of protein folding and of ligand binding mechanisms. It may also provide a basis for the development of artificial peptidases and of enzyme inhibitors.

Only a few synthetic receptors for peptide recognition in water have been described until today,[1] which either rely on binding of lipophilic side chains[1b] or on polar interactions with acidic amino acid side chains.[1c] The sequence-selective complexation in water described by Still et al. occurs on polymers, on which dyes[2] or lipophilic macrocycles are immobilized for peptide recognition.[3] A larger variety of mostly macrocyclic host compounds for peptides work only in hydrophobic medium.[4] The associations with these systems rely essentially on hydrogen bonds and are restricted to chloroform as solvent. Furthermore, almost all known synthetic peptide recognition systems have been described for protected peptides only. We were looking for host compounds which can recognize any natural peptide in water with length and sequence selectivity and align the peptide in space while we simultaneously measure interactions with the amino acid side chains.

The textbook picture of peptide $\beta$-sheets would suggest that one could design receptors on the basis of complementary hydrogen bonds between the amino acids and use the terminal carboxylic or amino ends as salt-bridge elements. However, it is well-known that the construction even of **intra**molecular turns showing such complementary hydrogen bonds usually requires special anchors.[5] Oligopeptides, including also those with phenylalanines, show no evidence for any association in water, likely because of insufficient match between the ionic terminal sites or the aromatic residues. With longer or more lipophilic peptides, which would allow stronger interactions between their side chains, water-insoluble material is observed even at very low concentrations.

**Scheme 1.** Design and Structure of Water-Soluble Peptide Receptors




An ideal synthetic receptor for peptides should exert primary binding at the terminal ionic groups, with secondary interactions along the peptide chain providing for amino acid and sequence selectivity. A zwitterionic host with charges at sites complementary to the peptide analyte would serve this purpose but has the severe disadvantage of competing self-association. On the basis of these considerations, we designed host compounds with one binding site A for the peptide N-terminus, and another site B for the C-terminus. The structure should ensure that A and B will not interact with each other in order to avoid self-association. The choice was a 18-crown-6 ether for site A, which is well-known to complex protonated nitrogen groups; for site B a peralkylammonium group was introduced, which has been checked for the absence of any measurable interactions with crown ethers.[6]

A suitable ligand **1** (Scheme 1) containing terminal crown ether and ammonium units was prepared in an overall yield of 45% by reaction of 4-formyl-benzo-18-crown-6[7] with 4-(dimethylamino)-aniline, reduction of the imine with sodium borohydride, and selective permethylation at one of the anilino groups. The new host compound is well soluble in water and to a lesser but sufficient degree also in methanol. For extraction and transport experiments it easily can be made lipophilic by the introduction of longer alkyl instead of methyl groups at the terminal nitrogen atom. Computer-aided molecular modeling using energy mini-

(1) (a) Wennemers, H.; Still, W. C. *Tetredron Lett.* **1994**, *35*, 6413. (b) Breslow, R.; Yang, Z.; Ching, R.; Tzrojandt, G.; Odobel, F. *J. Am. Chem. Soc.* **1998**, *120*, 3536. (c) Peczuh, M. W.; Hamilton, A. D.; Sanchez-Qesada, J.; de Mendoza, J.; Haack, T.; Giralt, E. *J. Am. Chem. Soc.* **1997**, *119*, 9327. (d) Yoon, S. S. *Bull. Korean Chem. Soc.* **1998**, *19*, 252, 254.

(2) (a) Chen, C.-T.; Wagner, H.; Still, W. C. *Science* **1998,** *279*, 851. (b) Hioki, H,; Still, W. C. *J. Org. Chem.* **1998**, *63*, 904. (c) Yoon, S. S.; Still, W. C. *Tetrahedron* **1995**, *51*, 567. (d) Wennemers, H.; Yoon, S. S.; Still, W. C. *J. Org. Chem.* **1995**, *60*, 1108. (e) Yoon, S. S.; Still, W.C. *Tetrahedron Lett.* **1994**, *35*, 8557.

(3) Recently, association constants of up to $10^7$ $M^{-1}$ in water were claimed for complexes between a nonapeptide and some pentapeptides (Sasaki, S.; Takagi, M.; Tanaka, Y.; Maeda, M. *Tetrahedron Lett.* **1996**, *37*, 85). However, the fitted UV curves described in this communication, measured at rather short wavelength, show an unusual behavior. In our hands attempted UV measurements of oligopeptide associations in water were hampered by nonlinear effects of small stray light even with high quality spectrometers.

(4) (a) For a short review, see: Schneider, H.-J. *Angew. Chem.* **1993**, *105*, 890; *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 848. (b) Liu, Ruiping; Still, W. C. *Tetrahdron Lett.* **1993**, *34*, 2573. (c) Dowden, J.; Edwards, P. D.; Kilburn, J. D. *Tetrahedron Lett.* **1997**, *38*, 1095. (d) Waymark, C. P.; Kilburn, J. D.; Gillies, I. *Tetrahedron Lett.* **1995**, *36*, 3051. (e) Casnati, A.; Fabbi, M.; Pelizzi, N.; Pochini, A.; Sansone, F.; Ungaro, R. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2699.

(5) For leading references, see: (a) Nowick, J. S.; Smith, E. M.; Pairish, M. *Acc. Chem. Res.* **1996**, *29*, 1. (b) Schneider, J. P.; Kelly, J. W. *Chem. Rev.* **1995**, *95* 2169. (c) Kemp, D. S.; Li, Z. Q. *Tetrahedron Lett.* **1995**, *36*, 4175, 4179. (d) Gardner, R. R.; Liang, G.-B.; Gellman, S. H. *J. Am. Chem. Soc.* **1995**, *117*, 7, 3280. (e) Brandmeier, V.; Sauer, W. H. B.; Feigel, M. *Helv. Chim. Acta* **1994**, *77*, 70. (f) Gretchikhine, A. B.; Ogawa, M. Y. *J. Am. Chem. Soc.* **1996**, *118*, 1543 (some results in this work are questionable, see ref 5a); (g) Doig, A. J. *J. Chem. Soc., Chem. Commun.* **1997**, 2153.

(6) Rüdiger, V.; Schneider, H.-J.; Solov'ev, V. P.; Kazachenko, V. P.; Raevsky, O. A. *Eur. J. Org. Chem* **1998**, accepted.

(7) Reinhoudt, D. N.; De Jong, F.; Tomassen, H. P. M. *Tetrahedron Lett.* **1979**, *22*, 2067.

10.1021/ja982435g CCC: $15.00 
Published on Web 10/20/1998

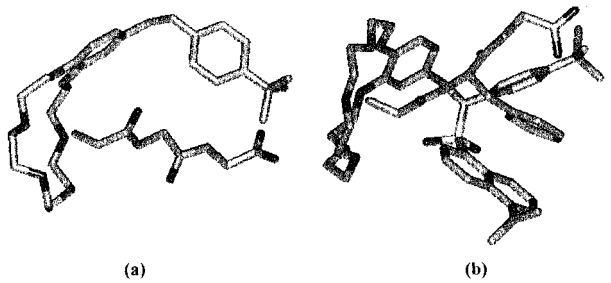



**Figure 1.** Force-field optimized stuctures of the complexes of receptor **1a** with triglycin (a), and of receptor **1b** with Gly-Phe-Gly (b). Hydrogen atoms are omitted for clarity. In each case the N-terminus of the peptide is anchored within the crown moiety by hydrogen bonding; the corresponding C terminus forms an ion pair permethylated ammonium group. The additional aromatic−aromatic interactions of the side chains is observed in the case of the complex shown in (b).

**Table 1.** Binding Constants of **1** with Peptides[a]

| entry | peptide | solvent[b] | $K$/M$^{-1}$[c] | $-\Delta G$[d] |
|---|---|---|---|---|
| 1 | Gly-Gly-Gly | W | 200 | 13.1 |
| 2 | Gly-Gly-Phe | W | 170 | 12.7 |
| 3 | Gly-Phe-Gly | W | 180 | 12.9 |
| 4 | Phe-Gly-Gly | W | 135 | 12.1 |
| 5 | Leu-Leu-Leu | W | 155 | 12.5 |
| 6 | Gly-Gly | W | 50 | 9.5 |
| 7 | Gly-Leu | W | 45 | 9.4 |
| 8 | Leu-Gly | W | 40 | 9.1 |
| 9 | Gly-Phe | W | 40 | 9.1 |
| 10 | Gly-Gly-Gly-Gly | W | 45 | 9.4 |
| 11 | Gly-Gly-GlyOMeH$^+$ | W | 30 | 8.2 |
| 12 | Gly-Gly-Gly | M | 13 000 | 23.5 |
| 13 | Gly-Gly-Phe | M | 8 500 | 22.4 |
| 14 | Gly-Phe-Gly | M | 10 500 | 23.0 |

[a] Measured by NMR titration of **1** with peptides at 25 °C; $[\text{peptide}]_0 = 10$ mM (in $D_2O$), or 1.0 mM (in $CD_3OD$); titrations were performed with eight measurements under neutral conditions (pH = 7.0 ± 0.2), error limit in $K < \pm 5\%$, in $\Delta G < \pm 0.5$ [kJ mol$^{-1}$]. [b] W, water; M, methanol. [c] Values were obtained from the average of two to four different NMR signals of peptides. [d] $\Delta G$ in [kJ mole$^{-1}$].

mization with the CHARMm force field[8] established with this host and a tripeptide an essentially strain-free fit between crown and $^+NH_3$ at one end, and between the $^+NMe_3$ and the $COO^-$ groups at the other end, with retention of the most stable extended conformation of the peptide (Figure 1).

NMR titrations with several peptides were performed as described earlier;[9] nonlinear least-squares fitting of the observed shift changes with several independent NMR signals showed satisfactory fitting with a 1:1 association model. The calculated constants (Table 1) are even in water above $K = 100$ M$^{-1}$, wherever geometric fit is to be expected. The $K$ variations show a moderate selectivity for the amino acid nature and for the sequence (cf. entries 1 and 4 in Table 1) and demonstrate the desired *length*-selectivity (cf. entries 1 and 10 in Table 1). The low association constant observed in the case of the triglycylester (entry 11) confirms the bidentate binding mechanism. In methanol one observes much higher binding constants, around $10^4$ M$^{-1}$, and significantly larger selectivities (Table 1). This is the result of tighter binding at both ends of the peptides and should lead to much stronger selectivity in a less polar environment.

Introduction of a dansyl substituent into host **1** yielded receptor **2** which not only provides for the convenient analysis by fluorescence-instead of NMR-spectroscopy but also allows for the first time the evaluation of lipophilic interactions of specific amino acid side chains. Figure 2 shows that the optical titrations also lead to good fit in a 1:1 association model and characterizes

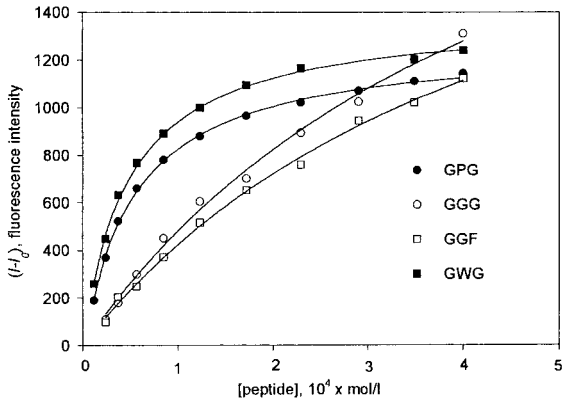



**Figure 2.** Fluorimetric titration curves of receptor **2** with tripeptides (GFG = Gly-Phe-Gly, GGG = Gly-Gly-Gly), (GGF = Gly-Gly-Phe), (GWG = Gly-Trp-Gly). Excitaion wavelength, $\lambda_{max} = 320$ nm.

**Table 2.** Binding Constants of **2** with Tripeptides in Water

| entry | guest | $K$/M$^{-1}$[a] | $-\Delta G$[d] | $I_{max}/I_0$,[b] |
|---|---|---|---|---|
| 1 | Gly-Gly-Gly | 210 | 13.2 | 3.5 |
| 2 | Gly-Gly-Gly | 190 | 13.0 | *c* |
| 3 | Gly-Gly-Phe | 215 | 13.3 | 3.2 |
| 4 | Phe-Gly-Gly | 220 | 13.6 | 2.8 |
| 5 | Gly-Phe-Gly | 1700 | 18.4 | 1.8 |
| 6 | Gly-Phe-Gly | 1620[c] | 18.3 | *c* |
| 7 | Trp-Gly-Gly | 260 | 13.8 | 3.0 |
| 8 | Gly-Gly-Trp | 310 | 14.3 | 2.8 |
| 9 | Gly-Trp-Gly | 2150 | 19.0 | 2.3 |
| 10 | Gly-Leu-Gly | 850 | 16.7 | 3.4 |
| 11 | Leu-Leu-Leu | 700 | 16.3 | 3.3 |
| 12 | Gly-Ala-Gly | 540 | 15.6 | 3.1 |
| 13 | Val-Val-Gly | 720 | 16.3 | 3.3 |

[a] Measured by fluorescent titration of peptides with **2** at 25 °C; $[\mathbf{2}]_0 = 10^{-5}$ M, $\lambda_{max} = 320$ nm; peptide concentration from $4 \times 10^{-3}$ M to $2.5 \times 10^{-4}$ M. Titrations with usually 10 to 12 measurements; pH = 7.0 ± 0.2, error limit in $K < \pm 5\%$, in $\Delta G < \pm 0.5$ [kJ mol$^{-1}$]. [b] $I_0$ is initial fluorescence intensity, $I_{max}$ is maximum (intrinsic) intensity at 100% complexation calculated from nonlinear least-squares fit. [c] Measured by NMR titration. [d] $\Delta G$ in [kJ mol$^{-1}$].

the obvious fluorometric differences between the different peptides. The data in Table 2 demonstrate that these interactions enhance the peptide associations by a power of magnitude with a phenyl or indole side substituent. Figure 1b illustrates the possible stacking between the aromatic parts of host and guest. An aliphatic side chain such as the one in valine also leads to an affinity increase compared to glycine, but to a lesser degree than aromatic substituents. The complex with a specific amino acid and its position in the peptide is characterized not only by differences in the association constants but also by changes in the intrinsic fluorescence emission intensities $I_{max}/I_0$ of the complexes (Table 2), and to a smaller degree even in changes in the emission maxima wavelength; this can be of interest for the development of corresponding sensors.

The results hold promise for the development of ligands, which in ***any*** environment can differentiate peptides without prior modification according to length, amino acid composition, sequence, and, after introduction of chiral elements, configuration. The synthesis of suitable receptors is straightforward and in progress. More lipophilic derivatives can show considerable biological activity at lower concentrations.[10]

**Acknowledgment.** M.A.H. thanks the Alexander von Humboldt foundation for a postdoctoral fellowship. Our work is supported by the Deutsche Forschungsgemeinschaft, Bonn, and the Fonds der Chemischen Industrie, Frankfurt.

JA982435G

(8) Brünger, A. T.; Karplus, M. *Acc. Chem. Res.* **1991**, *24*, 54; Brooks, C. L.; Karplus, M. *Methods Enzymol.* **1986**, *127*, 369.

(9) Schneider, H.-J.; Kramer, R.; Simova, S.; Schneider, U. *J. Am. Chem. Soc.* **1988**, *110*, 6442.

(10) This is paper 82 of our series Supramolecular Chemistry. Paper 81: Ali, A.; Gasiorek, M.; Schneider, H.-J. *Angew. Chem.* In press.