PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Synthetic Receptors for Peptides

W. Clark Still and Reiping Liu

Phil. Trans. R. Soc. Lond. A 1993 **345**, 97-104
doi: 10.1098/rsta.1993.0121

Synthetic receptors for peptides

By W. Clark Still and Reiping Liu

Department of Chemistry, Columbia University, New York, New York 10027, U.S.A.

A part of a project directed toward predicting the binding properties of receptors for molecules of biological interest, we have synthesized and studied several synthetic receptors for simple peptides. We describe here the enantioselective binding properties of one of these molecules and structural studies of its complexes.

1. Introduction

For several years now, members of our research group have been designing, synthesizing and studying new substrate-selective receptors for peptides. The aim of our research is to learn enough about bimolecular complexation of biological molecules to be able to design new molecules which interact in a predictable way with given molecular structures or receptors. Such a capability would put molecular design on a more rational basis than is currently available.

Ideally, we would like to predict association constants or free energies. However, molecular association is an intrinsically complex phenomenon because many thermodynamic properties of the binding partners change significantly during association. To calculate association constants with acceptable accuracy, we would need good estimates of the changes in entropy, solvation energy and interaction energy which occur upon binding. Unfortunately, the changes in such thermodynamic properties can be large and their estimates are often difficult to make with sufficient accuracy to predict actual binding energies within 1 kcal mol^{-1} (but see Jorgensen *et al.* 1991, 1992). Without this level of accuracy, calculations would not make a very useful design tool. In contrast to the problematic calculation of absolute binding energies, predicting the relative binding of two closely related substrates for the same receptor is a much easier problem because many calculational errors are systematic and cancel effectively.

Another problem in predicting molecular association involves the flexibility of the molecules undergoing association and their complex. To estimate the overall energy of any real molecular system, it is necessary to include appropriately weighted contributions from all significantly populated states (e.g. conformations) of the system. Including such contributions can be a very difficult problem, especially where there are a large number of states or when important states are separated by large energy barriers. Thus highly flexible receptor/substrate systems can present major problems for computational studies (Mitchell & McCammon 1991; van Gunsteren & Mark 1992). For that matter, even the interpretation of experimental studies are severely limited by flexibility when the contributing three-dimensional structures are not known with certainty.

The experiments we describe here involve the synthesis and study of macrocyclic molecular receptors for peptides. By comparing the results of these experiments with

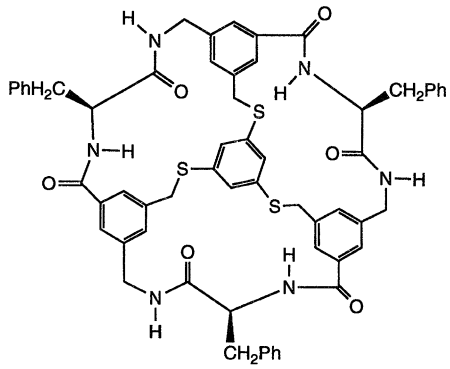
Phil. Trans. R. Soc. Lond. A (1993) **345**, 97–104
Printed in Great Britain

the corresponding calculations, we hope to validate the computational methods we wish to use as predictive tools for molecular design. To have the best chance of calculating relative binding energies accurately, we have tried to keep our recognition systems as simple as possible. Thus we have chosen to study enantioselective binding because it measures the relative binding of the most closely related substrates imaginable, i.e. mirror image isomers. Such selectivity should be minimally sensitive to errors in a molecular mechanics force field (Sanderson *et al.* 1989; Dappen *et al.* 1990). We have also tried to design our receptors to be conformationally inflexible so that the three-dimensional structure of at least one of the binding partners will be known. This choice of conformationally stable receptors should minimize errors associated with insufficient sampling of the various populated states of the system. Finally, we are primarily interested in receptors for peptides which display binding enantioselectivity of more than 0.5 kcal mol^{-1}. These last requirements reflect our focus on biologically relevant systems and our belief that any contemporary calculation is unlikely to be substantially more accurate than this.

Given the constraints outlined above, it is rather challenging to find enantioselective receptors which are appropriate for the studies we have planned. Though a number of enantioselective host molecules for neutral substrates have been reported in recent years, only a few exhibit binding enantioselectivities much above 0.5 kcal mol^{-1} and fewer still are conformationally well defined (Jeong *et al.* 1990; Liu *et al.* 1990; Murakami *et al.* 1991). Consequently we set out to design and synthesize new receptors which were suitable for our studies. In this paper, we describe our work on one class of macrocyclic receptors having the requisite properties including highly enantioselective binding (90–99% ee) of simple peptides.

2. Macrotricyclic receptors for neutral peptides

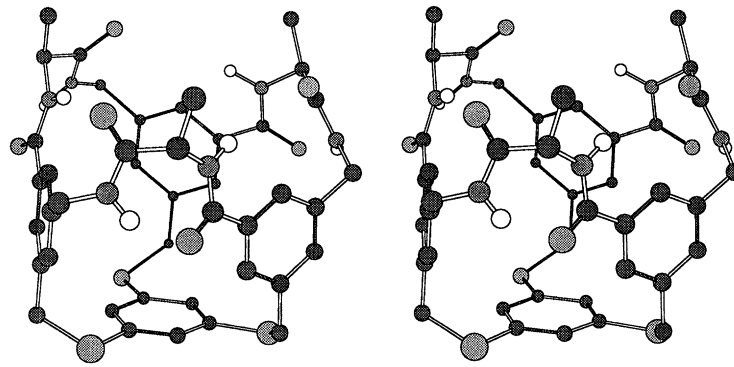
One of the most interesting receptors for neutral peptides we have prepared is the C_3-symmetric macrotricyclic host **1** (Hong *et al.* 1991). This structure was assembled from two different types of *meta*-trisubstituted aromatics which provided conformational rigidity and an L-phenylalanine linker which induced a well-defined chirality. The synthesis of **1** was efficient and proceeded via a trilactamization which closed all three macrocyclic rings simultaneously in 68% yield (Erickson *et al.* 1993).



1

According to conformational searching (Goodman & Still 1991) by molecular mechanics using the MacroModel implementation of AMBER (McDonald & Still 1992), **1** (modelled as the alanine analogue) exists almost exclusively in a single

family of closely related conformations. All members have the same large, cylindrical binding site whose walls are formed by the three unsymmetrical aromatic rings and the chirality-defining (phenyl)alanine connectors. All have γ-turn conformations for the connecting amino acids as shown in the stereopair diagram below. The various distinct conformations within the family differ only in the Ar–S–CH_2–Ar′ torsion angles which orient the central aromatic ring with respect to the rest of the receptor. This ring forms a non-polar floor which caps the otherwise cylindrical receptor, and its different orientations do not appear to alter the shape or the positions of the hydrogen bonding functionalities of the receptor. In this regard, receptor **1** is predicted to be conformationally heterogeneous, but it may function as if it were homogeneous since all of its low energy conformations are closely related and likely to have similar binding properties.



We evaluated the binding properties of receptor **1** by carrying out titrations (0.5 mM **1** in $CDCl_3$, 23 °C) with a variety of simple amino acid derivatives using ^{1}H NMR to monitor association. In nearly all cases, we were able to measure the movement of at least two protons (typically N–H and C(ar)–H) during the titrations and thus obtain multiple independent evaluations of binding free energies. A summary of the binding energies we measured (Hong *et al.* 1991; Liu & Still 1993) is listed in table 1 and the association free energies reported there are accurate to ± 0.2 kcal mol^{-1}.

As the table indicates, a number of the amino acid derivatives bind to **1** with high enantioselectivity. There are actually three classes of simple peptides which bind with enantioselectivity exceeding 1 kcal mol^{-1}. There are N-Boc amino acid N-methyl amides (entries 1, 4–7), N-Moc (methoxycarbonyl) amino acid t-butyl esters (entries 9–12), and certain sidechain-functionalized amino acid esters (entries 16, 18–20). Although we have been unable to obtain crystals of **1** or its complexes of X-ray crystallography, we have learned a great deal about the structure of **1**/peptide complexes from NMR and molecular mechanics.

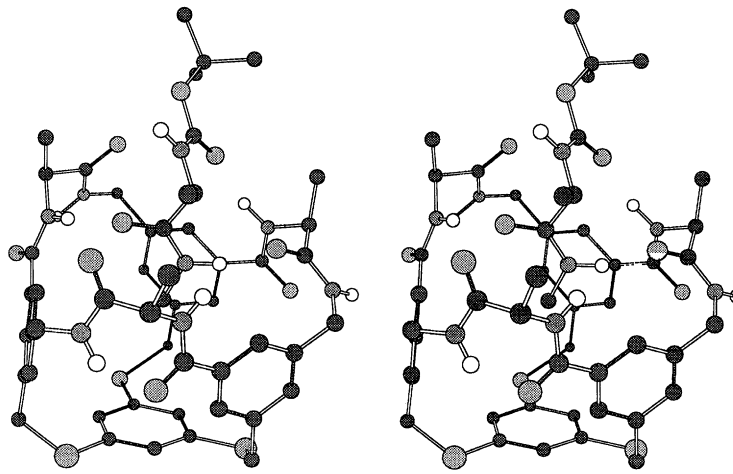
Among the various substrates showing highly enantioselective binding, we have studied complexes of the Boc, N-methyl amides most thoroughly. To begin with, we carried out a molecular mechanics conformational search of the **1**/Boc–L–Ala–NHMe complex. This search involved not only rotating the torsion angles of the host and its peptidic guest, but also varying the position (x, y, z translations ± 1 Å) and orientation (x, y, z rotations $\pm 180°$) of the guest relative to **1**. Interestingly, the search revealed that while there several forms of the complex within the first 2 kcal mol^{-1} of the ground state, all had the C-terminal N-methyl group of the peptide guest deeply buried within the binding cavity of the host as shown below in

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Table 1. *Binding of* **1** *and α-amino derivatives in* $CDCl_3$

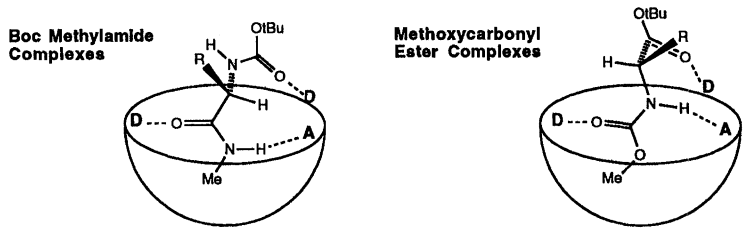
entry	substrate	binding energy, $-\Delta G$ / kcal mol^{-1} L	binding energy, $-\Delta G$ / kcal mol^{-1} D	enantioselectivity
1	Boc–Ala–NHMe	3.9	1.7	95% ee
2	Boc–Ala–NHBn	1.4	—	
3	Ac–Ala–NHMe	3.9	2.7	76% ee
4	Boc–Val–NHMe	4.4	1.5	99% ee
5	Boc–Leu–NHMe	4.1	1.5	98% ee
6	Boc–Ser–NHMe	> 6.1	3.8	> 96% ee
7	Boc–Thr–NHMe	> 6.2	3.2	> 98% ee
8	Boc–Ala–Ala–NHMe	LL: 2.9 DD: 2.0	LD: 2.8 DL: 3.3	64, 40% ee
9	MeO_2C–Ala–OtBu	4.8	2.3	97% ee
10	MeO_2C–Val–OtBu	3.7	1.5	95% ee
11	MeO_2C–Ser–OtBu	⩾ 7.0	4.7	⩾ 96% ee
12	MeO_2C–Ala–Ala–OtBu	LL: 4.7 DD: 2.2	LD: 4.4 DL: 2.7	97, 89% ee
13	Pr–Ala–OtBu	3.8	2.3	84% ee
14	Ac–Ala–OtBu	3.0	1.5	84% ee
15	Boc–Ala–OMe	1.5	1.2	15% ee
16	Boc–Ser–OMe	4.7	2.9	90% ee
17	Boc–His–OMe	3.5	2.7	58% ee
18	Boc–Asn–OMe	3.1	2.2	64% ee
19	Boc–Gln–OMe	4.2	2.2	93% ee
20	Boc–Glu(OMe)–OMe	2.9	1.4	84% ee

stereo. Such a geometry is compatible with experimental observation that increasing the size of the C-terminal amide substitute (e.g. from methyl to benzyl, entries 1 and 2) decreases binding significantly. It also suggests how hydrogen bond donating functionality on a bound peptide sidechain (e.g. the OH of serine or threonine) may associate with a nearby host carbonyl to increase binding (entries 1 versus 6, 7).



The structure found by conformational searching is also compatible with the ^{1}H NMR observation of a large upfield shift (from 2.8 to −0.8 p.p.m.) for the peptidic N-methyl resonance upon binding. ROESY and difference nOe studies on the

complex provided further confirmation of the proposed structure by showing, *inter alia*, intermolecular nOe signals indicating proximity of the bound guest's N-methyl and protons on the trimercaptobenzene ring in the bottom of the binding site.



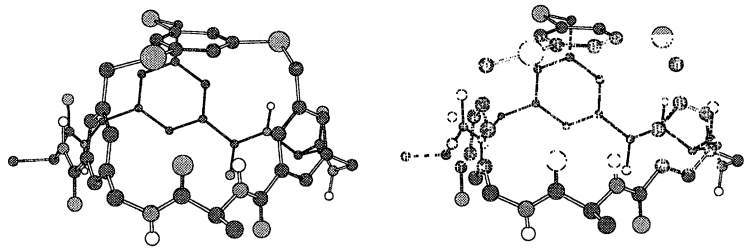
Thus our calculational and experimental results suggest a binding motif for the Boc methylamide derivatives of α-amino acids in which the lowest energy forms encapsulate the C-terminal carboxamide substituent within the binding cavity. According to molecular mechanics, complexation is driven energetically by the formation of three intermolecular hydrogen bonds. Such a binding mode is sketched above and suggested to us that a different set of peptide protecting groups might reverse the orientation of the peptide guest within the binding site. To make the orientation-reversing protecting groups closely resemble the previously studied Boc, methylamide derivatives, we chose an N-terminal methoxycarbonyl (Moc) and a C-terminal *tert*-butyl ester. The expected binding mode is shown above and was found as the global minimum by conformational searching.

As table 1 indicates, receptor **1** does indeed bind Moc, *tert*-butyl esters (entries 9–12) tightly and with high enantioselectivity. The binding site of **1** seems particularly well-suited to fit a methoxycarbonyl N-terminal substituent as small changes in its structure (e.g. replacing CH_3O by CH_3CH_2; entries 9 and 13) significantly reduce the binding of the preferentially bound L-peptide. Though structural evidence for the suggested binding mode is limited at this time, inverse titration experiments showed that the terminal Moc methyl of serine undergoes a 3.5 p.p.m. upfield shift (from 3.7 to 0.2 p.p.m.) in the ^{1}H NMR upon complexation with **1**. This large upfield shift appears analogous to a similar one observed with the Boc methylamide complexes and suggests that binding is accompanied with insertion of the Moc group into the benzene-lined binding cavity. Intermolecular ROESY and difference nOe experiments provided support to the proposed structure of the complex by showing proximity between the Moc methyl group and the trimercaptobenzene ring. We also observed nOe signals indicating close contacts between the guest *tert*-butyl and host phenylalanine sidechains in both the Boc, methylamide and the Moc, *tert*-butyl ester complexes of **1**. Thus, all available evidence points to the binding modes and structures proposed above.

Our results thus far have been interpreted in terms of a three-dimensional structure for **1** which comes from molecular mechanics conformational searching. Actually, there is a second family of conformations which we calculate by molecular mechanics to be 3.9 kcal mol^{-1} higher in energy and which appears to be separated from **1** by a large energy barrier. In the following discussion, we will designate this second family of conformers as receptor **2**. The two families arise from differing macrocyclization pathways and can be considered to be formed by ring closures over differing faces of the central trimercaptobenzene ring. Thus if family **1** results from cyclization to the top of this ring (as shown in the structural diagrams above), then family **2** originates from cyclization around the bottom. One of **2**s low energy

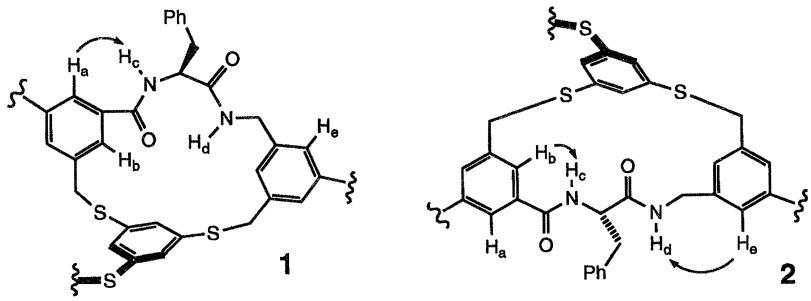
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

conformations is shown below and was found by simulated annealing (300 K → 50 K):



Compared with the previously described **1**, receptor **2** has a shallower binding cavity (depth 5.3 Å against 7.2 Å) and more conformational mobility according to stochastic dynamics calculations. Thus simulations of **2** at 300 K in $CHCl_3$ find a dynamic equilibrium between various γ-turn (C7) and extended (C5) conformations of the (phenyl)alanine connecting fragments, whereas the corresponding fragments of **1** appear almost exclusively in γ-turn conformations.

Though our synthesis of **1** was selective for the receptor whose properties are described above, examination of the crude macrotricyclization product mixture revealed the presence of a minor isomer which we isolated by chromatography in *ca.* 5% yield. We have assigned the structure **2** to this isomer. To confirm the structures for **1** and **2** which had been suggested by molecular mechanics, we carried out a series of ROESY and difference nOe experiments as a probe of the phenylalanine fragment conformation. The most diagnostic nOes we found are indicated in the diagrams below.



These studies indicate that the benzoic carboxamide H_c lies close to H_a in **1** but close to H_b in **2**, and that H_d is proximate to H_e in **2** but not in **1**. Furthermore, H_c in **1** undergoes a large (−1.1 p.p.m. against H_d −0.6 p.p.m.) downfield shift upon binding as would be expected if it were involved in hydrogen bonding with the guest peptide. With **2**, however, it is H_d which shifts downfield (−1.2 p.p.m. against H_c −0.4 p.p.m.) upon peptide binding. Such shifts would be expected if **1** or **2** bound peptides by forming hydrogen bonds with the most exposed (and least hydrogen bonded) amide groups of the host. These studies thus provide further support to the structural proposals outlined above.

Though **1** and **2** are structurally similar, their binding properties in $CDCl_3$ are very different as the binding data for **2** indicates in table 2.

While the D isomers we studied are weakly bound by both **2** and **1**, the L isomers are bound 2–3 kcal mol^{-1} more weakly by **2** than by **1**. Thus we observe essentially no enantioselective binding with **2**. Elucidating the structural origin of this reduced

Table 2. *Binding of* **2** *and* α*-amino acid derivatives in* $CDCl_3$

entry	substrate	binding energy, $-\Delta G$ / kcalmol^{-1} L	binding energy, $-\Delta G$ / kcalmol^{-1} D	enantioselectivity
1	Boc–Ala–NHMe	1.7	1.7	0% ee
2	Ac–Ala–NHtBu	2.2	2.0	16% ee
3	MeO_2C–Ala–OtBu	1.4	1.2	16% ee
4	MeO_2C–Ala–Ala–OtBu	LL: 1.8	DL: 1.6	—

selectivity will require further studies, but it is likely that **2**s shallower binding cavity and increased flexibility will be found to be responsible.

3. Conclusion

As the above studies indicate, we have not only discovered a novel, synthetic receptor for peptides, but we have also collected a great deal of data about its structure and binding properties. Because **1** has a well-defined three dimensional structure and exhibits highly enantioselective binding with several different classes of peptides whereas the closely related **2** has almost no enantioselectivity, these results should provide challenging tests for the computational methods which we hope to use as predictive tools. We will describe computational studies of **1** and **2** in the near future.

We acknowledge the National Science Foundation Grant CHE92 08245 for supporting the work described here.

References

Dappen, R., Karfunkel, H. J. & Leusen, F. J. J. 1990 Computational chemistry applied to design of chiral stationary phases for enantiomeric separation. *J. comput. Chem.* **11**, 181–193.

Erickson, S. D., Simon, J. A. & Still, W. C. 1993 Practical synthesis of a highly enantioselective receptor for peptides. *J. org. Chem.* **58**, 1305–1308.

Goodman, J. M. & Still, W. C. 1991 An unbounded systematic search of conformational space. *J. comput. Chem.* **12**, 1110–1117.

Hong, J.-I., Namgoong, S. K., Bernardi, A. & Still, W. C. 1991 Highly selective binding of simple peptides by a C_3 macrotricyclic receptor. *J. Am. chem. Soc.* **113**, 5111–5112.

Jeong, K.-S., Muehldorf, A. V. & Rebek, J. 1990 Molecular recognition: asymmetric complexation of diketopiperazines. *J. Am. chem. Soc.* **112**, 6144–6145.

Jorgensen, W. L. 1991 Computational insights on intermolecular interactions and binding in solution. *Chemtracts Org. Chem.* **4**, 91–119.

Jorgensen, W. L., Nguyen, T. B., Sanford, E. M. *et al.* 1992 Enhanced view of structure and binding for cyclophane–arene complexes through joint theoretical and experimental study. *J. Am. chem. Soc.* **114**, 4003–4004.

Liu, R., Sanderson, P. E. J. & Still, W. C. 1990 Enantioselective complexation of the alanine dipeptide by a C_2 host molecule. *J. org. Chem.* **55**, 5184–5186.

Liu, R. & Still, W. C. 1993 Highly selective binding of diverse neutral donor/acceptor substrates by a C_3 macrotricyclic receptor. *Tetrahedron Lett.* **34**, 2573–2576.

Murakami, Y., Ohno, T., Hayashida, O. & Hisaeda, Y. 1991 Novel cage-type azaparacyclophane bearing chiral binding sites. *J. chem. Soc. Chem. Commun.* 950–952.

McDonald, Q. M. & Still, W. C. 1992 AMBER* torsional parameters for the peptide backbone. *Tetrahedron Lett.* **33**, 7743–7746.

Mitchell, M. J. & McCammon, J. A. 1991 Free energy difference calculations by thermodynamic integration: difficulties in obtaining a precise value. *J. comput. Chem.* **12**, 271–275.

Sanderson, P. E. J., Kilburn, J. D. & Still, W. C. 1989 Enantioselective complexation of simple amides by a C_2 host molecule. *J. Am. chem. Soc.* **111**, 8314–8315.

van Gunsteren, W. F. & Mark, A. E. 1992 On the interpretation of biochemical data by molecular dynamics computer simulation. *Eur. J. Biochem.* **204**, 947–961.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES