Engineering of a synthetic receptor to alter peptide binding selectivity

Michael R Carrasco and W Clark Still*

Department of Chemistry, Columbia University, New York, NY 10027, USA

Background: Molecular recognition processes are ubiquitous in nature: substrate binding by enzymes, antigen recognition by antibodies, and hormone activation of receptors provide three classic examples. To better understand these large systems it is valuable to study smaller, well defined host molecules. Previously we found sequence-selective peptide binding with a class of C_3 symmetric synthetic receptors. In this work we rationally altered that host structure in order to produce a corresponding change in binding selectivity.

Results: A novel C_3 symmetric receptor was designed and synthesized such that, unlike previous host molecules, it contained hydrogen-bond accepting functionality within the binding cavity. Screening of this host against a combinatorial tripeptide library revealed an exquisite selectivity for sequences with D-Pro-D-Asn carboxyl termini. Computer simulations and NMR studies indicate that hydrogen bonding of the D-Asn side-chain amide to the amine functionality within the cavity is responsible for this selectivity.

Conclusions: Computer-aided design and combinatorial library screening methods combine to provide a powerful approach to induce and evaluate the results of rational changes in the molecular recognition properties of molecules. Using this approach, we modified the binding properties of a class of molecules to select for hydrogen-bonding residues instead of hydrophobic residues and concomitantly increased the overall sequence selectivity. Structural studies indicate that these changes indeed result from the type of binding mode proposed as part of the initial design. This approach can increase our understanding of molecular recognition processes, and should allow the rational design of larger, more selective systems.

Chemistry & Biology April 1995, **2**:205–212

Key words: combinatorial libraries, molecular modeling, molecular recognition

Introduction

One of the remarkable features of many natural receptors, such as enzymes and antibodies, is their ability to recognize and selectively bind multi-residue peptide sequences. Neurotensin binding by neurotensin receptors and enkephalin binding by opioid receptors provide two examples from neurological pathways (for reviews, see [1,2]). The specific binding of Acyl-D-Ala-D-Ala by the N-Ac-muramyl-tripeptide:D-Ala-D-Ala ligase in bacterial peptidoglycan synthesis represents a clinically important case of an enzyme of this type [3].

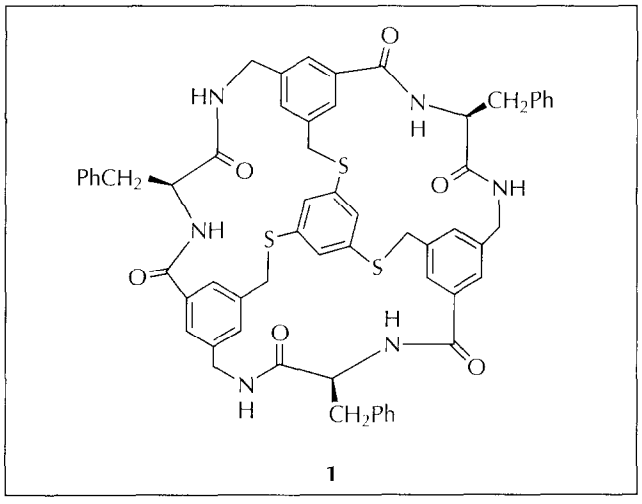
Recently, synthetic receptors have begun to reproduce the ability to bind selectively to ligands [4–8]. Among these are a class of C_3 symmetric hosts synthesized in our group and represented by the parent macrocycle **1** in Fig. 1 [4,6,7]. Composed of a hydrophobic cavity of benzene rings and a periphery of hydrogen bond donors and acceptors, this bucket-shaped molecule showed high enantio- and residue-selectivities in the binding of amino acids and short peptides. As might be expected, this selectivity was driven in part by the binding of lipophilic, unfunctionalized residues inside the cavity.

Although receptor **1** showed remarkable properties, its discovery was less than truly rational in that many other hosts were synthesized before the highly selective binder, receptor **1**, was found. Ideally, one would be able to design and synthesize a single host molecule to bind a specific target. Accomplishing this goal should not only allow the creation of synthetic receptors rivaling their natural counterparts but also could lead to better understanding and manipulation of those natural receptors. While accurate *de novo* design is not yet possible, we felt that we should be able to rationally modify known structures to achieve predictable changes in their binding abilities. Part of that confidence sprang from the ability to combine two powerful techniques: molecular modeling and encoded combinatorial library screening. Molecular modeling allows us to restrict our synthetic efforts to those molecules that are predicted to exhibit certain desirable properties. Usually these properties include a conformationally restricted structure, a well defined binding cavity, and the proper array of functionality complementary to that of certain types of guests. Combinatorial library screening methods now allow us to rapidly and efficiently determine the binding properties of the synthesized hosts. Thus, one can test a receptor with enormous numbers of guests (often $> 10^4$) simultaneously and, by analyzing the tightest binders, determine its selectivity.

Armed with these tools, we imagined modifying our macrocycle **1** to alter its selectivity. Because its lipophilic

*Corresponding author.

 ISSN 1074-5521



1

Fig. 1. Structure of C_3 macrocylic host **1**.

cavity had shown a preference for binding N-methyl amides of certain amino acids, we reasoned that including functionality capable of forming hydrogen bonds within the cavity would change that selectivity to guests with hydrogen-bonding capabilities such as those containing primary amides. We report here the design, synthesis, and study of the novel C_3 symmetric receptor **2a** (Fig. 2) based on this strategy.

Results and discussion

Design and synthesis of the macrocycle

Previous C_3 symmetric receptors we reported [4,6,7] had binding cavities that only provided lipophilic, generally unfunctionalized environments. As a result, these receptors primarily showed selective binding of hydrophobic residues. We expected that addition of hydrogen-bonding functionality within the binding cavity would change the types of amino-acid sequences that were preferentially bound. We also reasoned that increasing the functional group diversity in the cavity would increase the overall binding selectivity. In view of the C_3 symmetry, it seemed natural to incorporate this functionality by designing a host whose bottom contained a tertiary amine. Accordingly, a series of receptors were designed and submitted to Monte Carlo conformational searching to test for their conformational stability. The low energy conformations generated for each receptor were evaluated for whether they maintained a well defined binding cavity. While many of the designs showed low energy conformations with collapsed cavities, macrocycle **2** was predicted by the calculations to exist as a single, low energy conformation that corresponded to the desired bucket-shaped molecule with the amine lone pair pointing up into the cavity. That receptor was therefore chosen for synthesis.

Macrocycle **2** was prepared in nine steps as detailed in Fig. 3. The previously described bromide **3** [4] was oxidized to the corresponding aldehyde and then subjected to Wittig olefination followed by hydroboration/oxidation to provide the alcohol **6**. This alcohol served as a branch point for formation of the alkylation precursors: mesylation of compound **6** followed by bromide displacement gave bromide **7**, while mesylation followed by azide displacement and reduction provided the amine **8**. Alkylation of compound **8** with an excess of compound **7** then formed the tertiary amine **9**. Straightforward peptide synthesis procedures converted compound **9** to the (L)-phenylalanine-coupled product **10** and subsequently to the macrocyclization precursor **11**. All these reactions could be performed on large scale and in good yield (20 % overall to compound **11**). Deprotection of compound **11** and treatment with mild base under high dilution macrocyclization conditions then provided host **2** as two non-interconvertible conformers **2a** (30 %) and **2b** (40 %). Of these, only receptor **2a** showed appreciable binding properties (see below). The analogous receptors **12a,b**, in which the (L)-phenylalanines are replaced by O-allyl-protected (L)-tyrosines, were also synthesized. In this case, the binding conformer **12a** was allyl-deprotected ($Pd(PPh_3)_4$/dimedone) and coupled to a Disperse Red dye-*p*-toluic acid conjugate to provide receptor **13**, whose brilliant red color allowed its screening for binding against an encoded combinatorial library.

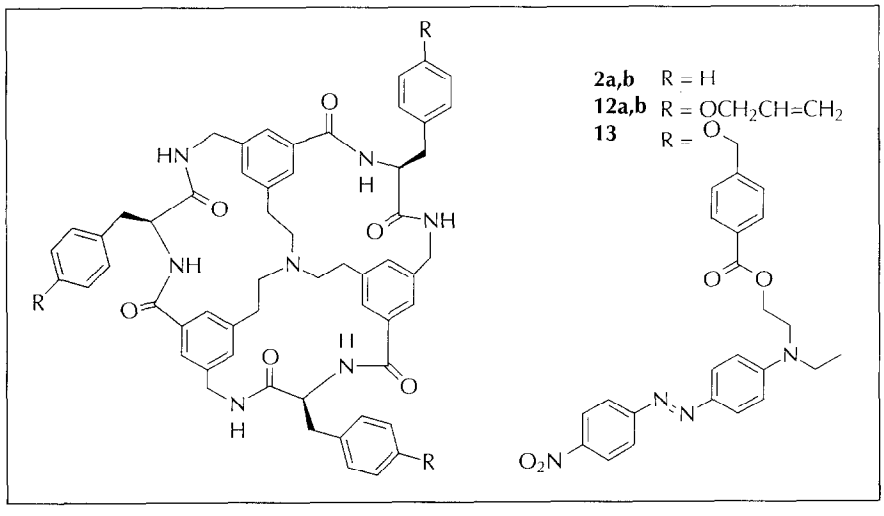


Fig. 2. Structure of macrocycles **2**,**12**, and **13**.

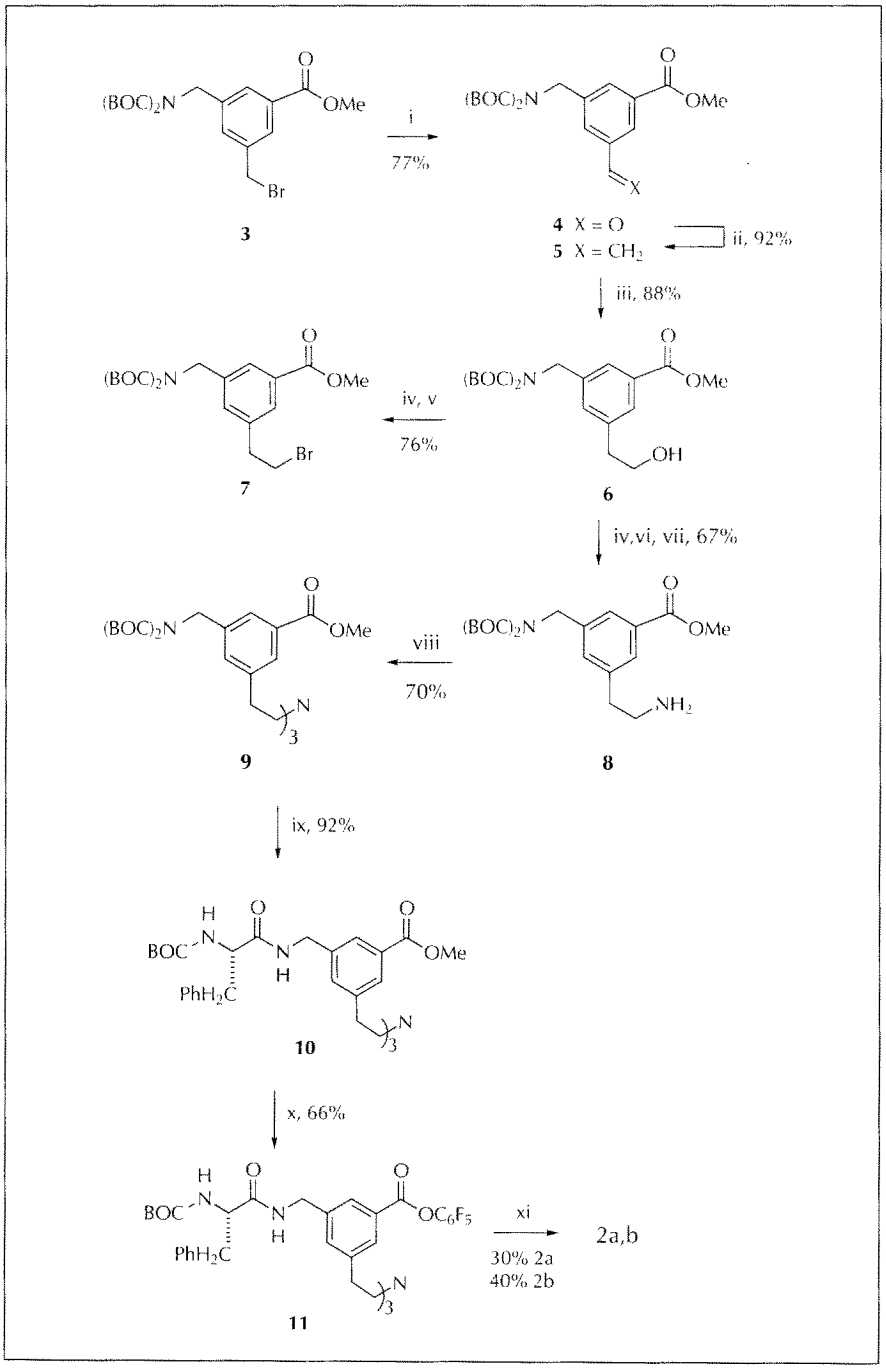


Fig. 3. Synthesis of macrocyclic receptor **2a**. Starting from compound **3** [4], the following conditions were used for the subsequent transformations: (i) $(n\text{-}Bu_4N)_2Cr_2O_7$, $CHCl_3$, Δ; (ii) $NaN(SiMe_3)_2$, Ph_3PCH_3Br, THF, 0° C; (iii) Disiamyl borane, THF, 0° C to rt; then NaOH, H_2O_2, 0° C; (iv) CH_3SO_2Cl, NEt_3, CH_2Cl_2; (v) LiBr, acetone; (vi) NaN_3, DMF; (vii) Ph_3P, Na_2CO_3, H_2O, CH_2Cl_2; (viii) Na_2CO_3, CH_3CN, 7 (250 mol %), Δ; (ix) AcCl, MeOH; then BOC-L-Phe-*p*-nitrophenyl ester, *i*-Pr_2NEt, DMF, 0° C to rt; (x) LiOH, dioxane/H_2O, 36 h; then 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide, C_6F_5OH, THF; (xi) trifluoroacetic acid, CH_2Cl_2, anisole; then *i*-Pr_2NEt, DMA, THF, 3 d.

Screening of the macrocycle against a combinatorial peptide library

Receptor **13** was screened against an encoded combinatorial peptide library using previously described procedures [9]. This >50 000-member library consisted of amino-terminally acylated tripeptide sequences attached to aminomethylated polystyrene beads. The general structure and listing of the various residues used are shown in Fig. 4. After equilibrating the library in a $CDCl_3$ solution of **13** for several days on a wrist-action shaker, ~0.5 % of the beads had become deep red in color. Eighty-five of these red beads were selected and decoded by electron capture gas chromatography, providing 65 unique binding sequences. The sequence-independent percentage of appearance of each residue at each position is shown graphically in Fig. 5. One immediately sees that the selectivity is highest in the amino acid 1 (aa_1) and aa_2 positions and decreases toward the cap end of the peptides. Moreover, the data show single site preferences for either D-Pro or Gly at aa_2 and D-Asn or D-Ser at aa_1. While these single site preferences alone are interesting, more remarkable is the pairwise selectivity that is seen within complete sequences. Based simply on the single residue binding frequencies, one would expect to find sequences containing both D-Pro-D-Asn and D-Pro-D-Ser terminal regions. D-Pro in aa_2, however, is never accompanied by D-Ser in aa_1, but indeed 83 % of the aa_2 = D-Pro sequences have aa_1 = D-Asn. Similarly, Gly in aa_2 is never accompanied by D-Asn in aa_1, but is found 78 % of the time accompanied by D-Ser. Thus, as well as showing a

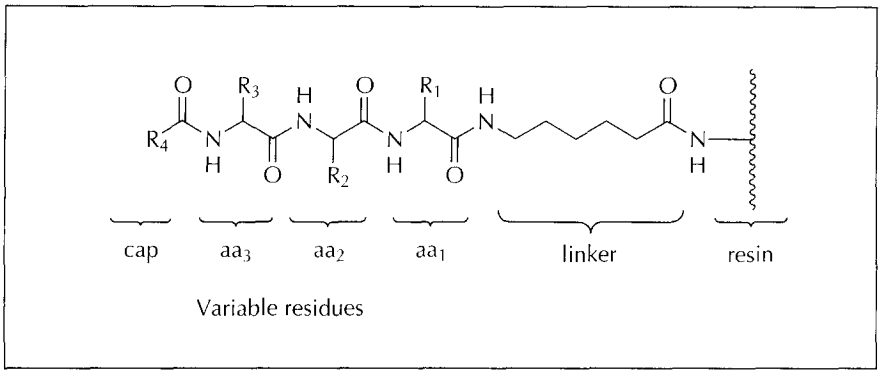


Fig. 4. Generic structure of combinatorial library beads. Listed below are the possible residues found in the variable positions. Amino acids used: Gly, D-Ala, L-Ala, D-Ser, L-Ser, D-Val, L-Val, D-Pro, L-Pro, D-Asn, L-Asn, D-Gln, L-Gln, D-Lys, L-Lys. Capping groups (caps) used: C(O)R, where R = methyl (Me), ethyl (Et), isopropyl (*i*-Pr), *tert*-Butyl (*t*-Bu), neopentyl (neo-Pe), CF_3, isobutyl (*i*-Bu), $MeOCH_2$ (MOM), cyclopropyl (c-Pr), cyclobutyl (c-Bu), cyclopentyl (c-Pe), $AcOCH_2$, phenyl (Ph), Me_2N, morpholino (Morph).

preference for certain residues, receptor **13** shows a preference for binding two distinct families of peptides, one ending in D-Pro-D-Asn and one in Gly-D-Ser.

The screening results also allow for a measure of the overall selectivity of a host molecule. Based on statistical analysis of the number of duplicate sequences found, we are able to estimate the total number of sequences in the library to which the receptor should bind during the assay. Thus we find that receptor **13** (~ 100 μM) should bind to 150 of the total 50 625 sequences in the library. By this criterion, receptor **13** (and therefore receptor **2a**) can be judged to be 5–10 times more selective than the previously described C_3 symmetric hosts.

Investigation of sequence selectivity

The pairwise sequence selectivity was further explored via solution-phase binding studies. Various N-acylated tripeptide sequences were synthesized in which only the two carboxy-terminal amino acids were altered. With the less selective positions held constant, we asked how changes in chirality, size, and hydrogen-bonding ability in the key amino-acid pairs affected the binding energies. The binding energies for receptor **2a** binding to each peptide were determined by ^1H NMR titrations in $CDCl_3$ (Table 1).

In each of the two tightly bound peptide families, small variations in the tight-binding parent sequence cause large losses of binding energy. For example, in the D-Pro-D-Asn family (entries 1–8), lengthening the side chain (entry 4) or changing the chirality of either key residue (entries 3, 6) results in at least a two orders of magnitude decrease in binding. The most drastic drops result from altering the nature of the Asn side chain. Changing from an amide to an N-methyl amide (entry 7) results in a >1000-fold decrease in binding; failure to remove the trityl protecting group (entry 8) obliterates binding altogether. Interestingly, all these changes cause a greater loss in binding energy than that seen when the constant amino-acid residue is removed (entry 2).

The Gly-D-Ser family (entries 9–19) shows related but qualitatively different results. Progressive enlargement of the Gly residue (entries 12, 18) causes a progressive decrease in binding energy. Similarly, chirality change of the carboxy-terminal residue (entry 17) results in a greater than 100-fold decrease in binding. However, alteration of the D-Ser side chain has a less pronounced effect. Loss of the hydroxyl group (entry 14) or replacing it with a methyl group (entry 13) only result in a 1–1.5 kcal mol^{-1} decrease in binding energy. More strikingly, lengthening of the side chain (entry 10) or methylation of the hydroxyl (entry 11) cause little appreciable decrease in binding. In this family the role of the amino-terminal amino acid appears more important: the complete loss of this residue produces a 2 kcal mol^{-1} loss of stability (entry 15).

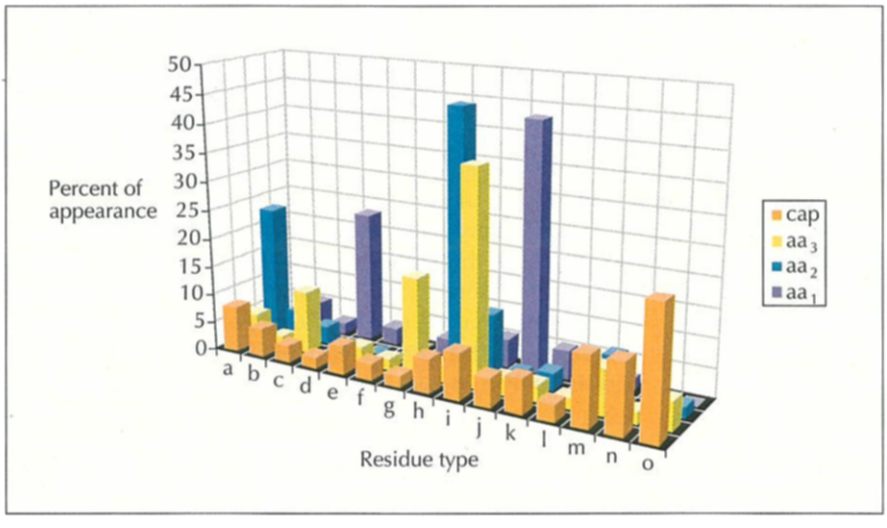


Fig. 5. Graph of the percentage appearance of residues by position. Key for cap position (given as R for generic C(O)R): a = Me, b = Et, c = *i*-Pr, d = *t*-Bu, e = neo-Pe, f = CF_3, g = *i*-Bu, h = MOM, i = c-Pr, j = c-Bu, k = c-Pe, l = $AcOCH_2$, m = Ph, n = Me_2N, o = Morph. Key for aa positions: a = Gly, b = D-Ala, c = L-Ala, d = D-Ser, e = L-Ser, f = D-Val, g = L-Val, h = D-Pro, i = L-Pro, j = D-Asn, k = L-Asn, l = D-Gln, m = L-Gln, n = D-Lys, o = L-Lys.

Table 1. Binding study results with receptor **2a**.

Entry	Sequence	Saturation (%)	ΔG (kcal mol^{-1})
1	R-D-Pro-D-Asn-R'	96	–6.3
2	*i*-PrCO-D-Pro-D-Asn-R'	86	–5.1
3	R-D-Pro-L-Asn-R'	65	–4.3
4	R-D-Pro-D-Gln-R'	50	–4.0
5	R-D-Pro-D-Homoser-R'	51	–4.0
6	R-L-Pro-D-Asn-R'	48	–3.9
7	R-D-Pro-D-Asn(γ-Me)-R'	25–30	~ –2.5
8	R-D-Pro-D-Asn(γ-Tr)-R'	—	> –1
9	R-Gly-D-Ser-R'	97	–6.6
10	R-Gly-D-Homoser-R'	98	–6.5
11	R-Gly-D-Ser(OMe)-R'	98	–6.1
12	R-L-Ala-D-Ser-R'	90	–5.4
13	R-Gly-D-Etgly-R'	92	–5.4
14	R-Gly-D-Ala-R'	85	–5.1
15	*t*-BuCO-Gly-D-Ser-NH(*n*-C_4H_9)	65	–4.5
16	R-Gly-L-Ala-R'	70	–4.2
17	R-Gly-L-Ser-R'	60	–4.2
18	R-D-Pro-D-Ser-R'	25	–3.2
19	R-Gly-D-Ser(O*t*-Bu)-R'	41	–2.3

All results were determined by NMR titrations in $CDCl_3$ as described; binding energies over 4.5 kcal mol^{-1} always represent averages of the calculated energies from the shifts of at least two different protons. Key: R = EtCO-L-Pro-; R' = -NH(*n*-$C_{12}H_{25}$); Homoser = homoserine; Etgly = ethylglycine; Tr = trityl.

Analysis of the binding modes

A combination of computer simulations, inferences from the binding data, and NMR structural studies have provided a picture of the binding modes for each family. Each family appears to bind in a distinct way. Of the two, the D-Pro-D-Asn family is the more relevant to the model envisioned in the original receptor design. Not only did these sequences appear with greater frequency in the library assay, but also only the D-Asn family appears to have a binding mode that uses the engineered tertiary amine.

Selectivity for the D-Pro-D-Asn motif results from recognition of the Asn side-chain amide by the amine at the bottom of receptor **2a**. A Monte Carlo conformational search of Ac-L-Pro-D-Pro-D-Asn-NH(*t*-Bu) with the lowest energy conformation of receptor **2a** provided 29 structures within 3 kcal mol^{-1} of the global energy minimum. All these conformations showed a hydrogen bond between the side-chain amide and the receptor amine. The lowest energy structure found is shown in Fig. 6. This hydrogen bond exists as a bifurcated arrangement of the two amide hydrogens with the receptor amine. The steric constraints of the cavity appear to prevent adoption of the optimal linear hydrogen bond geometry with a single amide hydrogen. The stability of this unusual arrangement was investigated by performing lengthy stochastic dynamics simulations starting from the previously calculated minimum energy complex. During several 1000 ps simulations at 300 K, the D-Asn side chain remained within the cavity and each of the amide protons stayed at a 3 Å average distance from the amine nitrogen. In addition to showing the hydrogen bond inside the cavity, the Monte Carlo-derived structure also shows how the remaining carbonyls and amides can array themselves to achieve the largest number of hydrogen bonds on the periphery. Interestingly, when the D-Asn residue was replaced by D-Gln, a Monte Carlo conformational search of the resulting complex provided 171 structures within 3 kcal mol^{-1} of the global minimum energy structure. Of these, only 67 structures placed the D-Gln side chain inside the receptor cavity, and the first structure showing this arrangement appeared 0.83 kcal mol^{-1} above the global minimum.

The calculated picture correlates well with the binding study results (Table 1). In general, alterations in the most important recognition elements cause the greatest drop in binding energies. Clearly, the largest losses of binding energy result from altering the Asn side chain. Slightly lengthening the side chain (entry 4) significantly lowers binding, as does changing the amide to a hydroxyl (entry 5). Most importantly, the greatest drop in binding is seen by simple methylation of the amide (entry 7). Were this amide simply providing a hydrogen bond on the periphery, only a small drop in binding would be expected. The only environment in which the methylated amide can no longer act as a hydrogen bond donor is within the steric confinement of the cavity.

NMR structural study results are also consistent with this binding mode. Several important observations were

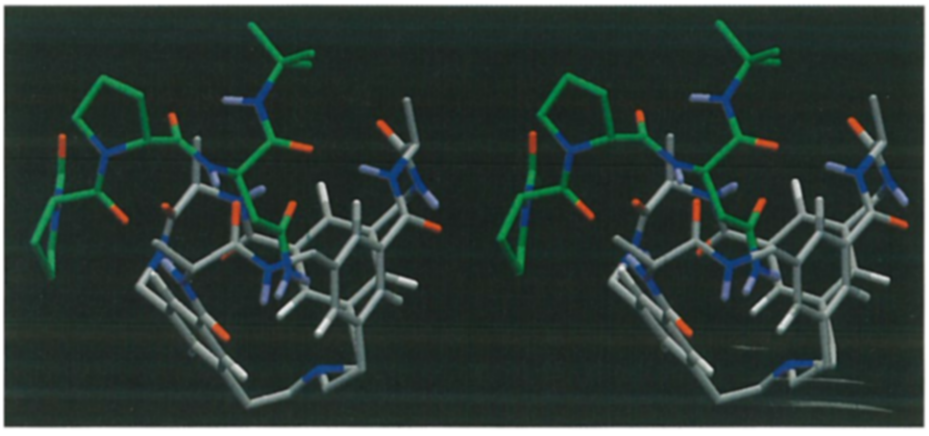


Fig. 6. Stereoview of the global minimum energy structure of macrocycle **2a** with Ac-L-Pro-D-Pro-D-Asn-NH(*t*-Bu).

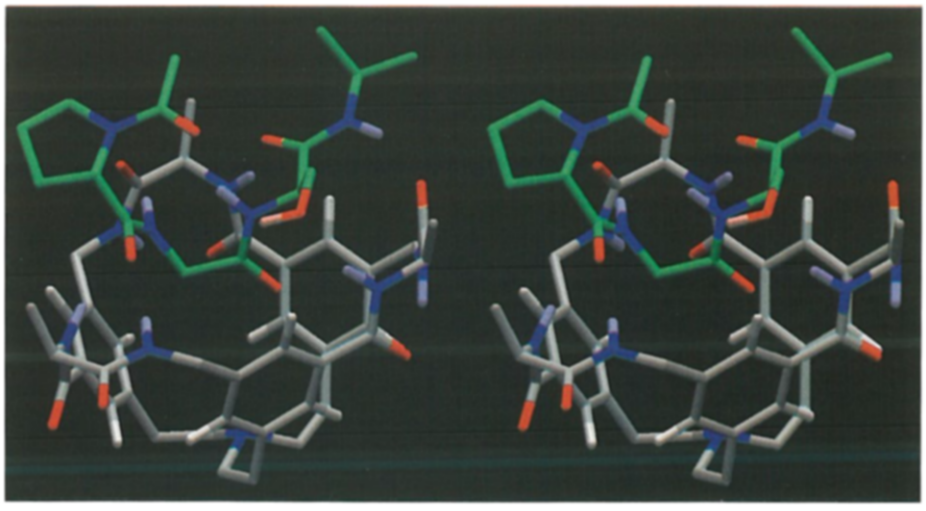


Fig. 7. Stereoview of the global minimum energy structure of macrocycle **2a** with Ac-L-Pro-Gly-D-Ser-NHMe.

seen during investigations of receptor **2a** with the guest EtCO-L-Pro-D-Pro-D-Asn-$NHC_{12}H_{25}$. First, during binding studies, the Asn side-chain amide protons, in contrast to the other guest amide protons, are the only ones that give evidence of being in a slow exchange environment. Secondly, the two Asn side-chain methylene protons exhibit large downfield shifts upon binding that correspond to 0.8 and 1.2 ppm. These shifts are difficult to rationalize by anything other than the notion that these protons are being held near the edge of a benzene ring [10], as would be the case in the postulated binding mode. Lastly, rotating-frame Overhauser effect spectroscopy (ROESY) cross peaks are observed between the L-Pro ring and the periphery of the receptor. Together, these results support the computer-generated structure qualitatively, if not precisely. In any case, the combined evidence clearly implicates binding of the Asn side chain to the tertiary amine within the macrocycle cavity.

The Gly-D-Ser sequences exhibit a different and unexpected binding mode. A Monte Carlo conformational search of Ac-L-Pro-Gly-D-Ser-NHMe with the low energy conformation of receptor **2a** provided six structures within 3 kcal mol^{-1} of the global minimum. In each of these structures, the peptide adopts a β-turn conformation, placing the glycine partly into the cavity, and uses the serine hydroxyl as a hydrogen bond acceptor on the periphery (Fig. 7). Although structural studies are not yet complete, the peptide binding selectivities are consistent with this model (Table 1). Lengthening the serine side chain (entry 10) and replacing the hydroxyl proton with a methyl group (entry 11) cause almost no loss of binding energy. Neither result can be rationalized with a mode involving side-chain recognition by the cavity and the tertiary amine; both are consistent with this side chain acting as an acceptor on the periphery. Further calculations have indicated that selectivity for this sequence may be determined by **2a** having a perfectly complementary hydrogen-bonding array for the normal solution conformation of this peptide.

Significance

The ability to predictably manipulate the binding properties of molecules would revolutionize many aspects of chemistry and biology. Areas such as drug design, *de novo* enzyme synthesis, and enzyme-inhibition studies would reap large rewards. While complete realization of this ability remains outside our grasp, significant progress has been made by studying the host-guest interactions of smaller, well defined systems. Using the methodologies of molecular modeling, organic synthesis, and combinatorial chemistry, it is now possible to design and synthesize novel receptors that are expected to show favorable binding properties and gain a broad picture of their binding properties by rapid screening against enormous numbers of substrates. So far, our limited knowledge precludes the reliable design of receptors having preordained selectivities for given guests. However, modification of known structures to tailor their selectivities is clearly achievable.

In this work we began with a highly selective receptor and modified its structure to change the nature and magnitude of its selectivity. As one would do with an enzyme, we targeted the alteration for the interior of the binding cavity, reasoning that changes in this region would have the greatest effect on binding properties. Replacing a benzene ring on the bottom with a tertiary amine, we added hydrogen bonding functionality to an apolar environment. This modification of the binding pocket led not only to a change in the types of residues bound but also increased the overall sequence selectivity of the receptor.

Thus we created and studied a novel, highly selective receptor and simultaneously emphasized the power of current methodologies. Accumulation of such studies in conjunction with improvements in techniques will lead to better understanding and development of other small molecule binding

systems. It should also lead to a better understanding and more accurate manipulation of their larger, biological counterparts.

Materials and methods

Computer modeling

Monte Carlo conformational searching was performed using the MCMM method [11] in MacroModel/Batchmin version 4.5 using GB/SA chloroform [12] and the AMBER* force field [13]. In the host-guest searches, an *ab initio*-fitted set of force field parameters for proline were added (Q. McDonald, personal communication). Typically, conformational searches were carried out for >40 000 steps in blocks of 5000 steps. Searches were continued until a new block produced no new low energy structures. In each step, a random conformation was generated, energy minimized using a Polak–Ribier conjugate-gradient algorithm, evaluated for fulfilling chirality and torsional constraints, and retained if its calculated energy was within 50 kJ mol^{-1} of that of the current lowest energy conformation. The structures remaining after the initial search were re-minimized to a gradient of less than 0.05 kJ mol^{-1} $Å^{-1}$, checked for uniqueness, and stored. In the host–guest searches, all allowable torsions of the guest were explored in addition to its rotations and translations relative to the host. No attempts were made to rotate peptide bonds in the guest from their initial *trans* positions. Additionally, no attempts were made to explicitly explore the conformational space of the host.

Molecular dynamics simulations were performed using the stochastic dynamics method [14] in MacroModel/Batchmin version 4.5. The force field and solvent model were as used in the conformational searching. Three separate simulations were run for 1000 ps with a time step of 1.5 fs and a bath temperature of 300 K. The starting geometry for each simulation was the lowest energy host–guest complex found in the Monte Carlo conformational searching. During each simulation, the distance between the D-Asn side-chain amide protons of the guest and the tertiary amine of the host were monitored.

Preparation of synthetic intermediates

Compounds were prepared by chemical synthesis as outlined in Fig. 3 using standard laboratory procedures. All compounds were characterized by infrared spectroscopy (IR), ^1H and ^{13}C NMR, and high resolution mass spectrometry (HRMS).

*Preparation of receptor **2a***

To a stirring solution of compound **11** (523 mg, 0.293 mmol) in CH_2Cl_2 (75 ml), anisole (2 ml) and trifluoroacetic acid (25 ml) were added. Stirring was continued for 3 h, and then the solvent was evaporated. Several times the residue was dissolved in toluene and then evaporated under high vacuum rotary evaporation. The glassy residue was then dissolved in N,N-dimethylacetamide (DMA) (8–9 ml) and drawn into a 10 ml gastight syringe. This solution was added via syringe pump over 50–60 h to a stirring solution of $NEt(i\text{-}Pr)_2$ (7.7 ml, 43.9 mmol) in THF (250 ml). Stirring was continued for 1 d after the addition was complete. The solvent was then evaporated, taking care to remove all residual DMA. Column chromatography ($MeOH/CHCl_3$, 5:95) followed by size exclusion gel filtration (Sephadex LH-20, $MeOH/CHCl_3$, 50:50) gave receptor **2a** (83 mg, 0.089 mmol, 30 %) as a white solid. The non-binding isomer **2b** was also formed in the reaction (30–45 %). Data for receptor **2a**: R_F 0.19 (silica gel, $MeOH/CHCl_3$ 5:95); ^1H NMR (400 MHz, $CDCl_3$) δ 7.20–7.41 (m, 24H, Ar + $ArCH_2NHCO$), 6.86 (s, 3H, Ar), 6.22 (d, J = 8.0 Hz, 3H, $ArCONHCH$), 4.72 (dd, J = 7.5, 15.4 Hz, 3H, $NHCH(CH_2Ph)CO$), 4.37 (dd, J = 8.3, 14.3 Hz, 3H, $ArCH_2N$), 4.02 (dd, J = 4.1, 14.3 Hz, 3H, $ArCH_2N$), 3.49 (bt, J = 12.2 Hz, 3H, $ArCH_2CH_2N$), 3.33 (dd, J = 7.1, 14.2 Hz, 3H, $NHCH(CH_2Ph)CO$), 3.19 (m, 3H, $ArCH_2CH_2N$), 3.04 (dd, J = 7.7, 14.1 Hz, 3H, $NHCH(CH_2Ph)CO$), 2.82 (bd, J = 17.1 Hz, 3H, $ArCH_2CH_2N$), 2.14 (bd, J = 14.0 Hz, 3H, $ArCH_2CH_2N$); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.1, 168.7, 141.5, 138.4, 136.6, 133.4, 132.9, 129.2, 128.9, 127.1, 124.8, 123.1, 54.5, 51.7, 42.9, 35.5, 31.8; IR (neat) 3314, 3062, 2932, 1660, 1651, 1599 cm^{-1}; HRMS calc'd for $C_{57}H_{58}N_7O_6$ (MH^+): 936.4449, found 936.4464.

Preparation of peptide sequences

All peptide sequences were synthesized chemically by standard solution-phase methods. Each completed sequence was characterized by ^1H and ^{13}C NMR and HRMS.

Binding studies

All binding titrations were performed in $CDCl_3$ using a 400 MHz NMR spectrometer to record the ^1H spectra. The concentrations of the initial host solutions were ~1.0 mM. The concentrations of the added guest solutions varied from 3.5 mM to over 20 mM, depending on the magnitude of the binding constant being measured. In a typical procedure, 700 μl of the host solution was placed in an NMR tube and its spectrum was recorded. Six additions of guest solution were made, starting with 12 μl and doubling the added amount each time to give a final addition of 384 μl and a total addition of 756 μl. Spectra were taken after each successive addition, and the chemical shifts of various host protons were recorded. Nonlinear fitting of these measurements to a binding curve of the type A + B ↔ AB gave the binding energies.

Supplementary material available

Complete preparation details and characterization for compounds **4–11**, physical data for **2b**, and the 85 sequences decoded in the combinatorial assay are available.

Acknowledgements: We thank Dr Quentin McDonald for performing many of the calculations. M.R.C. wishes to thank the NSF for a predoctoral fellowship and the Ford Foundation for a minority dissertation fellowship. This work was also supported by NIH grant HL25634.

References

1. Elliot, P.J. & Nemeroff, C.B. (1986). The neurobiology of neurotensin. In *Neural and Endocrine Peptides and Hormones*. (Moody, T.W., ed.), pp. 219–245, Plenum Press, New York.
2. Pasternak, G.W. (1988). Studies of multiple morphine and enkephalin receptors: evidence for mu_1. In *Advances in Experimental Medicine and Biology, Vol. 236: Neuroreceptors and Signal Transduction*. (Kito, S., Segawa, T., Kuriyama, K., Tohyama, M. & Olsen, R.W., eds), pp. 81–93, Plenum Press, New York.
3. Neuhaus, F.C. & Struve, W.G. (1965). Enzymatic synthesis of analogs of the cell wall precursor. I. Kinetics and specificity of uridine diphospho-N-acetylmuramyl-L-alanyl-D-glutamyl-L-lysine:D-alanyl-D-alanine ligase (adenosine diphosphate) from *Streptococcus faecalis* R*. *Biochemistry* **4**, 120–131.
4. 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.
5. Yoon, S.S. & Still, W.C. (1993). An exceptional receptor for peptides. *J. Am. Chem. Soc.* **115**, 823–824.
6. Borchardt, A. & Still, W.C. (1994). Synthetic receptor for internal residues of a peptide chain. Highly selective binding of (L)X-(L)-Pro-(L)X tripeptides. *J. Am. Chem. Soc.* **116**, 7467–7468.
7. Yoon, S.S. & Still, W.C. (1994). Expeditious synthesis of a novel

C_3-symmetric receptor that binds peptides sequence-selectively. *Angew. Chem. Int. Ed. Engl.* **33**, 2458–2460.
8. Albert, J.S., Goodman, M.S. & Hamilton, A.D. (1995). Molecular recognition of proteins: sequence-selective binding of aspartate pairs in helical peptides. *J. Am. Chem. Soc.* **117**, 1143–1144.
9. Borchardt, A. & Still, W. C. (1994). Synthetic receptor binding elucidated with an encoded combinatorial library. *J. Am. Chem. Soc.* **116**, 373–374.
10. Johnson, Jr., C.E. & Bovey, F.A. (1958). Calculation of nuclear magnetic resonance spectra of aromatic hydrocarbons. *J. Chem. Phys.* **29**, 1012–1014.
11. Chang, G., Guida, W.C. & Still, W.C. (1989). An internal coordinate Monte Carlo method for searching conformational space. *J. Am. Chem. Soc.* **111**, 4379–4386.
12. Still, W.C., Tempczyk, A., Hawley, R.C. & Hendrickson, T. (1990). Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112**, 6127–6129.
13. McDonald, D.Q. & Still, W.C. (1992). AMBER* torsional parameters for the peptide backbone. *Tetrahedron Lett.* **33**, 7743–7746.
14. van Gunsteren, W.F. & Berendsen, H.J.C. (1988). A leap-frog algorithm for stochastic dynamics. *Mol. Simul.* **1**, 173–185.

Received: **8 Mar 1995**. Accepted: **28 Mar 1995**.