Pergamon

Tetrahedron 56 (2000) 3309–3318

TETRAHEDRON

A Synthetic Receptor Motif Designed for Extended Peptide Conformations

Kevin Ryan, Leland J. Gershell and W. Clark Still*

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

Received 23 December 1999; accepted 22 March 2000

Abstract—The intrinsic flexibility of short peptides makes them difficult targets for molecular recognition. Although in general too short to form stable secondary structures, many are able to sample extended conformations. Using monomer cycloalkyl-1,2-*trans*-diamines alternated with cycloalkyl-1,2-*trans*-dicarboxylic acids, we have constructed a novel receptor motif to be complementary to extended tripeptides. Receptor design, synthesis, structural analysis and binding studies are presented.

Introduction

The inherent flexibility of short peptides (2–15 amino acids) makes them difficult targets for molecular recognition. Artificial receptors designed to bind them face, in the case of most sequences, a plethora of different unbound conformers at any given instant. A general rule in host/guest chemistry is that a host must be pre-organized to bind a target lest too much conformational entropy and enthalpy be lost by the host (and the guest) upon binding to be favorable.[1] This means that, ideally and impossibly, a receptor should be preorganized to bind many different peptide conformers. A more realistic expectation is that a well designed receptor should be able to bind a small family of related conformers that are frequently sampled by the unbound peptide. Despite this problem of target flexibility, during the past few years we[2] and others[3,4] have made progress in the sequence-specific recognition of short peptide sequences in organic solvents and, to a lesser extent, in water.[5,6] Many of our own receptors have been constructed from macrocycles designed to form pre-organized molecular clefts lined with hydrogen-bond donors and acceptors.

In the process of screening our receptors against encoded combinatorial tripeptide libraries, we have noted a modest but recurring preference by some receptors for heterochiral tripeptide sequences that contain proline (see Refs. 5, 7–10 for example). Proline is the amino acid most often associated with β-turn structures in peptides—especially heterochiral peptides[11]—and its pyrrolidine ring makes it the most conformationally restricted of the proteinogenic amino acids. Although very little information is available regarding the three-dimensional structure of our receptors or receptor–peptide complexes, we wondered if the predilection of some of our receptors for proline might be significant. For example, does the conformational restriction of proline contribute significantly to the pre-organization of the host–guest complex? While we design pre-organization into our receptors, pre-organization in the target may also enhance binding to the point where it is brought within the limit of detection in our assay, provided there is at least a modicum of hydrogen-bonding complementarity. An equally well-matched tripeptide of greater flexibility might fall short of this threshold. Another possibility is that a tripeptide containing a proline-induced turn may be more compact and therefore be more likely to match the size of a receptor's binding pocket, thereby availing itself of a maximum number of hydrogen-bonds. Whatever the reason, this trend has prompted us to take into consideration more seriously the variable of peptide backbone conformational preference when designing new molecules in our overall pursuit of a collection of sequence-selective receptors that covers the widest variety of sequences. Although we cannot expect the tripeptide targets in our libraries to form well-defined structures like α-helices and β-sheets, we can expect many sequences to be capable of sampling extended conformations, the antithesis of reverse turn structures. Here we report the design and synthesis of a new receptor motif intended to be complementary to the extended peptide conformation, and its sequence-selective binding properties as revealed through on-bead screening of an encoded tripeptide library.

Results

Receptor design

Our design is based on oligomers formed by the condensation

Keywords: receptors; peptides; combinatorial chemistry; molecular recognition.

* Corresponding author. Tel.: +1-212-854-2577; fax: +1-212-854-4231; e-mail: clark@chem.columbia.edu

0040–4020/00/$ - see front matter
PII: S0040-4020(00)00252-0

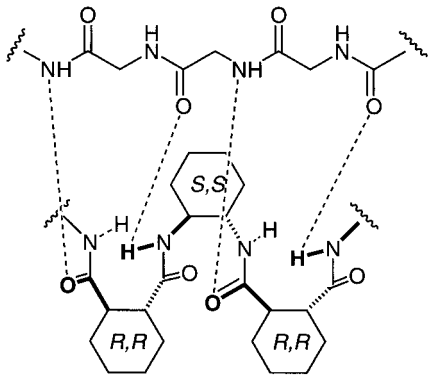
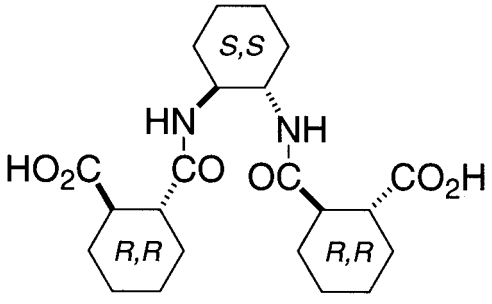


Figure 1.

of cycloalkyl-1,2-*trans*-dicarboxylic acids of one stereochemistry with cycloalkyl-1,2-*trans*-diamines of the opposite stereochemistry (Fig. 1). Molecular modeling[12] indicated that oligomers constructed from these alternating monomers should prefer to adopt a flat, extended conformation. As depicted in Fig. 1, the dipole–dipole directed alignment of the amide linkages presents an alternating array of hydrogen-bond donors (amide N–H) and acceptors (amide carbonyl) above and below the plane defined by the rings. This arrangement is reminiscent of the peptide backbone of the antibiotic vancomycin, which binds the (D)Ala–(D)Ala dipeptide in an extended conformation.[13] Models show that this diastereomer (or its enantiomer) favors the extended structure while other all-*trans* diastereomers do not. Models also show that the distance between alternating hydrogen-bond donor and acceptor is such that the array is complementary to that of a peptide in the extended conformation, with the fragment shown large enough to span a tripeptide sequence.

We verified the conformational preference for the unbound receptor motif by studying the crystal structure of compound **1**, a close analog of the receptor prototype. Compound **1**, a dicarboxylic acid-terminated version of a 1.5-mer receptor motif, crystallized in an extended conformation that was virtually superimposable on an energy minimized model of the receptor. As shown in Fig. 2, the cyclohexyl rings define a flat V-shaped surface. Cutting through the middle of this surface is the hydrogen-bond donor–acceptor array, which projects both above and below the plane. In the lattice, each molecule is extensively hydrogen bonded to others above and below the plane of the cyclohexyl rings (not shown). We found this class of receptors to be prone to insolubility and, when solubilized, to non-crystalline aggregation in aprotic solvents. This behavior may be due to the predisposition towards an extended though imperfectly formed lattice driven by the formation of hydrogen-bonds and permitted by the flat, relatively rigid conformational preference of this diastereomer.

An important feature of the receptor design that was verified by the crystal structure is that the receptor has two discrete faces, each presenting an alternating hydrogen-bond donor–acceptor array. The extended peptide conformation itself is in this respect similar as it has two discrete edges of alternating hydrogen-bond donors and acceptors which, for example, allow the extended conformation to assemble into β-sheets of more than two strands in proteins. In an effort to recognize simultaneously both edges of an extended peptide conformation, we elaborated the design by linking two receptors with an aromatic spacer to form a bidentate version of the receptor. Our intention was to allow two receptors to converge on, and hydrogen bond to, a single extended peptide. Three such compounds, **5**, **6**, and **7**, are shown in Fig. 3. The sulfonamide linkage was



1

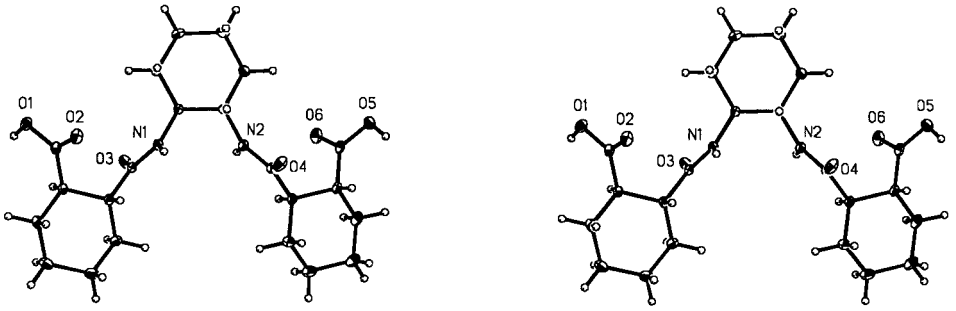


Figure 2.

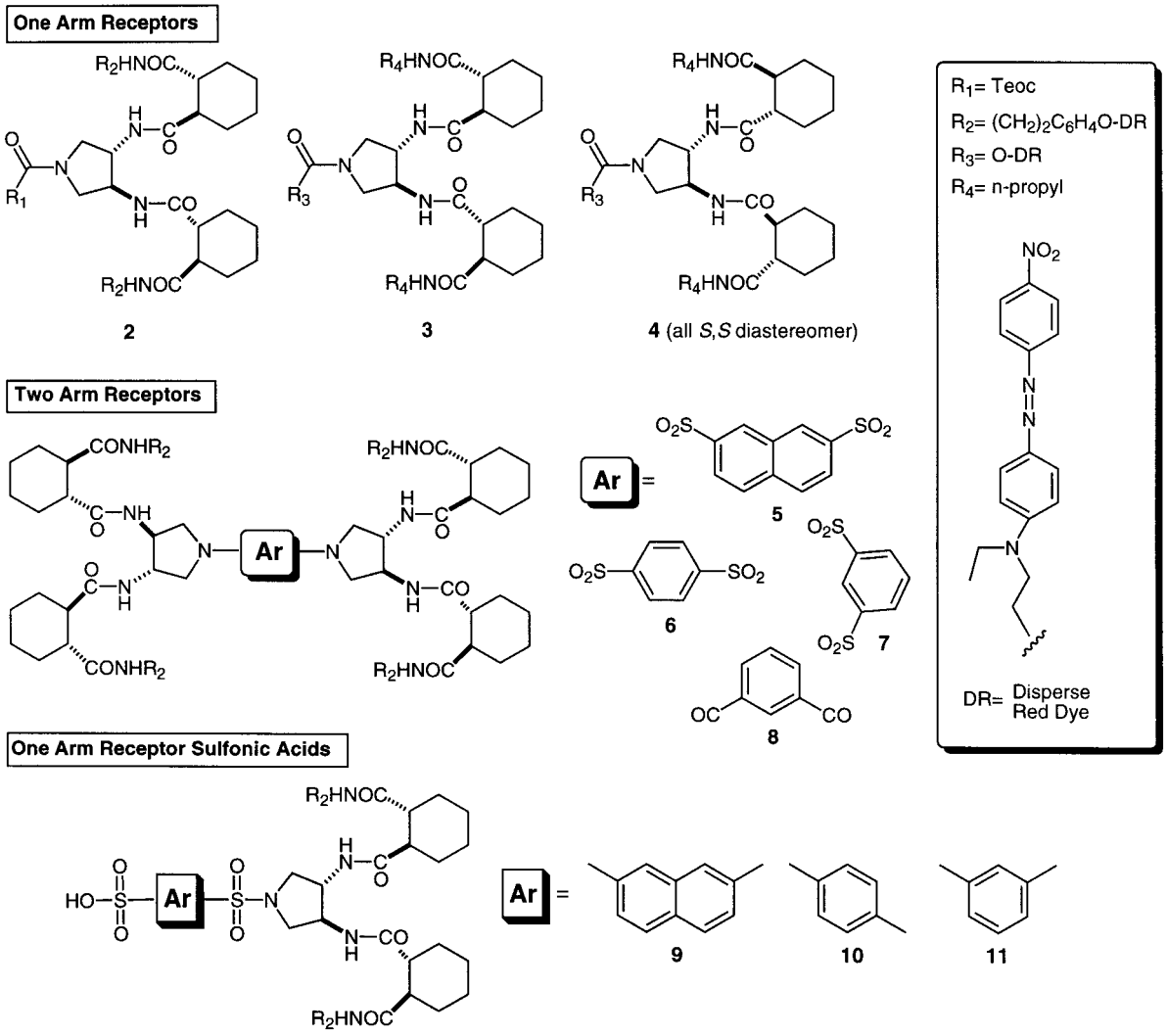


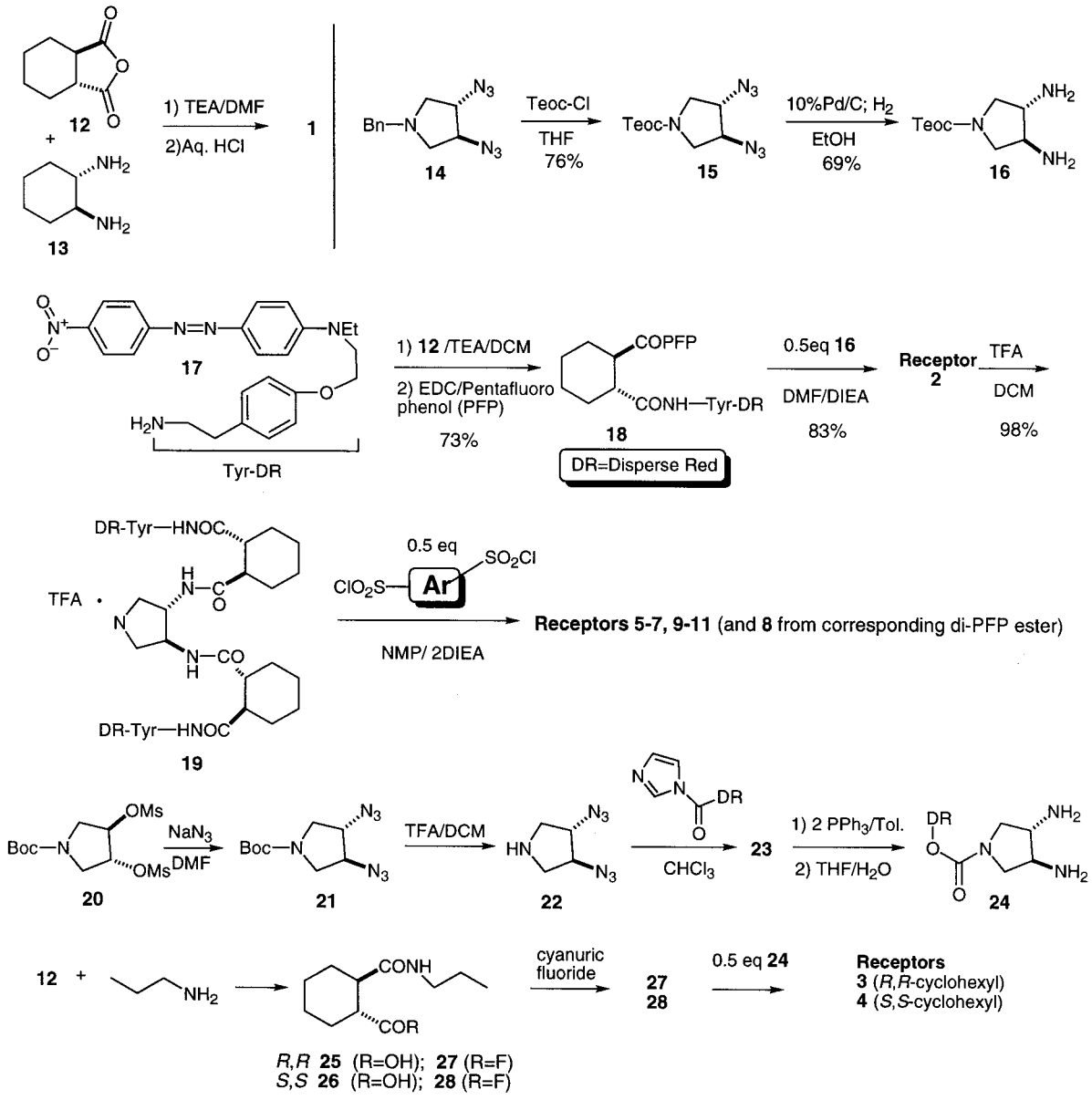
Figure 3.

selected for the tetrahedral configuration about the sulfur atom. Modeling studies indicate that this, and not the planar carboxamide linkage of analog **8**, allows the two receptor arms to align in a parallel manner to form a cavity in which an extended peptide might bind. The third side of the cavity may be defined either by the edge or more likely the face of the aromatic ring.[14] In addition to joining the two receptors and allowing a convergent conformer, the linkers were selected for their size and rigidity. Conformational searching in the absence of a guest indicated that the aromatic spacers in compounds **5** and **6**, and to a lesser extent **7**, should prevent intramolecular hydrogen bonding between the two receptor arms, thereby leaving them available and far enough apart to accommodate a tripeptide guest.

The one- and two-arm receptors synthesized for this study are shown in Fig. 3. Disperse Red (DR) dye-conjugates were made for testing against an encoded tripeptide library displayed on polystyrene beads. A pyrrolidine-3*S*,4*S*-diamine was used in place of the cyclohexyldiamine of **1** to allow attachment of the dye in receptors **3** and **4**, and to provide a site for elaboration into the bidentate form of receptors **5**–**11**. The cyclopentyl diamine compounds were also found to be less prone to insolubility and aggregation. In designing receptors for a general backbone conformation we have avoided the more difficult task of trying to design sequence specificity. By using a combinatorial library of tripeptides we initially leave this question to the assay. In the event that a receptor shows a preference for a certain sequence or family of sequences, study of these new leads can be used to design the next generation of receptors based on this novel motif. Compounds **9**–**11** are partial exceptions to this because, by virtue of their sulfonic acid group, they are expected to show some preference for tripeptides containing positively charged side chains.

Receptor synthesis

As shown in Scheme 1, the core of the receptor structure was made by condensing the appropriate cycloalkyl diamine, either **13**, **16** or **24** with a singly activated form of the diacid, either anhydride **12**, pentafluorophenol ester **18** or acid fluorides **27** and **28**, respectively. The chiral cycloalkyl diamines were either purchased (**13**) or made from the known compounds **14**[15] or **20**.[16] To prepare **16**, the benzyl group of **14** was exchanged for the Teoc group (2-(trimethylsilyl)ethoxycarbonyl) using an established



Scheme 1.

procedure.[17] This permitted the subsequent catalytic hydrogenation of the azide groups to yield diamine **16**. Two different diamino pyrrolidines were needed in order to append the DR to different sites (receptor **2** vs. **3**) on the receptor core. In the case of receptor **2**, the dye was attached to the cyclohexyl rings using a tyramine (Tyr) linker made from a Mitsunobu coupling between Boc-protected tyramine and DR to yield **17** after TFA deprotection and neutralization. For receptor **3** and its diastereomer **4**, the DR was attached to the pyrrolidine ring nitrogen of diazide **22** using DR and carbonyl diimidazole.[18] In this case, the diazide was subsequently reduced with triphenylphosphine (to preserve the diazo group) before coupling **24** to the acid fluorides **27** and **28**.

In order to make the two-arm receptors **5**–**8**, the Teoc group of receptor **2** was removed to yield **19**, which was then coupled to the various aromatic disulfonyl chlorides or, in the case of receptor **8**, the dipentafluorophenol ester of benzene-1,3-dicarboxylic acid. The sulfonyl chlorides were either purchased (**7**) or made according to standard procedures.[19,20] The single-arm receptor sulfonic acids **9**, **10** and **11** were isolated from the coupling reactions during silica gel chromatography.

Results and Discussion

Side-chain protected library

The receptor–dye conjugates were screened against two encoded, *N*-acetylated tripeptide libraries—one with the amino acid side chains protected and one with the side chains deprotected. The results are shown in Table 1. In the case of the side-chain protected library assayed in 5% *N*-methylpyrrolidinone (NMP)/chloroform, the basic receptor motif as represented by compounds **2** and **3** showed sequence selective binding of the *N*-acetylated dipeptide

Table 1. Tripeptide library binding results (notes: general conditions: 5% NMP/buffered chloroform, >48 h, rt.; library structure: Ac-AA1-AA2-AA3-NH(CH2)5CONH-polystyrene)

Cpd	Conc. μM	Lib. copies	Side-chain protected[h]					Side-chain deprotected				
			AA3	AA2	AA1	Freq[e]	Selectivity[b]	AA3	AA2	AA1	Freq[e]	Selectivity[b]
2	99	3	L-Asn	D-Val	L-Ala	4/4	E	L-Val	L-Gln	L-Gln	2/2	G
2[a]	30	3	L-Asn	D-Val	D-Ala	3/6	E	D-Val	D-Lys	D-Ala	1/4	G
			D-Ser	D-Val	D-Ala	1/6		Gly	D-Lys	D-Ala	1/4	
			D-Ser	D-Val	D-Val	1/6		D-Ala	D-Lys	D-Val	1/4	
			D-Asn	D-Gln	D-Val	1/6		L-Ser	D-Val	D-Val	1/4	
3	65	6	L-Asn	D-Val	D-Ala	5/8	G	N/A				
			L-Asn	D-Val	D-Ser	3/8						
3	50	3	L-Asn	D-Val	D-Ala	2/5	G	N/A				
			L-Asn	D-Val	D-Ser	2/5						
			L-Asn	D-Val	L-Ala	1/5						
5	25	6	L-Asn	D-Val	L-Val	2/4	G	[f]				
			L-Asn	D-Val	D-Ala	1/4						
			L-Asn	D-Val	D-Val	1/4						
6	30	3	L-Asn	D-Val	D-Val	2/2	E	L-Val	L-Gln	L-Gln	3/6	G
			L-Asn	D-Val	D-Ser	2/2		L-Val	L-Gln	D-Ser	1/6	
			L-Asn	D-Val	D-Ala	2/2		D-Val	D-Gln	L-Ala	1/6	
								L-Ser	D-Val	D-Val	1/6	
6	20	3	L-Asn	D-Val	D-Ala	3/5	G	L-Val	L-Gln	L-Gln	1/5	G
			L-Asn	D-Val	L-Val	1/5		L-Val	L-Gln	L-Ser	1/5	
			L-Asn	D-Val	[c]	1/5		L-Val	L-Gln	D-Ala	1/5	
								L-Ser	D-Val	D-Val	1/5	
								L-Ser	D-Val	[c]	1/5	
7	25	6	L-Val	L-Lys	L-Asn	4/4	M	L-Val	L-Gln	L-Gln	5/8	M
								L-Val	L-Gln	D-Ala	2/8	
								L-Val	L-Gln	D-Ser	1/8	
8	10	3	L-Asn	D-Val	D-Ala	2/2	E	L-Val	L-Gln	L-Gln	2/2	G
9[d]	30	6	L-Asn	D-Val	L-Ala	2/11	P	D-Lys	D-Val	L-Val	2/5	P
			L-Asn	D-Val	D-Ala	2/11		D-Val	D-Val	D-Lys	1/5	
			D-Gln	L-Asn	D-Val	4/11		D-Ala	D-Val	D-Lys	1/5	
			D-Pro	L-Asn	D-Val	3/11		L-Val	L-Gln	L-Gln	1/5	
10	20	3	L-Asn	D-Val	D-Ala	2/5	E	N/A				
			L-Asn	D-Val	D-Lys	1/5						
			L-Asn	D-Val	D-Ser	1/5						
			L-Asn	D-Val	L-Ser	1/5						
11	25	6	L-Asn	D-Val	L-Val	2/6	M	D-Ala	D-Val	D-Lys	4/8	P[g]
			L-Ala	L-Val	D-Gln	4/6		D-Ser	D-Val	D-Lys	2/8	
								D-Lys	D-Val	D-Val	2/8	

[a] <1% NMP.
[b] E=Excellent (decoded beads were the only beads with red color. Other beads had no significant color) G=Good (decoded beads were the only beads with red color. 5–15% of other beads had very light orange color) M=Moderate (decoded beads were the only beads with red color. More than 15% of other beads had light to moderate orange color) P=Poor (>10% beads had red color).
[c] Unreadable code.
[d] A representative portion (50%) of the darkest beads were sampled. Many other red and orange beads present.
[e] Number of times this sequence was found/total number of darkest red beads.
[f] Beads from the deprotected library induced precipitation of this receptor.
[g] Due to poor selectivity, only 30% of the darkest beads were picked and decoded.
[h] Amino acid protecting groups: Ser(O*t*Bu), Lys(Boc), Gln(Trt), Asn(Trt); N/A=not assayed.

(L)Asn(trt)-(D)Val, with a preference for any of four amino acids—(D/L)Ala, (D)Val or (D)Ser(O*t*Bu)—in the third (C-terminus) position. The degree of selectivity was good to excellent, as judged by the small number of red beads and the near absence of color in the remaining, non-red beads at the equilibrium concentrations tested, 30–100 μM. The core of the receptor itself appeared to be responsible for binding since changing the site of the attachment of the dye (compound **2** vs. **3**) had little influence on the sequences selected. Furthermore, the all *S,S-trans* diastereomer **4** showed no binding at concentrations up to 100 μM. Thus, this new receptor motif, despite its low core molecular weight (<500), efficiently selected a family of four tripeptides out of a collection of 15^3=3,375.[21]

We had anticipated that the linkers of the bidentate receptors **5**–**8** would influence the sequence specificity of the receptor, if only for steric reasons. As shown in Table 1, with the exception of **7**, this was clearly not the case. Instead, the same (L)Asn(trt)-(D)Val dipeptide was selected, again with good to excellent selectivity, and with an almost identical preference for the same third (C-terminal) amino acid. Although we cannot rule out the possibility that **2** and **3** might bind with a stoichiometry of 2:1 receptor:tripeptide and thus mimic a bidentate form, it seems unlikely at micromolar concentrations. Instead, it is possible that conformers where the two receptor arms of **5** and **6** are aligned parallel (to form a cleft) are not favored and that the receptor arms act independently. Results with receptor **8** support this. Here, the tetrahedral sulfonamide linkages have been replaced by planar carboxamides, a change that molecular models indicated would greatly disfavor cleft formation, insuring that the arms would act independently. As expected, receptor **8** bound the same sequence as the single

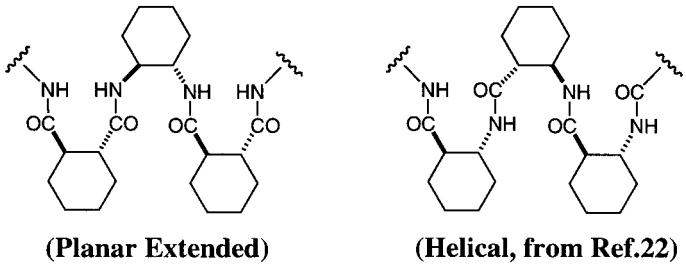


Figure 4.

arm receptors. Further evidence that the receptor arms in the bidentate receptor act independently was provided by the single arm receptor sulfonic acids, **9**–**11**. These receptors, although in general less selective, strongly favor the (L)Asn(trt)-(D)Val dipeptide. Compound **9** was unique in that it permitted a frame-shift of the dipeptide to the AA2-AA3 position. Nevertheless, the possibility that a tripeptide could be sandwiched in between the two arms simultaneously without changing the sequence specificity cannot be completely ruled out by these results and remains an interesting, if unlikely, possibility. Nor can we completely rule out the possibility that two arms from different bidentate receptors converge on a single peptide.

In the *meta* substituted disulfonamide **7**, models show that interaction between the arms in the unbound form is more favorable than in the other bidentate receptors. The aligned but partly intramolecularly hydrogen-bonded arms may present a very different binding surface and thus explain in part the selection by this compound of a quite different tripeptide sequence: (L)Val–(L)Lys(Boc)–(L)Asn(trt).

Side chain deprotected library

The affinity of the new receptor motif for (L)Asn(trt)-(D)Val was lost upon removal of the amino acid protecting groups, as shown in the results for the side-chain deprotected library in Table 1. Binding experiments revealed that in the deprotected library this sequence was replaced largely by the homochiral sequence (L)Val-(L)Gln, with a strong preference for a second (L)Gln in the third position. Once again, the preference of the core receptor **2** extended to the bidentate receptors, indicating that a functional cleft is either not formed or does not change the sequence specificity found for the isolated receptor arm. Contrary to the results with the protected library, this specificity even held for the *meta*-substituted disulfonamide **7**. In general, the sequence selectivity was slightly less impressive against the deprotected library as more beads had a higher background orange color than in the protected library under the same conditions.

The single arm receptor sulfonic acids **9**–**11** bear a single negative charge under the conditions of our assay. Although this proved to have little effect when tested against the protected library, we expected these receptors to show a preference for lysine-containing peptides when tested against the deprotected library. This was indeed found to be the case for the two sulfonic acid receptors tested, **9** and **11**. However, rather than binding lysine-containing peptides indiscriminately, these receptors only bound that amino acid in the specific sequence context of (D)Lys-(D)Val or (D)Val-(D)Lys. We thus conclude that in combination with a strong determinant of specificity such as the sulfonate anion, the core receptor's sequence specificity can be modulated. It should be noted that the (L)Val-(L)Gln-(L)Gln sequence appeared once with sulfonic acid **9**. It is possible that this sequence is still able to bind to the receptor arm but that the (D)Lys-(D)Val or (D)Val–(D)Lys dipeptides, whose beads were of a deeper red than those selected by **2**, **6**, **7** and **8**, bound more strongly and were picked and decoded preferentially.

The novel receptor motif we have designed and tested here was made from building blocks that resemble the *R*,*R*-*trans*-cycloalkyl β-amino acids recently employed in the construction of well-defined oligomeric helices, or foldamers (Fig. 4).[22,23] A comparison of the two structures offers an interesting example of how small changes at the monomer level can bring about large changes at the oligomer level. Whereas the foldamer researchers used homochiral β-amino acids, we used two different alternating monomers—an *R*,*R*-diacid (e.g. **12**) and an *S*,*S*-diamine (e.g. **16**). Both designs exploited the conformational restriction of the 1,2-*trans*-disubstituted cycloalkyl ring system to construct a relatively rigid, well-defined structure. In our design, the conformational restriction is enforced by the cycloalkyl rings and by the up–down alignment of the amide dipoles. Strong intramolecular hydrogen-bonds are prevented and the amide groups are left free to hydrogen-bond intermolecularly with a guest tripeptide, and to do so on either of two discrete receptor faces. The foldamers, on the other hand, offer an impressive example where intramolecular hydrogen-bonds are used to stabilize a designed secondary structure, leaving them in this case unavailable for molecular recognition. Taken together, the two examples demonstrate how a careful consideration of monomer stereochemistry, conformational restriction and hydrogen-bonding can lead to the successful design of oligomers having a defined shape and consequently well-suited to a particular goal—either molecular recognition or de novo secondary structure design.

As shown in Fig. 1, the receptor prototype was initially designed to be complementary to a peptide in an extended conformation. However, efforts to co-crystallize several receptor analogs with cognate tripeptides were not successful, and the poor solubility of the receptor core prevented detailed NMR studies. We have therefore not yet been able to ascertain whether or not the peptides are indeed bound in an extended conformation. The excellent sequence-selectivity, which may in fact argue against

specific binding to a conformation accessible to, and frequently sampled by, many sequences, was somewhat surprising given the low molecular weight of the receptor and the clear lack of a binding cleft at least in the single arm cases. Nevertheless, most receptor analogs displayed good to excellent selectivity and, within each library, very similar specificity. The single arm receptor sulfonic acids **9** and **11** comprised a special subset where it was shown that the specificity could be altered with only a small loss in selectivity by appending a functional group (here, sulfonate) possessing distinct amino acid side chain preference (unprotected lysine). This latter result is promising because it leads to the possibility of using this novel receptor motif in a modular, combinatorial approach in the search for receptors with tunable specificity for other tripeptide sequences.

Experimental

General methods

Unless otherwise indicated, all reagents and solvents were purchased from Aldrich or Fluka and were of the highest available purity and dryness. Thin layer chromatography was done on pre-coated Silica Gel 60 plates (EM Science), and solvent system A denotes: 8 benzene/8 chloroform/3 MeOH/0.3 water, vol/vol.

(1*R*,2*R*)-(−)-1,2-Cyclohexane dicarboxylic acid anhydride (12). The racemic diacid was purchased from Aldrich and resolved according to the literature procedure.[24] Briefly, the diacid was mixed with a molar equivalent of (*R*)-(+)-α-methylbenzylamine in hot *n*-propanol and repeatedly recrystallized from that solvent. After separation from the amine by extraction, the anhydride was formed in THF by the addition of 1.0 equiv. of dicyclohexyl carbodiimide (DCC) to the isolated free diacid.[25] A portion of the chiral diacid precursor used initially in these studies was provided as a gift from Merck–Schuchardt. The ^{1}H NMR and the specific rotation of the resolved material were identical to the sample provided by Merck–Schuchardt.

Dicarboxylic acid (1). To a stirred solution of **13** (29 mg, 0.25 mmol) and triethylamine (0.105 ml, 0.76 mmol) in DMF (0.4 ml) under argon was added **12** (78 mg, 0.51 mmol) in a mixture of DCM (0.4 ml) and DMF (0.4 ml). The resulting suspension was stirred overnight and the solvents were evaporated under reduced pressure. The residue was dissolved in saturated sodium bicarbonate solution and washed with chloroform. Concentrated HCl was added dropwise to precipitate the product, which was then washed with water and dried under vacuum to yield 65 mg of a white solid (61%). Crystals for X-ray analysis were prepared from DMSO by vapor diffusion with water.

(3*S*,4*S*)-Diazido-(*N*-(trimethylsilyl)ethoxycarbonyl)-pyrrolidine (15) (Ref. 17). To a stirred, chilled (−60°C) solution of **14** (0.8 g, 3.28 mmol) in THF (6.5 ml) was added dropwise a solution of Teoc-Cl (1.78 g, 9.9 mmol) over a period of 20 min. The reaction was allowed to warm to rt overnight. After 24 h the mixture was purified by silica gel flash chromatography (10% ethyl acetate/90% hexanes, vol/vol) to yield **15** (747 mg, 76%): IR (neat): ν_{max} 2955, 2895, 2102, 1694, 1428, 1351, 1250, 1178, 1110 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$, no TMS) δ [ppm] 4.19 (m, 2H), 3.96 (br s, 2H), 3.68 (m, 2H), 3.45 (m, 2H), 0.98 (m, 2H), 0.02 (s, 9H); Chemical ionization (CI) (NH_3)-LRMS, calcd for $C_{10}H_{19}N_7O_2Si$ (M+NH_3) 314.42, found 315.

(3*S*,4*S*)-Diamino-(*N*-(trimethylsilyl)ethoxycarbonyl)pyrrolidine (16). 37 mg of 10% Pd/C was suspended in an ethanol solution (1.5 ml) of **15** (397 mg, 1.33 mmol) under argon. The argon was replaced by a hydrogen balloon and the reaction was stirred for 18 h. The mixture was filtered through celite and purified by silica gel flash chromatography (49.5% methanol/49.5% ethyl acetate/1% triethylamine (TEA) to yield **16** (224 mg, 69%): IR (thin film): ν_{max} 3347 (br), 2952, 2896, 1679, 1434, 1359, 1249 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$, no TMS) δ [ppm] 4.15 (m, 2H), 3.70 (m, 2H), 3.04 (m, 4H), 1.49 (br s, 5.4H, including 4 amino H), 0.98 (m, 2H), 0.01 (s, 9H); CI(NH_3)-LRMS, calcd for $C_{10}H_{23}N_3O_2Si$, 245.39, found 246.

Pentafluorophenol ester (18). To a stirred solution of **12** (123 mg, 0.8 mmol) in dichloromethane (DCM, 2 ml) at 0°C was added dropwise a solution of **17** (346 mg, 0.8 mmol, 5 ml DCM) and TEA (0.11 ml, 0.8 mmol) over a period of 20 m. After 2 h the solvent was evaporated in vacuo and the crude product was purified by flash chromatography (gradient of 1.5–5% methanol/DCM/1 drop glacial acetic acid per 100 ml solvent). Fractions containing the product (R_f=0.4 in 5% MeOH/0.5% HOAc/DCM) were collected and evaporated in vacuo. The residue (453 mg, 0.76 mmol) was dissolved in DCM (8 ml) along with 139 mg pentafluorophenol (0.76 mmol). To this was added a 4 ml DCM solution of 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC, 160 mg, 0.83 mmol, 1.1 equiv.). After 4 h the reaction was evaporated in vacuo and purified by silica gel flash chromatography (3% acetone/DCM) to yield **18** (437 mg 73%): IR (thin film): ν_{max} 2931, 2861, 1780, 1650, 1600, 1518, 1338, 1134 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ [ppm] 8.33 (d, 2H), 7.90, (dd, 4H), 7.08 (d, 2H), 6.81 (dd, 4H), 5.48 (br t, 1H), 4.16 (m 2H), 3.82 (m, 2H), 3.61 (q, 2H), 3.46 (m, 2H), 3.06 (dt, 1H), 2.72 (ddd, 2H), 2.38 (dt, 1H), 2.29 (br d, 1H), 1.83 (br m, 3H), 1.49 (m, 4H), 1.29 (t, 3H); FAB-LRMS calcd for $C_{38}H_{36}F_5N_5O_6$ 753.71, found 754.

Teoc-protected one-arm receptor (2). A solution of **18** (216 mg, 0.29 mmol) in dry DMF (2 ml) was added dropwise under argon over a 2 min period to a DMF (0.4 ml) solution containing **16** (36.5 mg, 0.14 mmol) and diisopropylethyl amine (DIEA, 0.052 ml, 0.29 mmol). The reaction was followed by TLC (5% MeOH/95% chloroform/0.1% acetic acid), which, in addition to the product, showed formation of the intramolecularly cyclized imide product from **18** (R_f=0.9) and, initially, the monosubstituted product (R_f=0.1). After stirring for 36 h the precipitated product (R_f=0.4) was filtered off, washed with DMF and dried in vacuo to yield a dark red solid, **2** (150 mg, 75%). Flash chromatography of the filtrate residue yielded an additional 14 mg: ^{1}H NMR (400 MHz, DMSO-d_6) δ [ppm] 8.66 (d, 4H), 7.95 (m, 6H), 7.84 (d, 4H), 7.71 (br dd, 2H), 7.05 (m, 4H), 6.92 (d, 4H), 6.82 (d, 4H), 4.22 (br

dd, 4H), 3.95 (m, 2H), 3.90 (m, 2H), 3.83 (br dd, 4H), 3.59 (q, 4H), 3.46 (m 2H), 3.08 (br dd, 4H), 3.02 (m, 2H), 2.58 (m, 4H), 2.30 (br m, 4H), 1.54 (br s, 6H), 1.20 (t+m, 16H), 0.75 (dd, 2H), −0.12 (s, 9H); FAB-HRMS, calcd for $C_{74}H_{93}N_{13}O_{12}S$ 1383.6836, found 1383.6884.

Deprotected one-arm receptor (19). Trifluoroacetic acid (TFA, 0.2 ml) was added to a stirred suspension of **2** in DCM (2 ml). The reaction was followed by TLC (solvent system A, R_f=0.45). After 15 h the solvent (and excess TFA) was evaporated under reduced pressure to yield a dark red solid, **19**, TFA salt (134 mg, 95 %): IR (KBr): ν_{max} 3286, 2934, 2861, 1654, 1612, 1552, 1510, 1437, 1395, 1339 cm^{-1}; ^1H NMR (400 MHz, DMSO-d_6) δ [ppm] 8.78 (br m, 2H), 8.35 (d, 4H), 7.91 (d, 4H), 7.85 (d, 4H), 7.70 (dd, 2H), 7.05 (d, 4H), 6.95 (d, 4H), 6.82 (d, 4H), 4.22 (dd, 4H), 4.09 (m, 2H), 3.86 (m, 4H), 3.59 (q, 4H), 3.38 (br m, 2H), 3.09 (m, 4H), 2.82 (br m, 2H), 2.55 (m, 4H), 2.29 (br m, 4H), 1.68 (br d, 6H), 1.18 (t+m, 16H); FAB-LRMS, calcd for $C_{68}H_{81}N_{13}O_{10}$ (free base) 1240.5, found 1242.

General sulfonamide coupling procedure (e.g. 5 and 9). To a stirred solution of **19** (12.3 mg, 9 μmol) in *N*-methylpyrrolidinone (NMP, 0.4 ml) was added DIEA (5 μl, 27 μmol), followed by naphthyl-2,7-disulfonyl chloride (1.5 mg, 4.5 μmol) in NMP (20 μl). The reaction was followed by TLC (solvent system A). After 9 h, additional disulfonyl chloride was added (0.49 mg, 0.3 μmol, in 11 μl DMF). After stirring for a total of 45 h, enough silica gel to absorb all the reaction solvent was added. The entire reaction mixture was transferred to the top of a chloroform/silica gel flash chromatography column. The products were eluted with a 0–5% MeOH gradient in chloroform. (Note: The poor solubility of **19** and the products necessitated the use of the polar, high-boiling solvent NMP. In some cases the NMP co-eluted with the purified product and could not be removed, e.g. **5**. The sulfonyl chlorides also slowly decomposed in this solvent.)

*Receptor (**5**):* 2.2 mg (including an unknown amount of NMP) R_f=0.6 (10% MeOH/chloroform); FAB-LRMS, calcd for $C_{146}H_{166}N_{26}O_{24}S_2$ 2733.2, found 2734.

*Receptor (**9**):* 5.5 mg (81%), R_f=0.1 (10% MeOH/chloroform); FAB-LRMS, calcd for $C_{78}H_{87}N_{13}O_{15}S_2$ 1510.74, found, 1512.

*Receptor (**6**):* 5.0 mg (51%), R_f=0.74 (10% MeOH/DCM); MALDI-TOF MS, calcd for $C_{142}H_{164}N_{26}O_{24}S_2$ 2683.11, found 2706.99 (M+Na).

*Receptor (**10**):* 2.1 mg (38%), R_f=0.4 (10% MeOH/DCM); MALDI-TOF MS, calcd for $C_{74}H_{85}N_{13}O_{15}S_2$ 1459.57, found 1497.67 (M+K).

*Receptor (**7**):* 5.4 mg (including an unknown amount of NMP) R_f=0.75 (eluted first in 5% MeOH/chloroform then in solvent system A: FAB-LRMS, calcd for $C_{142}H_{164}N_{26}O_{24}S_2$ 2683.1, found 2684.

*Receptor (**11**):* 6.2 mg (67%): R_f=0.35 (eluted as in **7**); FAB-HRMS, calcd for $C_{74}H_{86}N_{13}O_{15}S_2$ (M+1) 1460.5808, found 1460.5769.

*Receptor (**8**).* Made similarly, but the dipentafluorophenol ester of benzene-1,3-dicarboxylic acid was used in place of the aromatic disulfonyl chlorides. 8.7 mg (90 %): FAB-LRMS, calcd for $C_{144}H_{164}N_{26}O_{22}$ 2611.0, found 2612.

(3*S*,4*S*)-*N*-Boc-3,4-diazidopyrrolidine (21). To a solution of (3*R*,4*R*)-*N*-Boc-pyrrolidinediol dimesylate **20** (8.60 g, 23.93 mmol) in DMF (200 ml) was added NaN_3 (16.8 g, 0.258 mol). The mixture was heated to 90°C for 24 h, then evaporated to dryness in vacuo. The residue was diluted with EtOAc (100 ml) and washed with water (3×20 ml). The aqueous portion was back-extracted with EtOAc (1×20 ml). The organic portions were combined, washed with brine (1×10 ml), dried ($MgSO_4$) and concentrated. The residue was filtered through a plug of silica gel (20% EtOAc/hexanes) and concentrated to give **21** as a yellow oil (4.12 g, 68%). TLC R_f=0.81 (25% EtOAc/hexanes); IR (film) ν_{max} 2970, 2097, 1695 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) d 3.94 (m, 2H), 3.65 (m, 2H), 3.34 (m, 2H), 1.45 (s, 9H); LRMS (CI) Calcd for $C_9H_{15}N_7O_2$ 253.26, found 254.

(3*S*,4*S*)-3,4-Diazidopyrrolidine, TFA salt (22). To a stirred solution of **21** (1.00 g, 3.95 mmol) in DCM (25 ml) was added trifluoroacetic acid (2.5 ml). After TLC indicated complete conversion, the solvent and excess TFA were removed in vacuo, to yield **21** (1.05 g, quant.), which was used without further purification. TLC R_f=0.4 (10% MeOH/DCM); ^1H NMR (400 MHz, DMSO-d_6) d 9.43 (br s, 2H), 4.53 (dd, 2H), 3.39 (dd, 2H), 3.23 (dt, 2H).

Disperse red-pyrrolidine diazide (23). To a solution of disperse red (300 mg, 0.95 mmol) in THF (15 ml) was added 1,1′-carbonyldiimidazole (189 mg, 1.15 mmol). After 6 h a solution of **22** (TFA salt, 400 mg, 1.50 mmol) and diisopropylethylamine (0.30 ml, 1.72 mmol) in THF (2 ml) was added. The mixture was brought to reflux overnight. The crude mixture was evaporated to dryness in vacuo and purified by silica gel chromatography (40% EtOAc/hexanes) to give **23** as a red powder (461 mg, 0.934 mmol, 98%]. TLC R_f=0.13 (25% EtOAc/hexanes); IR (KBr) ν_{max} 2100, 1694, 1604, 1516, 1427, 1390, 1338, 1142, 1106 cm^{-1}; ^1H NMR (400 MHz, DMSO-d_6) d 8.34 (d, 2H), 7.91 (d, 2H), 7.83 (d, 2H), 6.91 (d, 2H), 4.30 (m, 2H), 4.21 (t, 2H), 3.71 (t, 2H), 3.50 (m, 4H), 3.24 (m, 2H), 1.15 (t, 3H); LRMS (CI) calcd for $C_{21}H_{23}N_{11}O_4$ 493.48, found 494.

Disperse red-pyrrolidine diamine (24). A solution of **23** (40 mg, 0.08 mmol) and triphenylphosphine (50 mg, 0.19 mmol) in toluene (1.5 ml) was brought to reflux. After 10 min, the reaction mixture was cooled to 50°C, and a solution of water (0.025 ml) in THF (0.3 ml) was added. Reflux was resumed and continued overnight. After evaporation of the solvents the crude product was purified by flash chromatography (1% Et_3N/10% MeOH/DCM) to yield a red solid (34 mg, 0.077 mmol, 95%). TLC R_f=0.29 (1% Et_3N/10% MeOH/DCM); IR (KBr) ν_{max} 1509, 1438, 1193, 1120, 858, 785, 722, 695, 651 cm^{-1}; ^1H NMR (400 MHz, DMSO-d_6) d 8.35 (d, 2H), 7.93 (d, 2H), 7.84 (d, 2H), 6.93 (d, 2H), 4.17 (t, 2H), 3.70 (t, 2H), 3.52 (m, 4H), 2.93 (m, 4H), 1.16 (t, 3H); LRMS (CI) calcd for $C_{21}H_{27}N_7O_4$ 441.48, found 442.

(1*R*,2*R*)-Cyclohexanedicarboxylic acid mono-*n*-propylamide (25). To a stirred solution of **12** (150 mg, 0.97 mmol) in DCM (10 ml) was added dropwise *n*-propylamine (0.25 ml, 3.02 mmol). A white precipitate soon appeared and the reaction mixture was stirred for 2 h, followed by

acidification with 1 M HCl (10 ml). The layers were separated, and the aqueous phase was extracted with EtOAc (3×5 ml). The organic portions were combined, washed with brine (1×5 ml), dried ($MgSO_4$), filtered, and concentrated to give **25** as an off-white solid (183 mg, 0.86 mmol, 88%). TLC R_f=0.5 (75% EtOAc/hexanes); IR (KBr) ν_{max} 3309, 2936, 1694, 1645, 1549, 1276, 1261, 921, 681 cm^{-1}; ^1H NMR (400 MHz, DMSO-d_6) d 11.86 (br s, 1H), 7.69 (t, 1H), 2.93 (m, 2H), 2.41 (td, 1H), 2.28 (td, 1H), 1.91 (m, 1H), 1.76 (m, 1H), 1.68 (m, 2H), 1.34 (m, 2H), 1.20 (m, 4H), 0.80 (t, 3H); LRMS (CI) calcd for $C_{11}H_{19}NO_3$ 213.27, found 214.

(1*S*,2*S*)-Cyclohexane-1,2-dicarboxylic acid mono-*n*-propylamide (26). Prepared in a manner analogous to **25** from the enantiomer of **12**.

Acid fluoride (27). To a stirred solution of (1*R*,2*R*)-cyclohexanedicarboxylic acid mono-*n*-propylamide **25** (50 mg, 0.23 mmol) and pyridine (23 ml, 0.28 mmol) in DCM (2 ml) was added dropwise to a solution of cyanuric fluoride (63 mg, 0.47 mmol) in DCM (1 ml). A white precipitate formed during the addition. The resulting suspension was stirred for 1.5 h, diluted with DCM (3 ml), washed with water (2×1 ml), dried ($MgSO_4$), filtered and concentrated to afford the acid fluoride as a white crystalline solid (49 mg, 0.23 mmol, 97%). This material was used immediately in the subsequent step. Acid fluoride **28** was prepared analogously.

Receptor (3). The acid fluoride **27** was added to a stirred solution of diamine **24** (10 mg, 0.023 mmol) and diisopropylethylamine (12 ml, 0.070 mmol) in DMF (0.5 ml). While stirring overnight, a dark red precipitate formed. TLC showed a new red-colored band with R_f=0.4 (10% MeOH/DCM). The mixture was evaporated to dryness and purified by preparative TLC (5% MeOH/$CHCl_3$) to yield a red solid (5 mg, 6 mmol, 26%): FAB-HRMS, calcd for $C_{43}H_{61}N_9O_8K$ 870.4280, found 870.4265 (M+K).

Receptor (4). Made in a manner analogous to **3**, from diamine **24** and acid fluoride **28**. Red solid (7 mg, 8 mmol, 37%); TLC R_f=0.61 (10% MeOH/DCM); FAB-HRMS, calcd for $C_{43}H_{61}N_9O_8K$ 870.4280, found 870.4286 (M+K).

Binding studies

Stock solutions of the receptors were made in neat *N*-methylpyrrolidinone (NMP) using gentle to moderate heating and/or sonication where necessary. A weight of approximately 3 or 6 copies of the library was shaken in a receptor solution of 5% NMP/chloroform buffered with 0.5 mM trioctylamine/0.25 mM trifluoroacetic acid at room temperature. After equilibration for at least 48 h the beads were inspected for accumulation of the red receptor-dye conjugate and the concentration was adjusted until, upon further equilibration, a small group of beads (usually 3–12) were stained red or dark orange. Equilibration was allowed to continue for at least 12 h after the number of red beads appeared constant. Unless otherwise noted in Table 1, all of the darkest beads were then picked and decoded by capillary GC as previously described.[26] Using this procedure, the amount of receptor held in the beads at equilibrium is insignificant. The final concentration of the receptor was estimated by taking an absorbance measurement (chloroform, 476 nm, ext. coeff.≈35,938 per chromophore) of the solution after diluting the NMP to 0.5%, vol/vol.

Acknowledgements

This work was supported by Grant CHE97-07870 from the National Science Foundation. Fellowships from the National Institutes of Health (from the NIAID to K. R. and from the Medical Scientist Training Program to L. J. G.) are gratefully acknowledged. We thank Prof. Gerard Parkin and Dr Tony Hascall for performing the X-ray crystallographic analysis. We are grateful to Vinka Parmakovich, Barbara Sporer and Yasuhiro Itagaki of the Columbia University Mass Spectroscopy Facility. We also thank M. David Weingarten for helpful discussions.

References

1. Cram, D. J. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1009.
2. Still, W. C. *Acc. Chem. Res.* **1996**, *29* (3), 155.
3. Hossain, M. A.; Schneider, H. J. *J. Am. Chem. Soc.* **1998**, *120* (43), 11208.
4. Lowik, D. P. W. M.; Mulders, S. J. E.; Cheng, Y.; Shao, Y.; Liskamp, R. M. J. *Tetrahedron Lett.* **1996**, *37*, 8253.
5. Torneiro, M.; Still, W. C. *J. Am. Chem. Soc.* **1995**, *117*, 5887.
6. Davies, M.; Bonnat, M.; Guillier, F.; Kilburn, J. D.; Bradley, M. *J. Org. Chem.* **1998**, *63*, 8696.
7. Carrasco, M. R.; Still, W. C. *Chem. Biol.* **1995**, *2*, 205.
8. Maletic, M.; Wennemers, H.; McDonald, D. Q.; Breslow, R.; Still, W. C. *Angew. Chem., Int. Ed. Engl.* **1996**, *35* (13/14), 1490.
9. Wennemers, H.; Yoon, S. S.; Still, W. C. *J. Org. Chem* **1995**, *60*, 1108.
10. Shao, Y. F.; Still, W. C. *J. Org. Chem.* **1996**, *61* (18), 6086.
11. Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1.
12. Macromodel® V 6.0 Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caulfield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440. Monte Carlo conformational searches were performed in chloroform using default settings. Structures were minimized using the MM2 forcefield.
13. Loll, P. J.; Bevivino, A. E.; Korty, B. D.; Axelsen, P. H. *J. Am. Chem. Soc.* **1997**, *119*, 1516.
14. Beddoes, R. L.; Dalton, L.; Joule, J. A.; Mills, O. S.; Street, J. D.; Watt, C. I. F. *J. Chem. Soc., Perkin Trans II* **1986**, 787.
15. Reddy, D. R.; Thornton, E. R. *J. Chem. Soc., Chem. Commun.* **1992**, 172.
16. Nagel, U. *Angew. Chem., Int. Ed. Engl.* **1984**, *23* (6), 435.
17. Campbell, A. L.; Pilipauskas, D. R.; Khanna, I. K.; Rhodes, R. A. *Tetrahedron Lett.* **1987**, *28* (21), 2331.
18. O'Malley, G. J.; Spahl, B.; Cherill, R. J.; Kosley, R. W. *J. Org. Chem.* **1990**, *55*, 1102.
19. Naphthalene-2,7-disulfonyl chloride from the disodium salt: Muth, F. *Houbeyn-Weyll* **1955**, 565.
20. Benzene-1,4-disulfonyl chloride from the sodium salt using oxalyl chloride/DMF, via diazotization of the 4-amino benzenesulfonic acid with SO_2: Meerwein, H.; Dittmar, G.; Gollner, R.; Hafner, K.; Mensch, F.; Steinfort, O. *Chem. Ber.* **1957**, *90* (6), 851.

21. Yoon, S. S. Y.; Still, W. C. *Tetrahedron* **1995**, *51* (2), 567 and references therein. The library contained 153 potential sequences of the structure: Ac-AA3-AA2-AA1-HN-polystyrene, where AAn represents Gly or the D or L isomers of: Ala, Ser, Val, Pro, Asn, Gln, or Lys.
22. Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1996**, *118* (51), 13071.
23. Appella, D. H.; Christianson, L. A.; Klein, D. A.; Gellman, S. H. *Nature* **1997**, *387*, 381.
24. French Patent, Rhone-Poulenc, 1974, No. 74 06885.
25. Holzapfel, C. W.; Pettit, G. R. *J. Org. Chem.* **1984**, *50*, 2323.
26. Ohlmeyer, M. H. J.; Swanson, R. N.; Dillard, L. W.; Reader, J. C.; Asouline, G.; Kobayashi, R.; Wigler, M.; Still, W. C. *Proc. Natl. Acad. Sci., USA* **1993**, *90*, 10922.