Pergamon

Tetrahedron Letters, Vol. 35, No. 46, pp. 8557-8560, 1994
Elsevier Science Ltd
Printed in Great Britain
0040-4039/94 $7.00+0.00

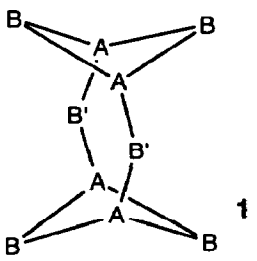
0040-4039(94)01894-4

Cyclooligomeric Receptors for the Sequence Selective Binding of Peptides. A Tetrahedral Receptor from Trimesic Acid and 1,2-Diamines.

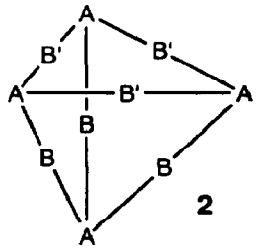
Seung Soo Yoon and W. Clark Still
Department of Chemistry, Columbia University, New York NY 10027 USA

Abstract: The first members of a new family of tetrahedral receptors (**2**) have been prepared from trimesic acid and certain chiral 1,2-diamines. A survey of the binding properties of one of these with a 50,000-member peptide library finds high selectivity for N-acetylated terminal residues (particularly Ac-(D)Gln) and the dipeptide sequence (L)Gln-(L)Pro.

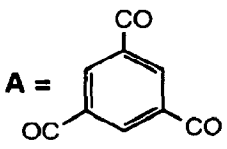
One of the most exciting recent advances in molecular recognition is the development of synthetic receptors or host molecules that selectively bind biological substrates.[1] Among such receptors, even simple designs can result in highly selective binding properties when they incorporate a conformationally well-defined binding site and functionality that is complementary to that of the substrate. In recent reports from this laboratory, we described a number of D_2-symmetric receptors (**1**) that are simple A_4B_6 cyclooligomers of trimesic acid (**A**$(OH)_3$) and 1,2-diamines (e.g. **B**H_2, **B'**H_2).[2] Such receptors are simple to prepare and were found to have highly enantio- and sequence-selective binding properties for certain amino acids and peptides. Here we describe receptor **2**, a new A_4B_6 cyclooligomer having tetrahedral symmetry (when **B'**=**B**). Though **2** and **1** have very different shapes, both have large, open binding cavities and unassociated arrays of hydrogen bond donors and acceptors. Like **1**, **2** binds peptides with high, but different, sequence selectivity in chloroform.

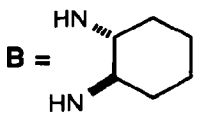


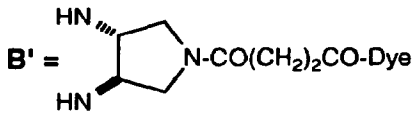
1



2

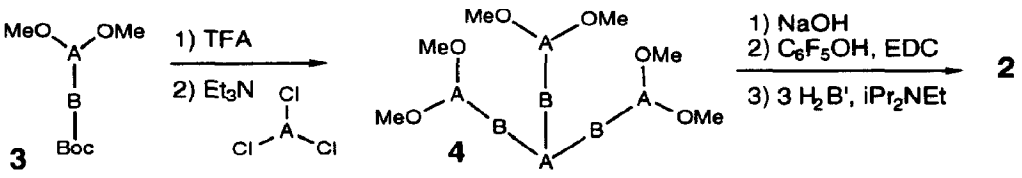






Receptor **2** is a new member of the growing family of cyclooligomeric host molecules based on trimesic acid and chiral 1,2-diamines.[3] Unlike our previously described A_4B_6 receptors, **2** has approximately tetrahedral symmetry with similar, if not equivalent, binding cavities on each tetrahedral face. We were particularly interesting in its topologically unusual binding sites that appeared more like large, interconnected pores than the more common, cup-like binding depressions or clefts. Such pore-like binding sites might create unusual peptide-binding opportunities as they appear large enough to allow a peptide substrate to thread itself through the receptor.

The synthesis of **2** is simple. First, the dimethyl ester of **A** is coupled (DCC) with mono-Boc (1*R*,2*R*)**B** to give **3**. Three equivalents of deprotected **3** are then coupled with one equivalent of 1,3,5-benzenetricarbonyl chloride (Cl_3**A**) to provide **4**. Next, the methyl esters are replaced by pentafluorophenyl esters and treated with three equivalents of the diamine of **B'** at high dilution (final [**2**] = 0.2 mM). This final macrotricyclization provides **2** by simultaneously closing three 27-membered rings in 51% yield after silica gel chromatography. For the dye, we use a brilliant red azo dye, Disperse Red 1. An analogous tricyclization of **4** with H_2**B** can be used to prepare fully tetrahedral **2**(**B'** = **B**) in 50% yield.



To survey **2**'s binding properties, we used our previously described solid phase color assay employing a substrate library of ~50,000 (maximally 15^4 = 50,625) terminally acylated, sidechain-protected tripeptides having the general structure:[3,4]

R(C=O)-AA3-AA2-AA1-NH$(CH_2)_5$CONH-polystyrene

The library was prepared on 50-80 μm polystyrene beads by encoded split synthesis[5] and each library bead carried only one type of acylated tripeptide substrate.

Binding surveys were performed by equilibrating **2** with the substrate library for 48 hours. Binding was indicated when a library bead concentrated the color of the receptor-tethered red dye from a ~50 μM solution of **2** in $CHCl_3$. By measuring **2**'s concentration in free solution at equilibrium and its loading on the selected beads, we

could estimate that the most deeply stained beads bound **2** with association constants of $\geq 10^4$. Fifty-five of these deep red beads were picked manually under a 4X microscope and the structures of their associated acylated tripeptide substrates were determined by ECGC decoding. The results are summarized below as the most frequently found residues at each substrate position:

Table 1. Frequencies of Residues Selectively Bound by Receptor **2**.

R	AA3	AA2	AA1
Me (45%)	(D)Gln (29%)	(L)Gln (59%)	(L)Pro (57%)
Me_2N (10%)	(D)Ser (10%)	(L)Ser (10%)	(D)Pro (12%)
			Gly (10%)

Because fifteen different residues were used at each site, simple statistics would predict that 7% of each different residue would be found at each site for unselective binding. Though several different residues were found with frequencies greater than that, only single residues at each site emerged from the study with statistical significances exceeding 90%. In fact, these most frequently found residues (Me, (D)Gln, (L)Gln and (L)Pro) are each selectively bound by **2** with a statistical significance ≥99.8%.[6] It is interesting that R was most commonly methyl (the smallest R in the library) and that slightly larger ethyl was not found at all. This finding suggests a well-defined binding site for R, probably one deeply buried within **2**'s interior. The binding selectivies of **2** are very different from those of **1**: previous studies[2c] established that **1** had virtually no selectivity for R but preferentially bound AA3-AA1 = (D)Asn, (L)Val, (L)Ser.

In addition to having significant selectivity at each position of our 4-residue substrates, receptor **2** preferentially binds these residues as two well-distinguished, independent two-residue sequences. Thus, R was found to be Me in 80% of the decoded beads having AA3 = (D)Gln. In all, 22% of the beads decoded had R,AA3 = Me,(D)Gln. Two-residue sequence selectivity at AA2 and AA1 was even stronger in that 58% of the decoded beads had AA2,AA1 = (L)Gln,(L)Pro. Furthermore, every decoded bead having AA1 = (L)Pro had AA2 = (L)Gln. Interestingly, no decoded beads were found with AA3,AA2 = (L)Gln,(L)Pro. The binding data also indicated that these two-residue sequences were selected by **2** independently as they are found as two distinct sets of bound peptides. Indeed, only one of the fifty-five decoded beads had the full consensus sequence Me, (D)Gln, (L)Gln, (L)Pro.

To verify and quantify the binding results described above, we resynthesized larger quantities of two of the preferred substrate sequences that were found with our solid phase color assay as well as several other comparison substrates that were not.

Table 2. Binding Energies of Resynthesized Tripeptides for **2** in $CHCl_3$.

Entry	R	AA3	AA2	AA1	$\Delta G_{binding}$	Found in Assay?
1	Me	(D)Gln	(L)Ala	Gly	-5.2 kcal/mol	yes
2	Me	(L)Gln	(L)Ala	Gly	-4.1 kcal/mol	no
3	CPen	(D)Gln	(L)Ala	Gly	-4.0 kcal/mol	no
4	CPen	(D)Gln	(L)Gln	(L)Pro	-5.1 kcal/mol	yes
5	CPen	(D)Gln	(D)Gln	(L)Pro	-3.8 kcal/mol	no
6	CPen	(D)Gln	(L)Gln	(D)Pro	-3.6 kcal/mol	no
7	CPen	(D)Gln	(D)Gln	(D)Pro	-3.0 kcal/mol	no

We then used UV to measure the binding energies each of these substrate sequences for **2**. The results are summarized above in Table 2. These energies confirm the findings of the solid phase color assay in that the two sequences (entries 1 and 4) found by the assay had the tighest binding to **2** among all sequences examined. The findings also demonstrate that our solid phase assay is able to distinguish binding energy differences as small as 1 kcal/mol.

Thus we find that receptor **2** binds simple peptides with high residue-, sequence- and stereoselectivity. The results also show that there is nothing special about the particular topology of our previously described receptor **1** so far as general peptide binding is concerned. Instead, the peptide-binding properties we find may stem simply from a well-defined, conformationally stable binding cavity and proximate, unassociated hydrogen bond donors and acceptors - structural features common to both **1** and **2**. Selectivity among different peptides, however, depends sensitively on the details of receptor structure.[7]

Notes and References.

1. Reviews: H.-J. Schneider, *Angew. Chem. Int. Ed.*, **32**, 848 (1993); T.H. Webb & C.S. Wilcox, *Chem. Soc. Rev.*, 383 (1993).
2. a. S.S. Yoon & W.C. Still, *J. Am. Chem. Soc.*, **115**, 823 (1993); b. S.S. Yoon & W.C. Still, *Tetrahedron Lett.*, **35**, 2117 (1994); c. S.S. Yoon & W.C. Still, *Tetrahedron*, in press.
3. A. Borchardt & W.C. Still, *J. Am. Chem. Soc.*, **116**, 373 (1994)
4. All possible combinations of *R* = methyl (Me), ethyl (Et), isopropyl (iPr), t-butyl (tBu), neopentyl (neoPe), trifluoromethyl (CF_3), isobutyl (iBu), methoxymethyl (MOM), acetoxymethyl (AcOM), cyclopropyl (cPr), cyclobutyl (cBu), cyclopentyl (cPe), phenyl (Ph), morpholino (Morph), dimethylamino (Me_2N) and *AA1-AA3* = Gly, D-Ala, L-Ala, D-Ser(OtBu), L-Ser(OtBu), D-Val, L-Val, D-Pro, L-Pro, D-Asn(N-trityl), L-Asn(N-trityl), D-Gln(N-trityl), L-Gln(N-trityl), D-Lys(N-Boc), L-Lys(N-Boc).
5. M.H.J. Ohlmeyer, R.N. Swanson, L.W. Dillard, J.C. Reader, G. Asouline, R. Kobayashi, M. Wigler & W.C. Still, *Proc. Natl. Acad. Sci. USA*, **90**, 10922 (1993).
6. This measure of statistical significance is the probability that the observed frequency of residue occurrence is not simply due to statistical fluctuation.
7. We acknowledge the support of NSF CHE92 08254.

(*Received in USA* 26 *August* 1994; *accepted* 21 *September* 1994)