Pergamon

Bioorganic & Medicinal Chemistry Letters 9 (1999) 2673–2678

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

CONSTRUCTION OF SEQUENCE-SELECTIVE PEPTIDE RECEPTORS FROM CONFORMATIONALLY RESTRICTED *ETA*- AND *THETA*- AMINO ACIDS

Kevin Ryan and W. Clark Still*

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

Received 8 June 1999; accepted 10 August 1999

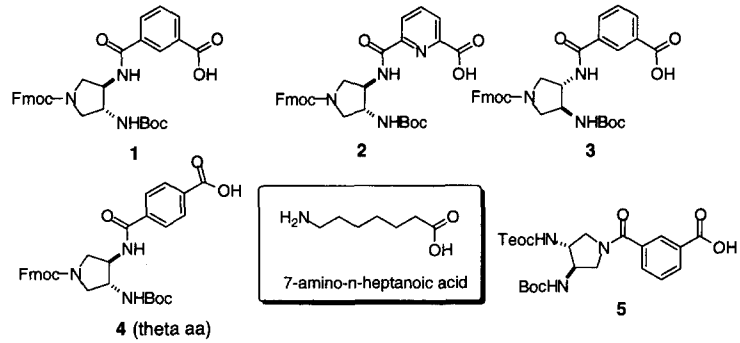
Abstract: Oligomeric chemoreceptors for tripeptides have been constructed from a set of novel *eta*- and *theta*-amino acids. Screening against an encoded combinatorial tripeptide library has revealed two new small molecule chemoreceptor motifs with highly sequence-selective binding properties.

The design of chemoreceptors for sequence-specific peptide recognition is often guided by the principle that a semi-rigid or preorganized[1] molecular cleft is needed to effect selective host-guest binding.[2] While this principle has led to much success, the difficulty in generating a large variety of rationally designed cleft motifs for binding a broad array of substances prompted us to explore a different approach. Rather than design an entire receptor we chose to make a collection of conformationally restricted bifunctional monomers which could be linked together into a variety of oligomers. Larger molecules made from constrained monomers should themselves be constrained into relatively well-defined shapes, which we hoped would have steric and electronic complementarity to certain peptide sequences. Furthermore, we reasoned that if a group of small oligomeric peptide-binding motifs could be identified, it would be possible to use these and related compounds to make still larger oligomers to target longer peptide sequences, perhaps by using a block synthesis combinatorial approach. As a first step toward this goal we have designed a set of novel conformationally restricted *eta*- and *theta*-amino acids. Synthesis of homo-and hetero-oligomers from this initial set has led to the discovery of two surprisingly small compounds that exhibit sequence-selective tripeptide binding.

Our initial strategy was to take fragments from previously successful receptors[2] and to fashion them into conformationally restricted amino acids in a form suitable for coupling (Figure 1). In comparison to the unrestricted 7-amino-n-heptanoic acid, each bond in the monomer amino acids is restricted in some way, either by (i) incorporation into a ring; (ii) resonance between a carbonyl and an aromatic ring; (iii) amide bond formation; or (iv) by the possible dipole-dipole alignment of the amide and carbamate carbonyls. Although the conformational restriction of the monomers is extensive, it does allow for some adaptability to the surface of a potentially complementary peptide. The conformational rigidity is also intended to minimize the formation of self- and nearest neighbor intramolecular hydrogen bonds that would have to be broken before hydrogen bonding to a peptide target. Since our design is limited to the monomer level we can not rule out the possibility that non-nearest

0960-894X/99/$ - see front matter
PII: S0960-894X(99)00463-1

neighbor hydrogen bonds may form. While hydrogen bonding of this type can favor interesting secondary structures[3] we have tried to avoid it with receptors like ours where intermolecular hydrogen bonding is an important component in the recognition and binding of a target molecule. Note also that each amino acid possesses two orthogonally protected amino groups: one for main-chain extension and one as a potential site for acylation, leaving open the possibility of attaching various side-chains to increase oligomer diversity. In the work presented here the pyrrolidine ring nitrogen was used exclusively for chain elongation and the Boc group was left as is.



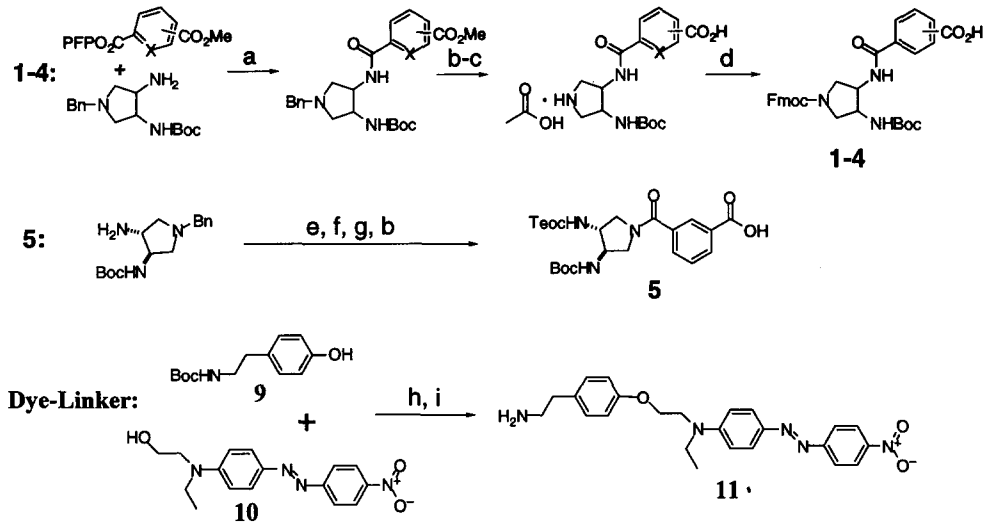
Oligomers: **6**: Fmoc-$(\mathbf{1})_5$-linker-dye
7: Fmoc-$(\mathbf{2})_5$-linker-dye
8: Fmoc-$(\mathbf{1})_2$-**Leu**-$(\mathbf{1})_2$-linker-dye
9: Fmoc-**2-5-4-Leu-3**-linker-dye
10: Teoc-----**5-4-Leu-3**-linker-dye
11: Ac--------**4-Leu-3**-linker-dye
12: Fmoc-**2-5-4**----------linker-dye
13: Fmoc----**5-4**----------linker-dye
14: Fmoc-------**4**----------linker-dye

Figure 1. List of monomers and oligomeric chemoreceptors

Synthesis of the monomer pool is summarized in Scheme 1.[4] Compounds **1–4** were made by a common route, beginning with coupling of the desired benzyl/BOC protected 1,2-*trans* *R,R*- or *S,S*-diaminopyrrolidine[5] with the monomethyl ester-protected, pentafluorphenol (PFP) ester-activated form of the commercially available dicarboxylic acids. The internal amide linkage of compound **5** is reversed relative to those of **1–4** and thus required the alternate route shown.

Oligomers **6–14** were synthesized in the C- to -N orientation beginning with a dye-linker conjugate of the C-terminal amino acid (Scheme 1).[6] Binding assays[7] of the oligomers were conducted with an encoded combinatorial tripeptide library prepared on 50 μm polystyrene beads.[8] The results for the largest oligomers (pentamers **6–9,** Table 1) show that oligomers **6** and **9** are capable of sequence selective peptide binding in chloroform at concentrations of less than 60 μM. As a first step towards understanding the origin of this

property, we set out to determine exactly which portion of each molecule is responsible for peptide binding. In the case of **6**, where the receptor portion of the molecule is homogeneous, we simply studied binding as a function of oligomer length in the series: Fmoc-(**1**)$_n$-dye-linker, where n = 1–5. Side by side comparisons showed that the relative binding strength, judged by depth of bead color at a single concentration, increases with oligomer length, and that the dimer (n = 2) is the minimum length which leads to binding (>75 μM). This finding was confirmed by oligomer **8**, where a middle leucine breaks the receptor into two dimers, which bound the same



Scheme 1. (a) X = CH, N; CH_2Cl_2, Hunig's base, rt, 66–89%; (b) MeOH, THF, excess NaOH, rt, 12 h; 5 °C conc HCl to pH 2–3, 73–96%; (c) H_2/Pd(OH)$_2$/C, 1 equiv HOAc, EtOH, 18 h, 50–87%; (d) Na_2CO_3 (pH 9–10), dioxane, 1.1 equiv Fmoc-OSuc; 5 °C, 1 N HCl to pH 2–3, 64-87 %. (e) 2-(TMS)ethyl *p*-nitrophenyl carbonate, Na_2CO_3, EtOH/Water, 1 h reflux, rt, O.N. 87%; (f) as in (c) then 5% Na_2CO_3 wash, 91%; (g) Benzene 1,3–dicarboxylic acid monomethyl ester, CH_2Cl_2, 1.1 equiv EDC, 65%; (h) DEAD/Ph_3P 60%; (i) 1 TFA:9 CH_2Cl_2 v/v, 10% Na_2CO_3, 98%.

family of tripeptides. The sequence preference did not change with length: the same family of three sequences shown in Table 1 was always selected, albeit in varying ratios (likely due to the statistical variability among library samples). The sequence selectivity for this series was very good with fewer than 5% of the unpicked beads showing a weaker light orange color. The concentration threshold for detectable binding of pentamer **5**, the strongest binder in the series, was 14 μM.

When investigating the propensity of receptor building blocks to form oligomers with desirable properties, it is as important to identify those compounds which are frequently found in successful oligomers as it is to identify those that are not, especially when planning the synthesis of a combinatorial library of the monomers or, in our case, their short oligomers. Among those amino acids in our original set we found that monomer **2** was notable for its tendency never to be found in a receptor motif of even poor selectivity. A striking example of this trend can be seen in a comparison of oligomers **6** and **7**. In **7**, the benzene rings of **6**, from monomer **1**, have

been replaced by the pyridine rings of monomer **2**. This small change completely abrogated the tripeptide binding ability of **6** (Table 1).

Table 1. Pentamer Receptor Binding Results [a]

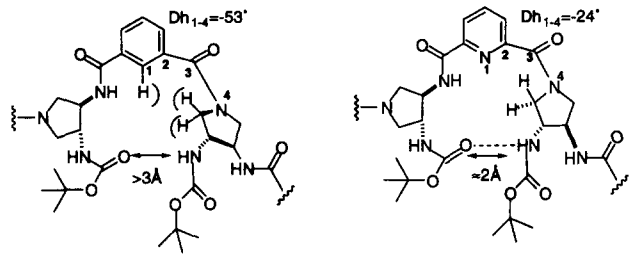
Com.	μM	Sequences Found--Protected Side Chain Library (% of stained beads)	Sequences Found--Deprotected Side Chain Library (% of stained beads)
6	33	(L)Val-(L)Pro-(L)Ala70% (D)Pro-(L)Ala-(L)Pro 15% (D)Pro-(L)Lys(Boc)-(L)Pro 15%	(D)Pro-(L)Ala-(L)Pro 80% (L)Val-(L)Pro-(L)Ala20%
7	780	No binding	No binding
8	35	(D)Pro-(L)Lys(Boc)-(L)Pro 71% (D)Pro-(L)Ala-(L)Pro 29%	Not assayed
9	55	(D)Val-(D)Asn(Trt)-(D)Lys(Boc) 30% (D)Lys(Boc)-(D)Asn(Trt)-(D)Lys(Boc) 70%	No Binding

[a]Assays were carried out in duplicate or triplicate. The results shown are from a single representative experiment.

In an effort to understand this we studied two-residue models of both homo-oligomers **6** and **7** by subjecting them to Monte Carlo conformational searching with energy minimization.[9] All lowest energy conformers generated from monomer **2** were stabilized by an interresidue hydrogen bond between the Boc carbonyl of one residue and the BocNH proton of the succeeding residue (Figure 2). In the case of the benzene analog **1**, however, the C-2 hydrogen of the phenyl ring appeared to force the pyrrolidine ring to twist out of the plane of the aromatic ring, as measured by the dihedral angle $Dh_{1\text{-}4}$ (Figure 2). This resulted in a significantly larger distance between the Boc groups of successive residues, thereby preventing the nearest neighbor hydrogen bond seen in the pyridine compound. Otherwise, the lowest energy conformations of the two-mer of **1** were very similar to those of the two-mer of **2**. Based on this analysis, we suppose that homo-oligomers of the pyridine monomer **2** have a propensity towards intramolecular hydrogen bonding which disfavors the formation of intermolecular hydrogen bonds with the tripeptides in our library. Monomer **2** should also have the same effect in the context of heterogeneous oligomers built from **1**, **2** and other *R,R*-diaminopyrrolidine monomers By this example it can been seen that subtle changes in monomer structure can result in large changes in oligomer binding properties, and that preliminary studies such as these are needed to fortify a library with the monomers most likely to result in productive oligomers, *i.e.* , those that are both rigid and able to form intermolecular hydrogen bonds.

To analyze the binding properties of the other sequence-selective ligand **9**, we made and tested the abbreviated oligomers **10–14**. Sequence-selective binding was found to be due exclusively to the dimer **5-4**, and this portion of the molecule bound the same two tripeptides regardless of the oligomer sequence context in which

it was tested. In addition, the sequence selectivity was excellent, with virtually no other library beads showing detectable color at concentrations where the selected beads were moderate to dark orange in color. The minimum concentration to achieve detectable binding was 31 μM.



From Monomer 1 From Monomer 2

Figure 2. Conformational difference between dimers of **1** and **2.**

In summary, we have designed and synthesized a novel set of conformationally restricted *eta-* and *theta-*amino acids for use in an oligomeric combinatorial approach to find chemoreceptors with sequence-selective oligopeptide binding ability. While studying representative oligomers made from these monomers we have found two novel low molecular weight compounds with the ability to bind specific tripeptide sequences. The minimum receptor structures along with their cognate peptides are shown in Figure 3. These small organic motifs with tripeptide-binding properties may, in conjunction with other low molecular weight receptors found in our laboratory,[2] provide the basis for the development of receptors which will be capable of recognizing peptide sequences of greater length.

Additionally, we have found one monomer, **2**, in our initial set that appears poorly suited for use in chemoreceptor library construction. This property appears to be due to a conformational preference favoring

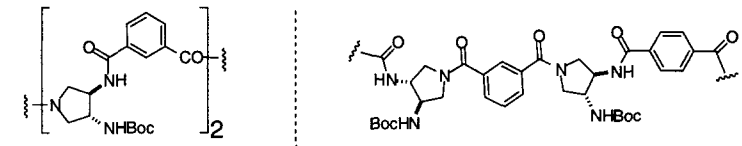


Figure 3. Novel chemoreceptor motifs and their cognate tripeptides

interresidue, intramolecular hydrogen bonding. The dipeptide made from **2**, when considered along with its counterpart made from **1** (Figure 2), underscores the two distinct roles that hydrogen bonds can play in chemoreceptor design using oligomeric molecules. Whereas hydrogen bonds are needed to provide binding

strength with recognition targets, they can also serve to stabilize well-defined secondary structures, as evidenced by the helical foldamers.[3] Using hydrogen bonding in these two distinct ways in one chemoreceptor to achieve secondary structure based recognition will require subtle control over conformational preferences in both the bound and unbound forms of potential foldameric chemoreceptors.

Acknowledgment: This research was supported by NSF grant CHE97-07870, and an NIH NRSA post-doctoral fellowship to K.R.

References and Notes

1. Cram, D. J. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1009.

2. Still, W. C. *Acc. Chem. Res.* **1996**, *29*, 155.

3. Gellman, S. H. *Acc. Chem. Res.* **1998**, *31*, 173.

4. Compounds **1–5** were characterized by ^{1}H NMR and FABHRMS. Compounds **6–14** were characterized by ^{1}H NMR and FABMS.

5. Reddy, D. R.; Thornton, E. R. *J. Chem. Soc., Chem. Commun.* **1992**, 172.

6. Each round of chain elongation included: EDC coupling using the pyrrolidine ring nitrogen in CH_2Cl_2 or a CH_2Cl_2/DMF mixture, Fmoc deprotection with 2.9 M diethylamine in DMF, and preparative TLC where necessary. In the synthesis of oligomers containing amino acid **5**, the Teoc group was removed by treatment with Bu_4NF in DMF (24 h).

7. A weight of approximately six copies of the library was shaken in a receptor solution of 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. At least 12 h after the number of red beads appeared constant, all of darkest beads were picked and decoded.[10]

8. Yoon, S. S.; and Still, W. C. *Tetrahedron* **1995**, *51*, 567, and references therein. The library contained 15^3 potential sequences of the structure: Ac-AA3-AA2-AA1-HN-polystyrene, where AAn represents Gly and the d or l isomers of: Ala, Ser, Val, Pro, Asn, Gln, and Lys.

9. Macromodel 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. (MM2 forcefield, chloroform, 2000 Monte Carlo steps, 1000 minimization steps using PRCG method.)

10. 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. U.S.A.* **1993**, *90*, 10922.