Peptide Receptors

DOI: 10.1002/ange.200601046

Sequence-Dependent Stereoselectivity in the Binding of Tetrapeptides in Water by a Flexible Artificial Receptor**

Carsten Schmuck and Peter Wich*

Dedicated to Professor Ronald Breslow on the occasion of his 75th birthday

The binding of small oligopeptides by artificial receptors under physiological conditions (in water) still remains a challenging task. Although early work in this area was limited to peptide recognition in organic solvents,[1] some progress has been made in the last few years for peptide binding in more polar solvents. Some examples of oligopeptide binding in aqueous solvents based on metal–ligand[2] or hydrophobic interactions,[3] and, most recently, purely electrostatic interactions (H-bonds and ion pairs)[4] have been reported. For the identification of artificial peptide receptors, combinatorial libraries have been successfully employed by us and other groups.[5] In some cases, receptors found in this way even demonstrate pronounced selectivity for closely related substrates at least in qualitative on-bead binding assays. However, a quantitative confirmation of the reported on-bead substrate selectivities in free solution is lacking. For example, Chamorro and Liskamp[6] reported the screening of a large combinatorial receptor library derived from a cyclotriveratrylene with three attached peptide arms for the binding of D-Ala-D-Ala. By using a color-coded substrate, efficient receptors could be identified from a qualitative screening in which the dipeptide binds more efficiently than the related depsipeptide D-Ala-D-Lac. No quantitative information on binding affinities or substrate selectivity was provided. In another example, Kilburn et al. recently reported on the screening of a large combinatorial library of "tweezer receptors" (>15000 members) for the binding of the tripeptide *N*-Ac-L-Lys-D-Ala-D-Ala. Again, by using a qualitative binding assay, receptors were identified and the most promising one was then resynthesized and studied in more detail.[7] Weak binding with low-mM affinities was found for two diastereomeric tripeptides with the resin-bound receptor, but no binding data in free solution could be obtained. Therefore, the demonstration of efficient stereoselectivity in the binding of

[*] Prof. Dr. C. Schmuck, Dipl.-Chem. P. Wich
Institut für Organische Chemie
Universität Würzburg
Am Hubland, 97074 Würzburg (Germany)
Fax: (+49) 931-888-4625
E-mail: schmuck@chemie.uni-wuerzburg.de

[**] We thank the DFG (SFB 630) and the Fonds der Chemischen Industrie for ongoing financial support of our work.

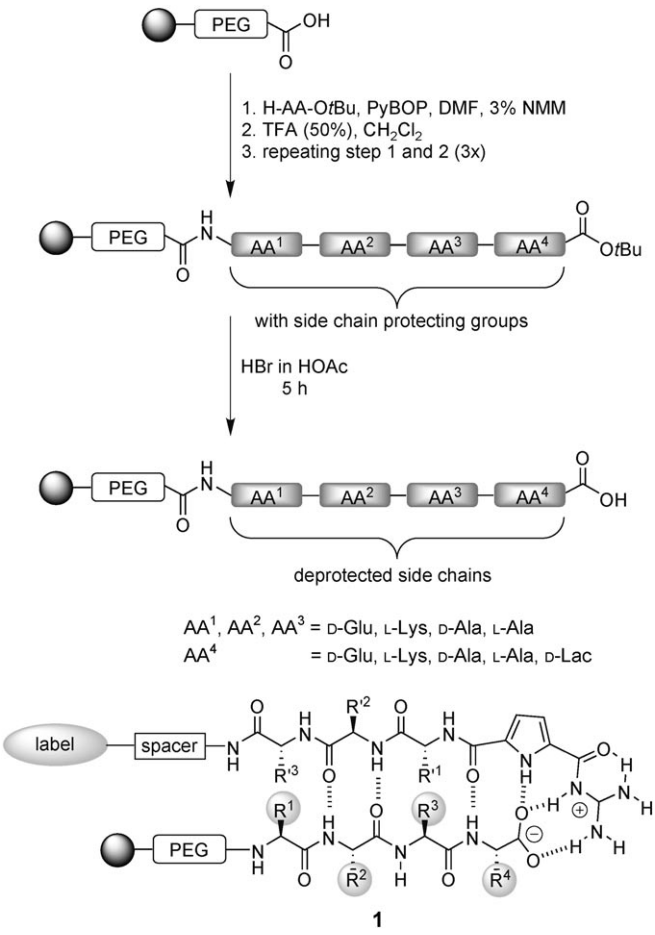
Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

oligopeptides in water still remains an unsolved problem. Herein, we report that the fully flexible peptide receptor **1** shows not only significant substrate selectivity but also a remarkable sequence dependent stereoselectivity in the binding of polar tetrapeptides in water.

Recently, we demonstrated that the tris-cationic receptor **1** (Gua-Lys-Lys-Phe; Gua = guanidiniocarbonyl pyrrole) efficiently binds the tetrapeptide D-Glu-L-Lys-D-Ala-D-Ala-OH with $K_a > 10^4\,M^{-1}$ in buffered water.[4] This polar tetrapeptide is interesting owing to its resemblance to a bacterial peptidoglycan which plays a crucial role in cell wall synthesis and antibiotic activity.[8] Receptor **1** was identified from a combinatorial screening of a small but focused receptor library with 512 members by using a fluorescence-labeled derivative of this tetrapeptide. From a screening of the whole library a detailed quantitative structure–binding relationship could be established. This relationship showed that complex formation is dominated by electrostatic interactions. Following this approach to use rather small but custom designed combinatorial libraries that can be fully analyzed quantitatively in detail, we have now screened receptor **1** for its binding selectivity against a combinatorial library of 320 different but closely related substrates.

The substrate library was prepared by inverse solid-phase peptide synthesis on TentaGel as shown in Scheme 1. To introduce structural diversity, a combinatorial split-and-mix synthesis was performed based on the IRORI radiofrequency tagging technology.[9] In each coupling step, the following four amino acids were used as their *tert*-butyl esters: D-Glu(OBzl), L-Lys(Cbz), D-Ala, and L-Ala (Bzl = benzyl, Cbz = carbobenzoxy). D-Lac was also incorporated at the fourth position of the C terminus of the substrate[10] giving rise to a total library size of $4\times4\times4\times5 = 320$ substrates. It was necessary during the course of the library synthesis to develop a modified protocol for the cleavage of the benzyl-protected glutamate side chains as the standard protocol normally used in the literature (25% HBr in HOAc, 16 h)[11] also led to a complete cleavage of the tetrapeptides from the solid support. This cleavage was owing to the TentaGel being more acid sensitive than the standard Merrifield resin.[12] Time-dependent cleavage studies with a model peptide (resin-Ala-Glu(OBz)-Lys(Z)-Phe-NH_2) synthesized on TentaGel showed that after more than 2 h of treatment with 33% HBr in HOAc the glutamate side chains were cleaved (as determined by a malachite-green test[13]), whereas the resin started to decompose only after 10 h. Therefore, after the library synthesis, we used a 5 h treatment with 33% HBr in HOAc to remove all benzylic protecting groups.

The 320 solid-phase-bound deprotected substrates were then screened for their binding affinity towards receptor **1** by first using a qualitative on-bead screening assay through incubation of the library members with the Dansyl-labeled receptor **1** in aqueous buffer solution. This assay showed that even at concentrations as low as 5 μM, receptor **1** efficiently and with high selectivity binds to some of the substrates within the library (Figure 1). We then performed a quantitative screening assay of the whole substrate library. From the change in UV absorption of the solution before and after incubation with the labeled receptor and the known loading



Scheme 1. Synthesis of a combinatorial library containing 320 tetrapeptides and depsipeptides by inverse solid-phase peptide synthesis on tentagel and schematic representation of the complex formed with the triscationic guanidiniocarbonyl pyrrole peptide receptor **1** (spacer = ε-aminocaproic acid, $R'^1 = R'^2$ = Lys, R'^3 = Phe). AA = amino acid, Ac = acetyl, PyBOP = (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophoslate, DCM = dichloromethane, DMF = *N*,*N*-dimethylformamide, NMM = *N*-methylmorpholine, PEG = polyethylene glycol, Py = pyridine, TFA = trifluoroacetic acid.

of the resin, the binding constants can be calculated by using a procedure described earlier.[4a] Some selected binding constants are given in Table 1. The best binding substrate in this library is D-Glu-D-Glu-D-Glu-D-Glu-D-Glu-OH (substrate **2**, $K_a = 26\,500\,M^{-1}$), which is not surprising for a tris-cationic receptor molecule. The binding affinity is remarkably high, keeping in mind that both receptor and substrate are fully flexible molecules. This emphasizes that multiple charge interactions can be very efficient even in water. As the binding is dominated by electrostatic interactions, the weakest binding affinities with $K_a \le 100\,M^{-1}$ are observed for those substrates that contain only nonpolar (Ala/Lac) or positively charged (Lys) side chains (e.g. substrates **10** or **11**). Only overall negatively charged substrates are bound efficiently. Binding of L-Lys-L-Lys-L-Lys-L-Lys-OH could not even be detected in this assay. Overall, the binding constants vary by more than a factor of 300, which is a remarkable substrate selectivity for such a small and flexible receptor, especially as the 320 substrates within the library are structurally very similar. The

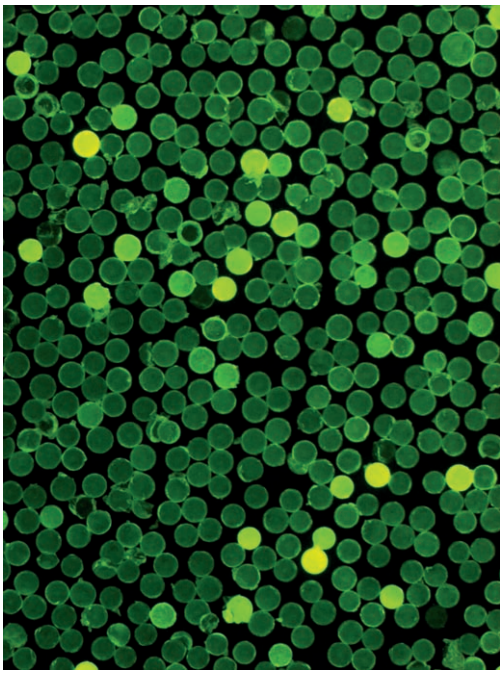


Figure 1. Incubation of the solid-phase bound substrate library with the fluorescence-labeled receptor **1** reveals a selective interaction with only some substrates within the library despite their structural close resemblance (aqueous buffer, pH 6.10, [**1**] = 5 μM).

Table 1: Selected association constants obtained from a quantitative on-bead assay for the binding of different substrates by receptor **1**.

Substrate	AA[1]	AA[2]	AA[3]	AA[4]	K_a [M^{-1}][a]
2	D-**Glu**	D-**Glu**	D-**Glu**	D-**Glu**	26500
3	L-Ala	D-**Glu**	D-**Glu**	D-**Glu**	8500
4	D-**Glu**	D-**Glu**	D-Ala	D-**Glu**	5000
5	D-**Glu**	D-Ala	L-Ala	D-**Glu**	4800
6	D-Ala	L-Ala	D-**Glu**	D-**Glu**	2600
7	D-Ala	D-Ala	L-Ala	D-**Glu**	1200
8	D-Ala	D-**Glu**	L-Ala	D-Ala	700
9	D-Ala	L-Ala	D-**Glu**	D-Ala	600
10	L-Lys	D-Ala	L-Ala	L-Lys	< 100
11	L-Ala	L-Ala	L-Ala	D-Ala	< 100

[a] Estimated error limit of the assay in $K_a \pm 30\%$.

substrates have three different side chains (Ala, Lys, and Glu) and differ only in the absolute configuration of either one building block (D/L-Ala) or the nature of one chemical linkage (Ala compared with Lac).

Even more surprising than the pronounced substrate selectivity is that receptor **1** also shows a distinct sequence-dependent stereoselectivity (Table 2): The receptor prefers D-Ala over its enantiomer L-Ala (K_a up to a factor of 10 greater) but only within certain sequences. Stereoselectivity is only observed for tetrapeptides that contain not more than one D/L-Ala located next to three D-Glu. This stereoselectivity is only found when the D/L-Ala is located in positions 2–4 of the tetrapeptide substrate, but not when it is in the N-terminal position 1. For example, although substrates **12** and **3** show nearly identical binding affinities ($K_a \approx 8000\,M^{-1}$), the association constants drops from $K_a = 5000\,M^{-1}$ to $< 1000\,M^{-1}$ if the

Table 2: Changes in the binding constant upon the exchange of D-Ala for L-Ala or D-Lac in various positions of the substrate as obtained from a quantitative on-bead screening.

Substrate	AA[1]	AA[2]	AA[3]	AA[4]	K_a [M^{-1}][a]
12	D-**Ala**	D-Glu	D-Glu	D-Glu	7900
3	L-**Ala**	D-Glu	D-Glu	D-Glu	8500
13	D-Glu	D-**Ala**	D-Glu	D-Glu	5200
14	D-Glu	L-**Ala**	D-Glu	D-Glu	600
4	D-Glu	D-Glu	D-**Ala**	D-Glu	5000
15	D-Glu	D-Glu	L-**Ala**	D-Glu	900
16	D-Glu	D-Glu	D-Glu	D-**Ala**	2100
17	D-Glu	D-Glu	D-Glu	L-**Ala**	200
18	D-Glu	D-Glu	D-Glu	D-**Lac**	300

[a] Estimated error limit of the assay in $K_a \pm 30\%$.

L-Ala is introduced in position 2 or 3 (compare substrates **13** and **14**, and **4** and **15**, respectively). The drop in affinity is even more pronounced if the L-Ala is positioned directly opposite the guanidiniocarbonyl pyrrole anion binding site (substrates **16** and **17**; $K_a = 2100\,M^{-1}$ and $K_a < 200\,M^{-1}$, respectively).

This remarkable and unexpected sequence dependent stereoselectivity was independently confirmed by UV-titration studies in free solution. First we determined the binding of Ac-D-Glu-D-Glu-D-Glu-D-Glu-OH **2** to the unlabelled receptor **1** in free solution. Aliquots of the tetrapeptide substrates were added to a solution of receptor **1** in buffered water ($[\mathbf{1}]_o = 50\,\mu M$, pH 6.10, 1.5 mM bis(2-hydroxyethyl)-imino-tris(hydroxymethyl)methane (Bis-Tris) buffer solution). From the decrease in the UV absorption of the pyrrole band at $\lambda = 300$ nm upon substrate addition, the binding constant was calculated by using a nonlinear least-square fitting procedure for a 1:1 complex formation (Figure 2). The complex stoichiometry was confirmed by a Job-plot extracted from the titration data.[14] The calculated association constant, $K_a = 23700\,M^{-1}$, for complex formation in free solution is in good agreement with the result from the on-bead screening ($K_a = 26500\,M^{-1}$).

We then tested the binding of two pairs of diastereomeric tetrapeptides, D-Glu-D-Glu-D/L-Ala-D-Glu-OH and D/L-Ala-D-Glu-D-Glu-D-Glu-OH, for their affinity to the unlabelled receptor **1** in free solution. Again UV titrations were performed by using the same conditions as described above. Stereoselectivity is only observed when the D-Ala/L-Ala exchange takes place at position 3 of the tetrapeptide but not at the N terminus (position 1). The two substrates D/L-Ala-D-Glu-D-Glu-D-Glu-OH (**12** and **3**) are bound with similar affinities by receptor **1** ($K_a = 5100\,M^{-1}$ and $4800\,M^{-1}$, respectively). However, from the two diastereomers D-Glu-D-Glu-D/L-Ala-D-Glu-OH (**4** and **15**) the one with D-Ala is preferred ($K_a = 4500\,M^{-1}$ versus $1400\,M^{-1}$, respectively) as was also found in the on-bead screening. Hence, even though the absolute binding constants determined on bead and in free solution differ (for example, owing to the different concentrations of buffer salts used in the experiments and also the influence of the solid support itself), the relative trends are the same.

Such a pronounced stereoselectivity owing to the change in the absolute stereochemistry of a single amino acid within a tetrapeptide is remarkable, even more so as the D-Ala/L-Ala

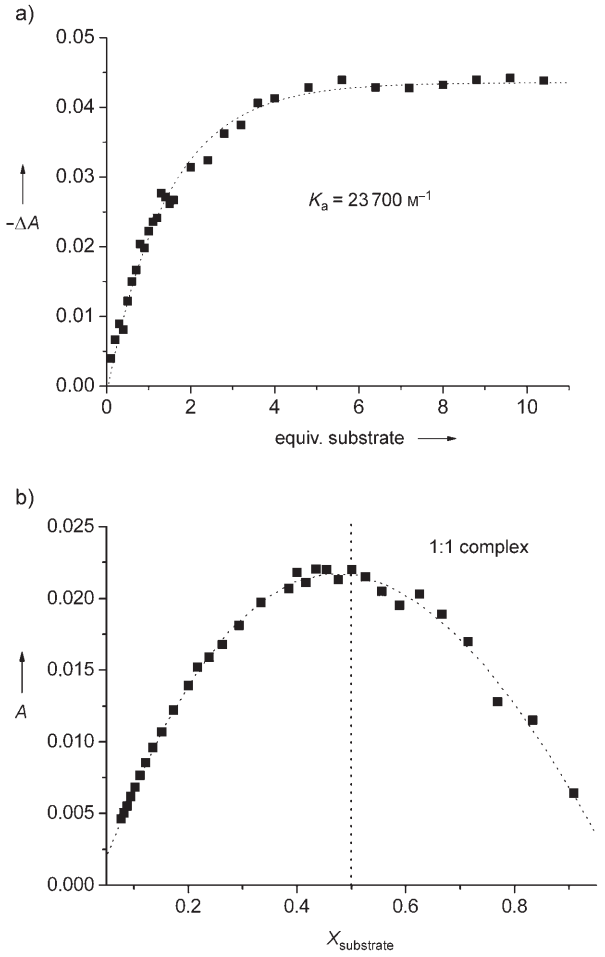


Figure 2. a) Binding isotherm at $\lambda = 300$ nm for the complexation of *N*-Ac-D-Glu-D-Glu-D-Glu-D-Glu-OH **2** by receptor **1** as obtained from a UV titration ($[\mathbf{1}]_0 = 50$ μM, 1.5 mM Bis-Tris buffer solution, pH 6.1; data corrected for absorption changes owing to dilution). The dotted line represents the curve-fitting for a 1:1 complexation. b) Job plot obtained from the titration confirming the 1:1 binding stoichiometry. $X_{substrate}$ = molar fraction of the substrate.

pair is the smallest stereochemical change possible within a peptide. Most likely, stereoselectivity requires a rather well defined complex structure and is only possible when the position of the D-Ala/L-Ala exchange is fixated at both sides by strong charge interactions between receptor and substrate. The alanine itself can only form H-bonds to the receptor that are, in water, too weak to play a significant role for complex formation. Hence, either a glutamate or the C-terminal carboxylate is needed next to this position. Both groups are negatively charged and allow for strong electrostatic interactions with the tris-cationic receptor, either with the lysines or the guanidiniocarbonyl pyrrole moiety, that stabilizes the complex around the Ala. This would explain why stereoselectivity is not found for other peptide sequences within the library besides those cited in Table 2. For example, if there is another alanine next to the exchange position, the complex around the D/L-Ala exchange is perhaps too flexible to allow for a discrimination. Hence, only sequences with just one D/L-Ala pair with three glutamates show stereoselectivity. Also, no stereoselectivity is observed when the alanine is in the N-terminal position next to three glutamates (substrates **12** and **3** in Table 2). In this case only the first three amino acids of the tetrapeptide substrate are efficiently bound by the receptor, whereas the N-terminal alanine does not contribute significantly to the stability of the complex. Most likely, this terminal alanine is simply a dangling end as depicted in Figure 3. Hence, it does not matter if a D-Ala or a L-Ala is present at this position.

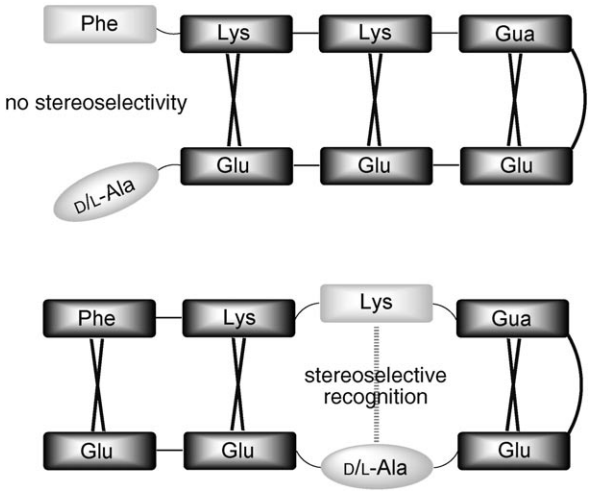


Figure 3. The stereoselectivity of receptor **1** depends on the sequence of the bound substrate. Stereoselectivity is only observed when the position where the D-Ala/L-Ala exchange takes place is fixated at both sides by strong electrostatic interactions between receptor (top) and substrate (bottom).

In good agreement with this multipoint-binding model is also the observation that the exchange of D-Ala for D-Lac at the C-terminal position of the substrate also significantly reduces the binding affinity. Although the peptide is bound with $K_a = 2100\,\text{M}^{-1}$, the affinity for the depsipeptide drops to $K_a \approx 300\,\text{M}^{-1}$ (substrates **16** and **18** in Table 2). This modification exchanges an attractive H-bond from a carbonyl CO of the receptor to an amide NH of the peptide substrate for a dipol–dipol repulsion with the ester oxygen of the depsipeptide. For such a small exchange to show up in the binding affinities, a well-defined complex structure around this position is needed. The same exchange is the origin of an upcoming bacterial resistance against the antibiotic vancomycin.[15] In this case the complex formed between vancomycin and the dipeptide fragment D-Ala-D-Ala is also well defined owing to the rigid structure of the antibiotic. Because of the extensive hydrophobic shielding of the substrate within the complex, the drop in affinity for the depsipeptide is even more pronounced. However, for an artificial receptor this is, to the best of our knowledge, the highest selectivity between this biological model dipeptide and the depsipeptide reported so far.

In conclusion, we have shown that first, the one-armed flexible peptide receptor **1** efficiently binds anionic tetrapeptides such as **2**, even in water, with association constants up to $K_a \approx 27\,000\,\text{M}^{-1}$. Second, even very closely related substrates are efficiently discriminated against as could be shown by

screening a combinatorial library of 320 substrates. Third, a remarkable stereoselectivity between D-Ala and L-Ala is found both on the solid support as well as in free solution, but only in certain sequences that allow formation of overall well-defined complexes. Forth, the receptor is also capable of distinguishing a peptide from a depsipeptide. We are currently exploring how to further improve the selectivity of such receptors in free solution by introducing additional hydrophobic shielding elements.

Received: March 16, 2006
Published online: May 31, 2006

Keywords: amino acids · artificial peptide receptors · combinatorial chemistry · molecular recognition · supramolecular chemistry

[1] For reviews on the early work of peptide recognition, see: a) H.-J. Schneider, *Angew. Chem.* **1993**, *105*, 890 – 892; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 848 – 850; b) M. W. Peczuh, A. D. Hamilton, *Chem. Rev.* **2000**, *100*, 2479 – 2493.

[2] For selected examples of oligopeptide binding in aqueous solvents based on metal-ligand interactions, see: a) A. T. Wright, E. V. Anslyn, *Org. Lett.* **2003**, *5*, 1341 – 1344; b) T. Mizutani, K. Wada, S. Kitagawa, *Chem. Commun.* **2002**, 1626 – 1627; c) S. Sun, M. A. Fazal, B. C. Roy, S. Mallik, *Org. Lett.* **2000**, *2*, 911.

[3] For examples of peptide recognition based on hydrophobic interactions, see: a) C. Schmuck, M. Heil, *ChemBioChem* **2003**, *4*, 1232 – 1238; b) M. Sirish, H.-J. Schneider, *Chem. Commun.* **1999**, 907 – 908; c) R. Xu, G. Greiveldinger, L. E. Marenus, A. Cooper, J. A. Ellman, *J. Am. Chem. Soc.* **1999**, *121*, 4898 – 4899; d) M. Davies, M. Bonnat, F. Guillier, J. D. Kilburn, M. Bradley, *J. Org. Chem.* **1998**, *63*, 8696 – 8703; e) R. Breslow, Z. Yang, R. Ching, G. Trojandt, F. Odobel, *J. Am. Chem. Soc.* **1998**, *120*, 3536 – 3537.

[4] For peptide recognition in water mainly based on electrostatic interactions, see: a) C. Schmuck, M. Heil, *Chem. Eur. J.* **2006**, *12*, 1339 – 1348; b) C. Schmuck, M. Heil, J. Scheiber, K. Baumann, *Angew. Chem.* **2005**, *117*, 7374 – 7379; *Angew. Chem. Int. Ed.* **2005**, *44*, 7208 – 7212.

[5] For review articles on the use of combinatorial receptor libraries, see: a) N. Srinivasan, J. D. Kilburn, *Curr. Opin. Chem. Biol.* **2004**, *8*, 305 – 310; b) B. Linton, A. D. Hamilton, *Curr. Opin. Chem. Biol.* **1999**, *3*, 307 – 312; c) Y. R. DeMiguel, J. M. K. Sanders, *Curr. Opin. Chem. Biol.* **1998**, *2*, 417 – 421; d) W. C. Still, *Acc. Chem. Res.* **1996**, *29*, 155 – 163.

[6] C. Chamorro, R. M. J. Liskamp, *Tetrahedron* **2004**, *60*, 11145 – 11157.

[7] J. Shepherd, T. Gale, K. B. Jensen, J. D. Kilburn, *Chem. Eur. J.* **2006**, *12*, 713 – 720.

[8] a) C. C. McComas, B. M. Crowley, D. L. Boger, *J. Am. Chem. Soc.* **2003**, *125*, 9314 – 9315; b) C. T. Walsh, *Nature* **2000**, *406*, 775 – 781; c) D. H. Williams, B. Bardsley, *Angew. Chem.* **1999**, *111*, 1264 – 1286; *Angew. Chem. Int. Ed.* **1999**, *38*, 1172 – 1193; d) M. Nieto, H. R. Perkins, *Biochem. J.* **1971**, *123*, 780 – 803.

[9] For a description of the IRORI tagging system, see: a) K. C. Nicolaou, J. A. Pfefferkorn, H. J. Mitchell, A. J. Roecker, S. Barluenga, G.-Q. Cao, R. L. Affleck, J. E. Lillig, *J. Am. Chem. Soc.* **2000**, *122*, 9954 – 9967; b) X.-Y. Xiao, Z. Parandoosh, M. P. Nova, *J. Am. Chem. Soc.* **1997**, *119*, 6029 – 6033.

[10] We determined the necessary coupling conditions of the *tert*-butyl ester of D-Lac compared to normal amino acids by synthesis of a model peptide sequence on *p*-methyl-benzyldryl-amine (MBHA) resin. The free amino function of this resin was transformed with the help of Z-Glu(O*t*Bu) into a carboxylate function allowing an analogous N to C solid-phase peptide-synthesis protocol like in the combinatorial library. The only difference is the higher acid sensitivity of the MBHA resin compared to TentaGel. It also allows the cleavage of the C-terminal *tert*-butyl esters with 50% TFA. In addition, the synthesized peptide can be cleaved with trifluoromethane sulfonic acid (TFMSA; 10% in TFA) and then analyzed with NMR spectroscopic techniques. These tests showed that it is possible to use the same conditions as those for like for amino acids to guarantee the quantitative coupling of D-Lac.

[11] S. Sano, S. Kawanashi, *J. Am. Chem. Soc.* **1975**, *97*, 3480 – 3484.

[12] However, the latter can not be used here as it does not swell in aqueous solutions and hence does not allow subsequent on-bead binding studies in water. For this purpose the hydrophilic TentaGel is the best choice.

[13] M. E. Attardi, G. Porcu, M. Taddei, *Tetrahedron Lett.* **2000**, *41*, 7391 – 7394.

[14] P. McCarthy, *Anal. Chem.* **1978**, *50*, 2165 – 2165.

[15] C. T. Walsh, S. L. Fisher, I.-S. Park, M. Prahalad, Z. Wu, *Chem. Biol.* **1996**, *3*, 21 – 28.

