VIP Carbohydrate Recognition

DOI: 10.1002/anie.200906620

Design, Synthesis, and Screening of a Library of Peptidyl Bis(Boroxoles) as Oligosaccharide Receptors in Water: Identification of a Receptor for the Tumor Marker TF-Antigen Disaccharide**

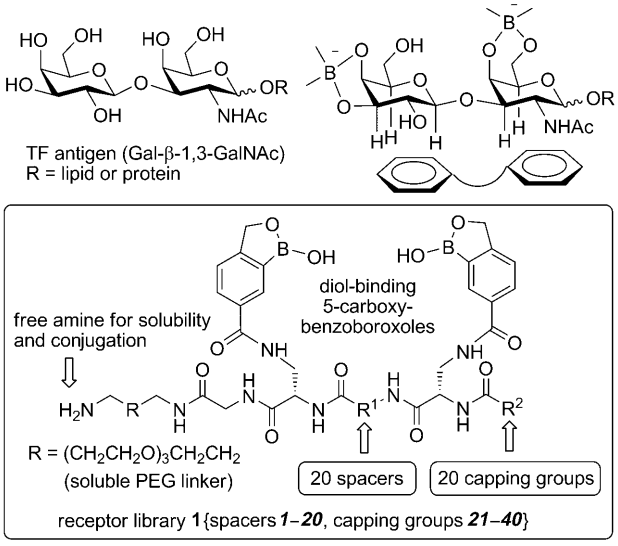
*Arnab Pal, Marie Bérubé, and Dennis G. Hall**

The important challenge of carbohydrate recognition in water presents several exciting opportunities in chemical biology and medicine, with potential applications in the analysis, purification, diagnostic, and physiological control of biologically important glycans. Although a number of synthetic receptors have been described for the recognition of complex carbohydrates in organic solvents,[1] it is notoriously difficult to achieve the same efficiency under physiological conditions (i.e., water at neutral pH).[2] The essence of the problem lies in the competition between the multiple hydroxy groups on the carbohydrates and the overwhelming ones from the bulk solvent, water. Recently, Davis and co-workers described a water-soluble, complex cage-like receptor that displayed low millimolar affinity and high selectivity for a distinct disaccharide in neutral water.[3]

To be general, however, a receptor approach must address the structural diversity of oligosaccharides while overcoming the difficulty of predicting their favored conformations. Combinatorial strategies appear to be ideal to address this problem. To this end, we sought to design a simple class of small receptors that could be synthesized in a modular fashion amenable to the preparation of libraries. Any approach to the recognition of carbohydrates in water should take advantage of the intrinsic orientation of the sugar's hydroxy groups on the rigid oxacarbocyclic skeleton. In this regard, boronic acids have the ability to form boronic esters reversibly with polyols and sugars in water.[4,5] While the use of boronic acids is regarded as one of the most promising approaches for the recognition of carbohydrates in water,[6] it is not without limitations. First and foremost, as boronic acids display a marked preference for binding furanose sugars,[7] cell-surface glycoconjugates are unfavorable targets because they are comprised of hexopyranosides.[8] To overcome this issue, Wang and co-workers successfully targeted a glycoprotein using a very large library of boronic acid-modified DNA aptamers termed "boronolectins".[9] To avoid selectivity issues that complicate the optimization of a discrete receptor, Lavigne and co-workers used arrays of unidentified peptide-boronic acids from a large mixture library to detect glycoproteins according to response patterns.[10]

In contrast to these strategies, we favor an approach targeting each glycan of interest with a single, low molecular weight receptor of defined composition. This approach exploits benzoboroxoles as hexopyranoside-binding agents,[11] and couples boronate formation with other modes of molecular recognition inspired from nature's carbohydrate-binding proteins, lectins. In this model study, a rationally designed library of synthetic receptors[12] is targeted against an important tumor-associated carbohydrate antigen, the Thomsen–Friedenreich (TF) disaccharide (Gal-β-1,3-GalNAc, Scheme 1).[13] Synthetic TF receptors are of particular interest because anti-TF antibodies have proven difficult to optimize.[14]

Because the TF antigen possesses two of the 4,6-diol or *cis*-3,4-diol units that bind preferentially with benzoboroxoles,[11] two such units were included on the receptors (Scheme 1). A peptide backbone was chosen for ease of synthesis and also for providing hydrogen bonding donor and



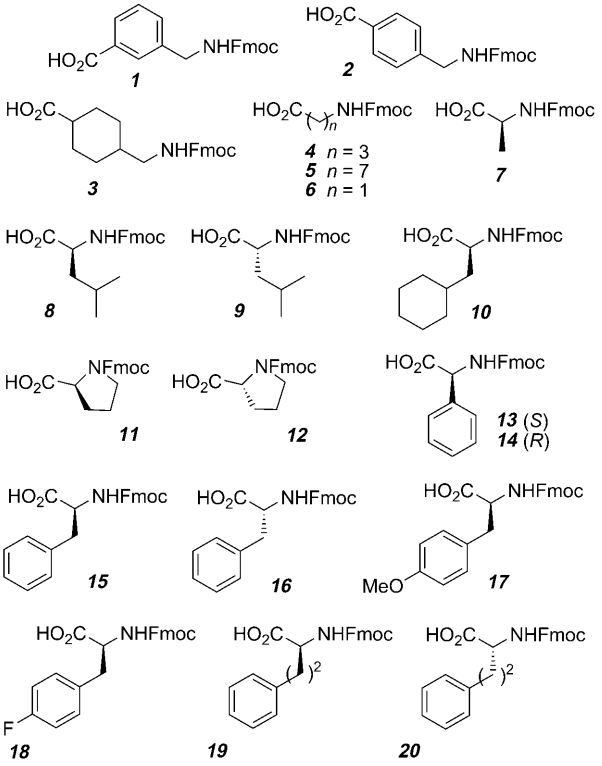
Scheme 1. Design of peptidyl bis(boroxoles) library **1**{***1–20,21–40***}.

[*] A. Pal, Dr. M. Bérubé, Prof. D. G. Hall
Department of Chemistry, Gunning-Lemieux Chemistry Centre
University of Alberta, Edmonton, Alberta, T6G 2G2 (Canada)
Fax: (+1) 780-492-8231
E-mail: dennis.hall@ualberta.ca
Homepage: http://www.chem.ualberta.ca/~dhall/

[**] This work was funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada (E.W.R. Steacie Memorial Fellowship to D.G.H.) and the University of Alberta. M.B. thanks the Killam Foundation for a postdoctoral fellowship (University of Alberta). We thank Eric Pelletier for support in library purification by HPLC, and Prof. David Bundle and Joanna Sadowska (Alberta Ingenuity Center for Carbohydrate Science) for help with the screening assay and access to the Lemieux Collection of Oligosaccharides.

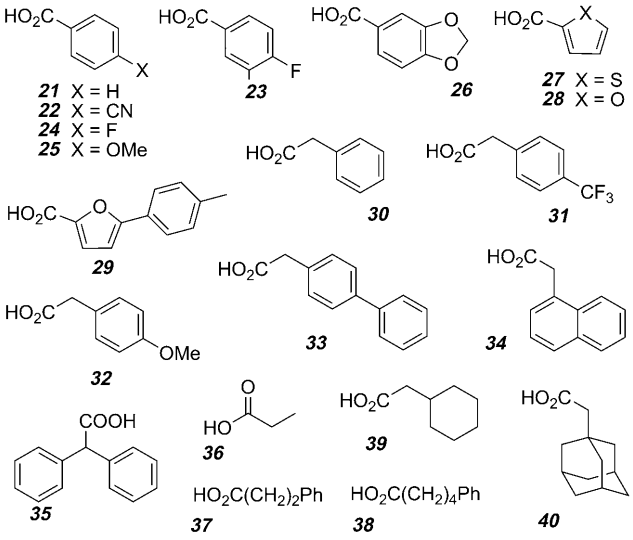
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200906620.

acceptor capabilities. The central amino acid position, flanked by two diaminopropionic acid (Dpr) residues for attachment of the benzoboroxole, was randomized with 20 natural and nonnatural amino acids ***1–20*** offering functional and geometrical diversity (spacer length and rigidity; Scheme 2).

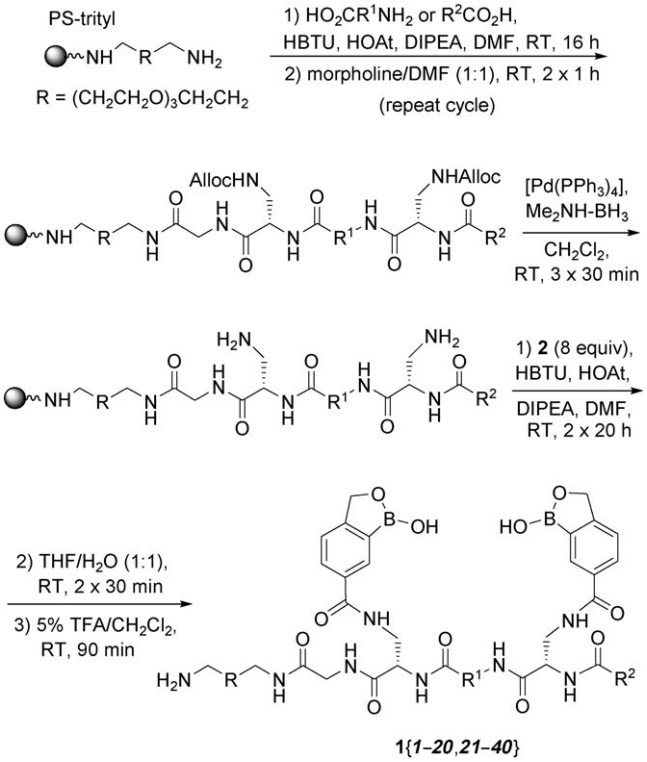


Scheme 2. Spacers ***1–20*** for library of receptors **1**{***1–20,21–40***}. Fmoc = 9-fluorenylmethoxycarbonyl.

The acyl capping group consisted in a selection of 20 carboxylic acids (Scheme 3). Several library components for both the spacer and terminal positions include aromatic subunits because they are known in carbohydrate-binding proteins to promote hydrophobic CH–π interactions with the nonpolar face of saccharides (Scheme 1).[15] The length and flexibility of the carboxybenzoboroxole-functionalized Dpr side arms should allow for effective interactions of these subunits with either faces of the disaccharide. The library was assembled by solid-phase peptide synthesis using trityl resin, and includes a short triethyleneglycol spacer and an anchoring primary amine for increased aqueous solubility and for eventual conjugation purposes (Scheme 4). The library of 400 receptors was prepared expeditiously by split–pool synthesis using the IRORI radio-frequency encoding technology with MicroKan reactors.[16] Key to this strategy is a late-stage coupling of the diaminopropionic acid side chain amines with 5-carboxybenzoboroxole (**2**).[17] Compared to benzoboroxole, an amide model of **2** provides a significant increase in hexopyranoside-binding affinity, which is attributable to the enhanced acidity gained through the electron-withdrawing carboxamide (Scheme 5). Once completed, library **1**{***1–20,21–40***} was cleaved with dilute trifluoroacetic acid,



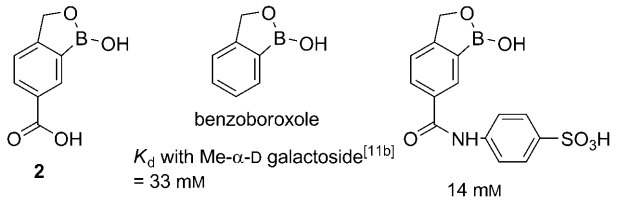
Scheme 3. Capping groups ***21–40*** for library of receptors **1**{***1–20,21–40***}.



Scheme 4. Preparation of library **1**{***1–20,21–40***}. HBTU = *O*-(benzotriazol-1-yl)-tetramethyluronium hexafluorophosphate; HOAt = hydroxy-7-azabenzotriazole; DIPEA = diisopropylethylamine; Alloc = allyloxycarbonyl; TFA = trifluoroacetic acid.

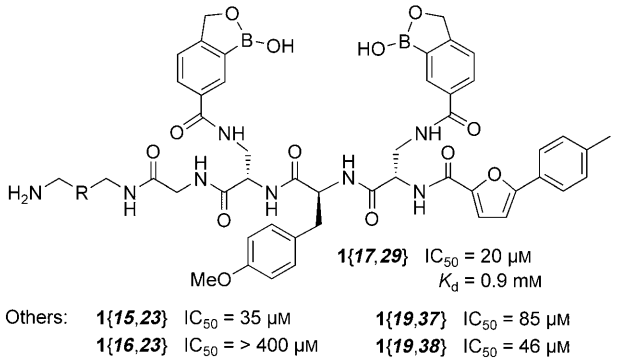
and all 400 members were purified by semipreparative HPLC.[17,18]

Screening of the library of 400 peptidyl bis(boroxole) receptors for binding to the TF antigen disaccharide was performed using a competitive ELISA in 96-well plates coated with Gal-β-1,3-GalNAc-O$(CH_2)_8$CO-BSA (ca. 5.5 units/protein).[17] The competing protein receptor in



Scheme 5. Pyranoside-binding with a 5-carboxybenzoboroxole amide.

this assay is the *Arachis hypogaea* (peanut) agglutinin lectin (PNA), a tetrameric protein known to bind to Gal-β-1,3-GalNAc with a dissociation constant K_d of 1×10^{-7} M.[19] The peroxidase-labeled PNA lectin (1.0×10^{-9} M) was incubated with each library member at a high concentration of 400 μM. Following washing operations and addition of the chromogenic substrate (3,3′,5,5′-tetramethylbenzidine), the absorbance was measured at 450 nm.[20] A total of 17 hits were confirmed reproducibly, and IC_{50} values were measured on the four most promising peptidyl bis(boroxoles).[17] The most potent receptor, **1**{***17***,***29***}, showed an IC_{50} of 20 μM (Scheme 6). The surprising disparity between receptors **1**{***15***,***23***} and **1**{***16***,***23***}, which differ only by the stereochemical configuration of the spacer, provided an early hint that selective and subtle molecular recognition is taking place with the TF antigen disaccharide.



Scheme 6. Most efficient receptors identified in the competitive ELISA screening of library **1**{***1–20***,***21–40***}.

The selectivity of receptor **1**{***17***,***29***} for the Gal-β-1,3-GalNAc disaccharide was assessed by monitoring the effect of various concentrations of added carbohydrates incubated under similar assay conditions in the presence of a fixed 20 μM concentration of **1**{***17***,***29***}, followed by washing and addition of the PNA lectin. As depicted in Figure 1, soluble Gal-β-1,3-GalNAc-O$(CH_2)_8CO_2$Et[17] interfered very strongly with the binding of **1**{***17***,***29***} to the Gal-β-1,3-GalNAc-O$(CH_2)_8$CO-BSA coated plates. The individual Gal and GalNAc glycosides as well as the Gal-containing Lewis B tetrasaccharide[17] competed only to a small extent at high concentrations, while the structurally unrelated oligosaccharides trehalose and cellobiose had no effect on the binding of **1**{***17***,***29***} to the Gal-β-1,3-GalNAc-O$(CH_2)_8$CO-BSA coated plates. The blood group B trisaccharide [Gal-α-1,3-(Fuc-α-1,2)Gal-β]-O$(CH_2)_8CO_2$Et competed to a significant extent, which is to

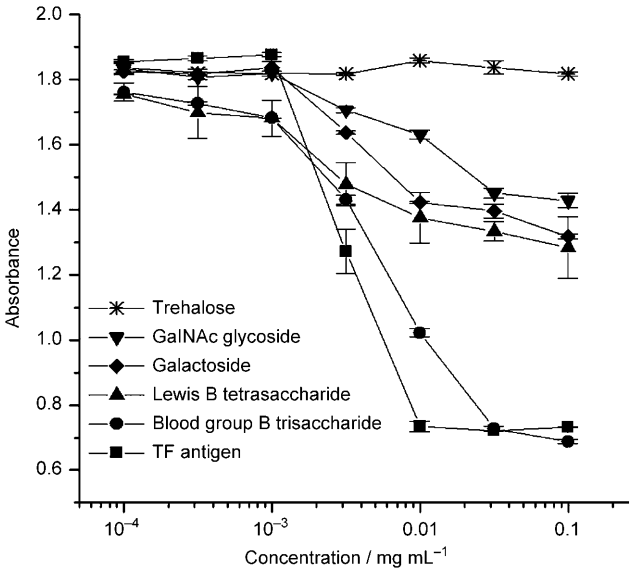
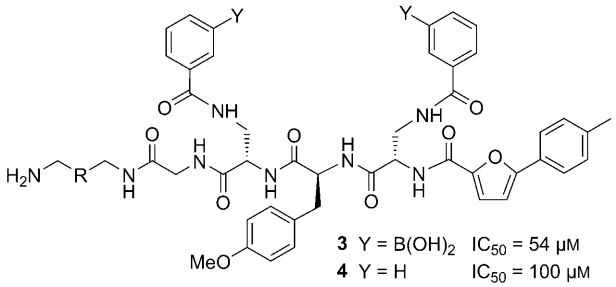


Figure 1. Competition experiments to assess the selectivity of receptor **1**{***17***,***29***} for saccharides. TF-antigen: Gal-β-1,3-GalNAc-O$(CH_2)_8CO_2$Et.

be expected given the "Gal-1,3-Gal like" polyol pattern offered by this saccharide.

Overall, these preliminary control experiments demonstrate that receptor **1**{***17***,***29***} is specifically targeting the Gal-β-1,3-GalNAc disaccharide and does so with a marked selectivity. An approximate dissociation constant between **1**{***17***,***29***} and Gal-β-1,3-GalNAc-O$(CH_2)_8CO_2$Et can be extracted from the IC_{50} value relative to the concentration of PNA lectin used in the assay and the known K_d for the complex between PNA and Gal-β-1,3-GalNAc.[19] The resulting estimated value of 0.5 mM is close to that of 0.9 mM obtained by induced circular dichroism observed on the peptidyl bis(boroxole) **1**{***17***,***29***} (measured in CH_3OH for solubility at high concentrations) and calculated according to an excellent fit to a 1:1 binding model.[17,21]

Receptor **1**{***17***,***29***} may bind its target disaccharide using several possible interactions, including boronate formation, hydrogen bonding, hydrophobic packing, and CH–π interactions. The fact that it contains the electron-rich *p*-methoxyphenylalanine and a furan as π donor components may be indicative of CH–π interactions with the electron-deficient hydrogens on the sugar rings. Because they are an important element of our receptor design, we assessed the role of the two boroxole units of **1**{***17***,***29***} by comparison with the corresponding bis(arylboronic acid) **3** and bis(phenylamide) **4** (Scheme 7), which were synthesized in a similar manner as in Scheme 4.[17] With respective IC_{50} values of 54 and 100 μM, it is not surprising to confirm that boroxoles are more favorable than normal boronic acids for complexing hexopyranosides. With a five-fold difference between **4** and **1**{***17***,***29***}, it seems unlikely, however, that both boroxole units of receptor **1**{***17***,***29***} are involved in strong covalent interactions with the two accessible diols of Gal-β-1,3-GalNAc. Other interactions from the peptide backbone such as hydrogen bonding or hydrophobic packing resulting from the aromatic R^1 and R^2 components, must contribute significantly. These issues



Scheme 7. Control compounds **3** and **4**: analogues of **1**{***17***,***29***} with modified benzamide side chains.

will be addressed in future studies on the structure of the complex.

In summary, we described a receptor design strategy that exploits several modes of molecular recognition, including the unique ability of benzoboroxoles to complex hexopyranosides. The synthesis is modular, thus well suited to targeting specific oligosaccharides using combinatorial receptor libraries. This approach was successful in identifying a low molecular weight receptor effective in neutral water and selective for the TF-antigen disaccharide, a pivotal cancer marker. Although it is remarkable that the moderate binding affinity of this receptor matches the efficiency of some lectins, it is still unsuitable for many applications. To this end, further studies will aim to exploit multivalency effects with oligomeric receptors and assess their efficiency in the labeling of TF-specific tumor cell lines.

Received: November 24, 2009
Published online: January 18, 2010

Keywords: benzoboroxoles · boronic acids · carbohydrate recognition · combinatorial chemistry · glycosides

[1] a) A. P. Davis, R. S. Wareham, *Angew. Chem.* **1999**, *111*, 3160–3179; *Angew. Chem. Int. Ed.* **1999**, *38*, 2978–2996; b) S. Striegler, *Curr. Org. Chem.* **2003**, *7*, 81–102; c) S. Kubik, *Angew. Chem.* **2009**, *121*, 1750–1753; *Angew. Chem. Int. Ed.* **2009**, *48*, 1722–1725.

[2] For recent examples, see: a) R. D. Hubbard, S. L. Horner, B. J. Miller, *J. Am. Chem. Soc.* **2001**, *123*, 5810–5811; b) M. Mazik, H. Cavga, *J. Org. Chem.* **2006**, *71*, 2957–2963; c) O. Alpturk, O. Rusin, S. O. Fakayode, W. H. Wang, J. O. Escobedo, I. M. Warner, W. E. Crowe, V. Kral, J. M. Pruet, R. M. Strongin, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9756–9760; d) T. Reenberg, N. Nyberg, J. Ø. Duus, J. L. J. van Dongen, M. Meldal, *Eur. J. Org. Chem.* **2007**, 5003–5009; e) H. Goto, Y. Furusho, E. Yashima, *J. Am. Chem. Soc.* **2007**, *129*, 9168–9169.

[3] Y. Ferrand, M. P. Crump, A. P. Davis, *Science* **2007**, *318*, 619–622.

[4] J. P. Lorand, J. O. Edwards, *J. Org. Chem.* **1959**, *24*, 769–774.

[5] G. Springsteen, B. Wang, *Tetrahedron* **2002**, *58*, 5291–5300.

[6] a) T. D. James, K. R. A. S. Sandanayake, S. Shinkai, *Angew. Chem.* **1996**, *108*, 2038–2050; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1910–1922; b) T. D. James, S. Shinkai, *Top. Curr. Chem.* **2002**, *218*, 159–200; c) W. Wang, X. Gao, B. Wang, *Curr. Org. Chem.* **2002**, *6*, 1285–1317; d) S. Jin, Y. Cheng, S. Reid, M. Li, B. Wang, *Med. Res. Rev.* **2010**, DOI: 10.1002/med.20155.

[7] M. Bielecki, H. Eggert, J. C. Norrild, *J. Chem. Soc. Perkin Trans. 2* **1999**, 449–455.

[8] A notable exception would be boronate formation with the sialyl unit of sialosides such as SLex: B. Wang, W. Yang, H. Fan, X. Gao, S. Gao, V. V. R. Karnati, W. Ni, W. B. Hooks, J. Carson, B. Weston, B. Wang, *Chem. Biol.* **2004**, *11*, 439–448.

[9] M. Li, N. Lin, Z. Huang, L. Du, C. Altier, H. Fang, B. Wang, *J. Am. Chem. Soc.* **2008**, *130*, 12636–12638.

[10] Y. Zou, D. L. Broughton, K. L. Bicker, P. R. Thompson, J. J. Lavigne, *ChemBioChem* **2007**, *8*, 2048–2051.

[11] a) M. Dowlut, D. G. Hall, *J. Am. Chem. Soc.* **2006**, *128*, 4226–4227; b) M. Bérubé, M. Dowlut, D. G. Hall, *J. Org. Chem.* **2008**, *73*, 6471–6479.

[12] For other combinatorial approaches to oligo(boronic acid) receptors: a) N. Y. Edwards, T. W. Sager, J. T. McDevitt, E. V. Anslyn, *J. Am. Chem. Soc.* **2007**, *129*, 13575–13583; b) S. Manku, D. G. Hall, *Aust. J. Chem.* **2007**, *60*, 824–828; c) P. J. Duggan, D. A. Offerman, *Tetrahedron* **2009**, *65*, 109–114.

[13] L.-G. Yu, *Glycoconjugate J.* **2007**, *24*, 411–420.

[14] P. Ravn, R. Stahn, A. Danielczyk, D. Faulstich, U. Karsten, S. Goletz, *Cancer Immunol. Immunother.* **2007**, *56*, 1345–1357, and references therein.

[15] Z. R. Laughrey, S. E. Kiehna, A. J. Riemen, M. L. Waters, *J. Am. Chem. Soc.* **2008**, *130*, 14625–14632, and references therein.

[16] a) E. J. Moran, S. Sarshar, J. F. Cargill, M. M. Shahbaz, A. Lio, A. M. M. Mjalli, R. W. Armstrong, *J. Am. Chem. Soc.* **1995**, *117*, 10787–10788; b) K. C. Nicolaou, X.-Y. Xiao, Z. Parandoosh, A. Senyei, M. P. Nova, *Angew. Chem.* **1995**, *107*, 2476–2479; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2289–2291.

[17] See the Supporting Information for details.

[18] Phenylglycine spacers **13** and **14** epimerize partly in the peptide coupling.

[19] K. J. Neurohr, N. M. Young, H. H. Mantsch, *J. Biol. Chem.* **1980**, *255*, 9205–9209.

[20] M.-G. Baek, R. Roy, *Bioorg. Med. Chem.* **2002**, *10*, 11–17.

[21] It was not possible to measure a K_d value by NMR spectroscopy due to extensive peak broadening, which can likely be attributed to slow exchange of boronate complexation on the NMR timescale.

