

Structural Basis of Multivalent Binding to Wheat Germ Agglutinin

David Schwefel,[†] Caroline Maierhofer,[‡] Johannes G. Beck,[‡] Sonja Seeberger,^{‡,§} Kay Diederichs,[†] Heiko M. Möller,^{*,‡} Wolfram Welte,^{*,†} and Valentin Wittmann^{*,‡}

Departments of Chemistry and Biology, Universität Konstanz, Universitätsstr. 10, 78457 Konstanz, Germany

Received February 25, 2010; E-mail: Mail@valentin-wittmann.de; Wolfram.Welte@uni-konstanz.de; Heiko.Moeller@uni-konstanz.de

Abstract: The inhibition of carbohydrate–protein interactions by tailored multivalent ligands is a powerful strategy for the treatment of many human diseases. Crucial for the success of this approach is an understanding of the molecular mechanisms as to how a binding enhancement of a multivalent ligand is achieved. We have synthesized a series of multivalent *N*-acetylglucosamine (GlcNAc) derivatives and studied their interaction with the plant lectin wheat germ agglutinin (WGA) by an enzyme-linked lectin assay (ELLA) and X-ray crystallography. The solution conformation of one ligand was determined by NMR spectroscopy. Employing a GlcNAc carbamate motif with α -configuration and by systematic variation of the spacer length, we were able to identify divalent ligands with unprecedented high WGA binding potency. The best divalent ligand has an IC₅₀ value of 9.8 μ M (ELLA) corresponding to a relative potency of 2350 (1170 on a valency-corrected basis, i.e., per mol sugar contained) compared to free GlcNAc. X-ray crystallography of the complex of WGA and the second best, closely related divalent ligand explains this activity. Four divalent molecules simultaneously bind to WGA with each ligand bridging adjacent binding sites. This shows for the first time that all eight sugar binding sites of the WGA dimer are simultaneously functional. We also report a tetravalent neoglycopeptide with an IC₅₀ value of 0.9 μ M being 25 500 times higher than that of GlcNAc (6400 times per contained sugar) and the X-ray structure analysis of its complex with glutaraldehyde-cross-linked WGA. Comparison of the crystal structure and the solution NMR structure of the neoglycopeptide as well as results from the ELLA suggest that the conformation of the glycopeptide in solution is already preorganized in a way supporting multivalent binding to the protein. Our findings show that bridging adjacent protein binding sites by multivalent ligands is a valid strategy to find high-affinity protein ligands and that even subtle changes of the linker structure can have a significant impact on the binding affinity.

Introduction

Multivalent interactions are frequently observed in biological systems where they govern important recognition processes.¹ Carefully studied examples include the adhesion of influenza viruses to bronchial epithelial cells,² adhesion of uropathogenic *Escherichia coli* strains to urethral endothelial cells,³ neutrophil–endothelium interactions during the inflammatory process,⁴ the binding of cholera toxin to gangliosides GM1 on cell surfaces,⁵ and the binding of transcription factors to multiple

sites on DNA.⁶ Multivalent interactions are characterized by the simultaneous binding of multiple epitopes of one entity (molecule, surface) with several binding sites of another and are responsible for the enhancement of weak interactions both in solution and at interfaces. Furthermore, they result in binding kinetics different from monovalent interactions⁷ and provide a means for the fine-tuning of such processes.

Many examples of multivalent binding involve the interaction between carbohydrates and proteins.⁸ In these cases, increased binding affinity due to multivalency has been termed the cluster glycoside effect.^{8e,9} Multidentate carbohydrate ligands are of significant medical interest in the diagnosis and inhibition of

[†] Department of Biology, Universität Konstanz.

[‡] Department of Chemistry, Universität Konstanz.

[§] Current address: Merck KGaA, Frankfurter Str. 250, 64293 Darmstadt, Germany.

- (1) (a) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2755–2794. (b) Huskens, J. *Curr. Opin. Chem. Biol.* **2006**, *10*, 537–543. (c) Whitty, A. *Nat. Chem. Biol.* **2008**, *4*, 435–439.
- (2) Lees, W. J.; Spaltenstein, A.; Kingery-Wood, J. E.; Whitesides, G. M. *J. Med. Chem.* **1994**, *37*, 3419–3433.
- (3) Connell, I.; Agace, W.; Klemm, P.; Schembri, M.; Märdil, S.; Svanborg, C. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 9827–9832.
- (4) (a) Varki, A. *J. Clin. Invest.* **1997**, *99*, 158–162. (b) Lasky, L. A. *Annu. Rev. Biochem.* **1995**, *64*, 113–139.
- (5) Fan, E.; Merritt, E. A.; Verlinde, C. L. M. J.; Hol, W. G. J. *Curr. Opin. Struct. Biol.* **2000**, *10*, 680–686.

(6) von Hippel, P. H. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 79–105.

(7) Rao, J.; Lahiri, J.; Weis, R. M.; Whitesides, G. M. *J. Am. Chem. Soc.* **2000**, *122*, 2698–2710.

(8) (a) Imberty, A.; Varrot, A. *Curr. Opin. Struct. Biol.* **2008**, *18*, 567–576. (b) Dam, T. K.; Brewer, C. F. In *Comprehensive Glycoscience*; Kamerling, J. P., Ed.; Elsevier: Oxford, 2007; pp 397–452. (c) Collins, B. E.; Paulson, J. C. *Curr. Opin. Chem. Biol.* **2004**, *8*, 617–625. (d) Sacchettini, J. C.; Baum, L. G.; Brewer, C. F. *Biochemistry* **2001**, *40*, 3009–3015. (e) Lee, Y. C.; Lee, R. T. *Acc. Chem. Res.* **1995**, *28*, 321–327. (f) Drickamer, K.; Taylor, M. E. *Annu. Rev. Cell Biol.* **1993**, *9*, 237–264.

(9) Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, 555–578.

such processes or as effectors of signal transduction pathways.^{8c,10} The enhancement of binding affinity toward target proteins as a result of ligand clustering can be remarkable and has been determined by different types of binding assays.⁹ In many cases, however, the molecular principles of how binding enhancement is achieved are not well understood. Several mechanisms can account for the observed binding enhancements.^{10i,11} Generally, the highest contribution is attributed to the chelate effect, that is, the spanning of adjacent binding sites by the multivalent ligand. It is obvious that the nature of the spacer between the binding epitopes, especially its conformation and flexibility, strongly influences the ligand's chelate formation capability. However, significant binding enhancements are also observed in cases where the spacer is too short to allow spanning of adjacent binding sites. In these cases, other mechanisms are assumed to be operative. One of them is statistical rebinding, that is, the rebinding of epitopes before complete dissociation of the complex due to high local concentration of epitopes. Often, cross-linking of proteins by the multivalent ligand leading to lattice formation is observed.¹² While this process has been studied in detail, the mechanisms of the often associated enhanced binding affinities are less well understood.

Structural information on multivalent carbohydrate–protein interactions with atomic resolution is rare. Cross-linking of lectins by multivalent ligands has been observed for galectin-1¹³ and concanavalin A dimers,¹⁴ monomers of the carbohydrate recognition domain of DC-SIGN,¹⁵ and BC2L-A dimers.¹⁶ The only examples of lectin structures with bound ligands that span several binding sites of a single (oligomeric) lectin stem from the AB₅ family of bacterial toxins. Kitov et al. designed a decavalent ligand, named STARFISH, for Shiga-like toxins with subnanomolar binding potency.¹⁷ Deviant from the original design, crystal structure analysis revealed that the ligand is sandwiched between two toxin B pentamer molecules. The pairs of trisaccharides at the tips of each of the five spacer arms connect pairs of opposite B subunits of the two toxin pentamers. Similar binding modes were observed for penta- and decavalent ligands for the cholera toxin, a

closely related AB₅ bacterial toxin.¹⁸ Divalent inhibitors of the cholera toxin B pentamer were shown to have high potency without the ability to span two binding sites.¹⁹

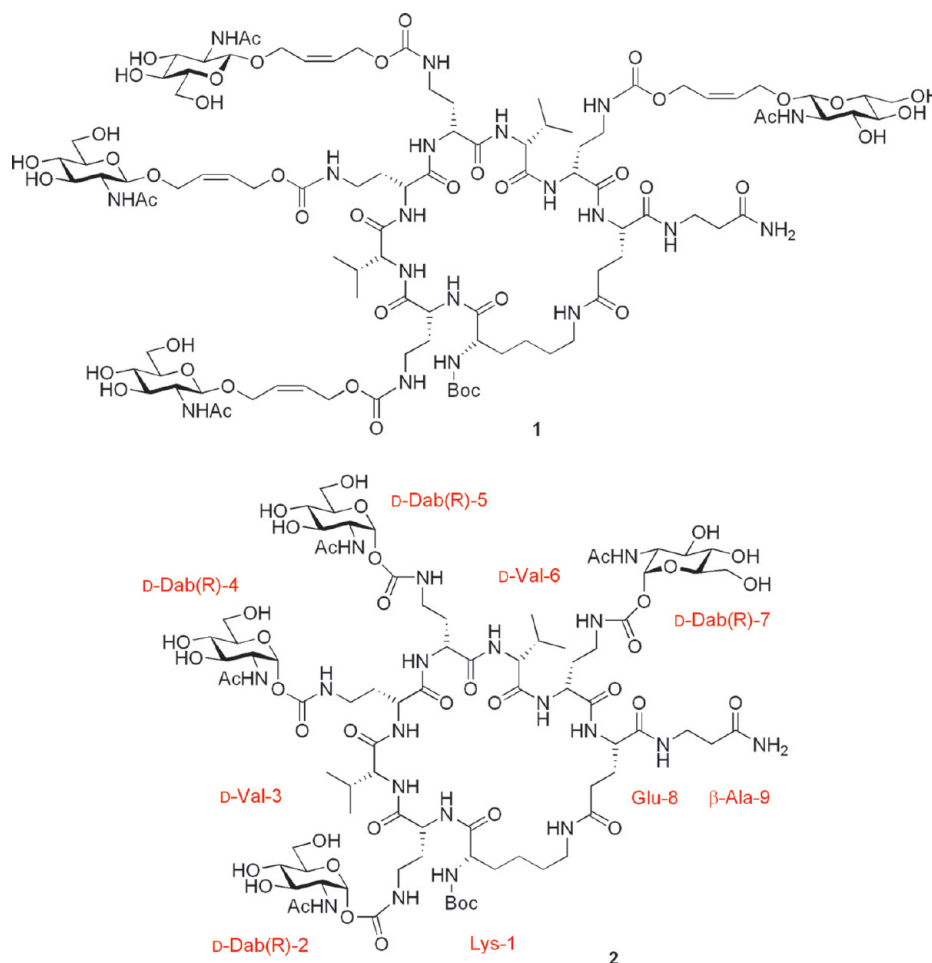
Recently, we introduced a diversity-oriented approach for the combinatorial synthesis and screening of multivalent lectin ligands with varying scaffold architectures.²⁰ In this approach, termed spatial screening of multivalent lectin ligands, cyclopeptides serve as scaffolds for the presentation of carbohydrate epitopes. From screening of a one-bead one-compound library of almost 20 000 cyclic neoglycopeptides, we were able to identify several tetra- to hexavalent wheat germ agglutinin (WGA) ligands with strongly increased binding potencies compared to monovalent *N*-acetylglucosamine (GlcNAc).^{20b} We concluded that it is the spatial presentation of the GlcNAc residues on the cyclic peptide scaffolds that is responsible for their high affinity.

WGA is a plant lectin that is enriched in the seeds of *Triticum vulgare* and exists in three closely related major isoforms, WGA1, WGA2, and WGA3. It is specific for terminal *N*-acetylneuraminic acid and GlcNAc and has been shown to inhibit fungal growth through interaction with fungal cell-wall components²¹ and to agglutinate transformed cells *in vitro*.²² WGA forms a 36-kDa stable homodimer with a twofold symmetry axis.²³ Each polypeptide chain forms four hevein domains (43 residues each), A–D. The carbohydrate binding sites were already analyzed by cocrystallization of the protein with sialyl lactose (Neu5Ac- α (2,3)-Gal- β (1,4)-Glc)²⁴ and with a sialoglycopeptide.²⁵ Furthermore, carbohydrate–protein interaction was examined by soaking of glutaraldehyde-cross-linked WGA crystals with *N,N*-diacetyl chitobiose (GlcNAc- β (1,4)-GlcNAc),^{23a,26} GlcNAc- β (1,6)-Gal,^{26b} and GlcNAc- β (1,6)-Gal- β (1,4)-Glc.^{26b} Combined evidence from these investigations revealed eight functional sugar binding sites per WGA dimer (four unique sites due to the twofold symmetry axis). However, simultaneous occupancy of all eight sites has not been observed in a single crystal structure. A binding site for GlcNAc or β -(1,4)-linked GlcNAc oligomers²⁷ is formed by each hevein domain via a cluster of three conserved aromatic residues of which the second is stacking to the sugar ring. Binding is complemented by polar residues from an adjacent domain of the other chain which provides hydrogen bonds. The sites can be adequately labeled as pairs of capitals (each followed by the number of the polypeptide chain it belongs to) indicating the hevein subunit which provides the aromatic residues and the polar residues, respectively. In D, the polar residues are absent, so

- (10) (a) *Neoglycoconjugates: Preparation and Applications*; Lee, Y. C.; Lee, R. T., Eds.; Academic Press: San Diego, CA, 1994; (b) Roy, R. *Top. Curr. Chem.* **1997**, *187*, 241–274. (c) Houseman, B. T.; Mrksich, M. *Top. Curr. Chem.* **2002**, *218*, 1–44. (d) Lindhorst, T. K. *Top. Curr. Chem.* **2002**, *218*, 201–235. (e) Ortiz Mellet, C.; Defaye, J.; García Fernández, J. M. *Chem.–Eur. J.* **2002**, *8*, 1982–1990. (f) Turnbull, W. B.; Stoddart, J. F. *Rev. Mol. Biotechnol.* **2002**, *90*, 231–255. (g) Choi, S.-K. *Synthetic Multivalent Molecules. Concepts and Biomedical Applications*; John Wiley & Sons: Hoboken, NJ, 2004; (h) Wittmann, V. In *Highlights in Bioorganic Chemistry: Methods and Applications*; Schmuck, C.; Wennemers, H., Eds.; Wiley-VCH: Weinheim, 2004; pp 203–213. (i) Kiessling, L. L.; Gestwicki, J. E.; Strong, L. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 2348–2368.
- (11) (a) Dam, T. K.; Brewer, C. F. *Biochemistry* **2008**, *47*, 8470–8476. (b) Pieters, R. J. *Org. Biomol. Chem.* **2009**, *7*, 2013–2025.
- (12) Brewer, C. F.; Miceli, M. C.; Baum, L. G. *Curr. Opin. Struct. Biol.* **2002**, *12*, 616–623.
- (13) Bourne, Y.; Bolgiano, B.; Liao, D.-I.; Strecker, G.; Cantau, P.; Herzberg, O.; Feizi, T.; Cambillau, C. *Nat. Struct. Biol.* **1994**, *1*, 863–870.
- (14) Dimick, S. M.; Powell, S. C.; McMahon, S. A.; Moothoo, D. N.; Naismith, J. H.; Toone, E. J. *J. Am. Chem. Soc.* **1999**, *121*, 10286–10296.
- (15) Feinberg, H.; Mitchell, D. A.; Drickamer, K.; Weis, W. I. *Science* **2001**, *294*, 2163–2166.
- (16) Lameignere, E.; Shiao, T. C.; Roy, R.; Wimmerova, M.; Dubreuil, F.; Varrot, A.; Imbert, A. *Glycobiology* **2010**, *20*, 87–98.
- (17) Kitov, P. I.; Sadowska, J. M.; Mulvery, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature* **2000**, *403*, 669–672.

- (18) (a) Merritt, E. A.; Zhang, Z.; Pickens, J. C.; Ahn, M.; Hol, W. G. J.; Fan, E. *J. Am. Chem. Soc.* **2002**, *124*, 8818–8824. (b) Zhang, Z.; Merritt, E. A.; Ahn, M.; Roach, C.; Hou, Z.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *J. Am. Chem. Soc.* **2002**, *124*, 12991–12998.
- (19) Pickens, J. C.; Mitchell, D. D.; Liu, J.; Tan, X.; Zhang, Z.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *Chem. Biol.* **2004**, *11*, 1205–1215.
- (20) (a) Wittmann, V.; Seeberger, S. *Angew. Chem., Int. Ed.* **2000**, *39*, 4348–4352. (b) Wittmann, V.; Seeberger, S. *Angew. Chem., Int. Ed.* **2004**, *43*, 900–903.
- (21) (a) Mirelman, D.; Galun, E.; Sharon, N.; Lotan, R. *Nature* **1975**, *256*, 414–416. (b) Laroy, K.; Weeks, G. *J. Cell Sci.* **1983**, *59*, 203–212. (c) Asensio, J. L.; Canada, F. J.; Siebert, H. C.; Laynez, J.; Poveda, A.; Nieto, P. M.; Soedjanaamadja, U. M.; Gabius, H. J.; Jimenez-Barbero, J. *Chem. Biol.* **2000**, *7*, 529–543.
- (22) (a) Aub, J. C.; Sanford, B. H.; Cote, M. N. *Proc. Natl. Acad. Sci. U.S.A.* **1965**, *54*, 396–399. (b) Aub, J. C.; Sanford, B. H.; Wang, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **1965**, *54*, 400–402.
- (23) (a) Wright, C. S. *J. Mol. Biol.* **1980**, *141*, 267–291. (b) Wright, C. S. *J. Mol. Biol.* **1989**, *209*, 475–487. (c) Harata, K.; Nagahora, H.; Jigami, Y. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1995**, *51*, 1013–1019.
- (24) Wright, C. S. *J. Mol. Biol.* **1990**, *215*, 635–651.
- (25) Wright, C. S.; Jaeger, J. *J. Mol. Biol.* **1993**, *232*, 620–638.
- (26) (a) Wright, C. S. *J. Mol. Biol.* **1984**, *178*, 91–104. (b) Muraki, M.; Ishimura, M.; Harata, K. *Biochim. Biophys. Acta* **2002**, *1569*, 10–20.
- (27) Wright, C. S.; Kellogg, G. E. *Protein Sci.* **1996**, *5*, 1466–1476.

Chart 1. Tetravalent WGA-Binding Neoglycopeptides



that the sites are A1, B1C2, C1B2, D1A2, A2, B2C1, C2B1, D2A1. The binding sites involving A and D exhibit lower affinity as those formed from B and C.²⁷ Consequently, the sites involving the B and C domains have been termed “primary” binding sites and the ones involving the A and D domains “secondary” binding sites. The secondary binding sites are generally assumed to bind carbohydrates with an affinity too weak to be detected in solution.

Out of the WGA ligands identified by our spatial screening procedure,^{20b} tetravalent neoglycopeptide **1** (Chart 1) showed an unprecedentedly strong increase in WGA binding potency by a factor of 1440 (360 per sugar residue) compared to monovalent GlcNAc as determined by an enzyme-linked lectin assay (ELLA).²⁸ In this assay, the concentration of ligand leading to an inhibition of 50% of the binding of horseradish peroxidase-labeled WGA to GlcNAc residues covalently immobilized to microtiter plates (IC_{50} value) is determined. Assuming that the spatial presentation of the GlcNAc residues on the cyclic peptide scaffold is responsible for its high affinity, variation of the length and flexibility of the spacer between the GlcNAc residues and the peptide backbone of **1** is expected to strongly affect the IC_{50} value of 16 μ M of the glycocluster.

In this report, we describe the synthesis of cyclic neoglycopeptide **2** (Chart 1) with a spacer length reduced by five bonds. In contrast to **1**, the GlcNAc residues in **2** are α -glycosidically linked to the peptide scaffold. The new ligand **2** exhibits an improved IC_{50} value

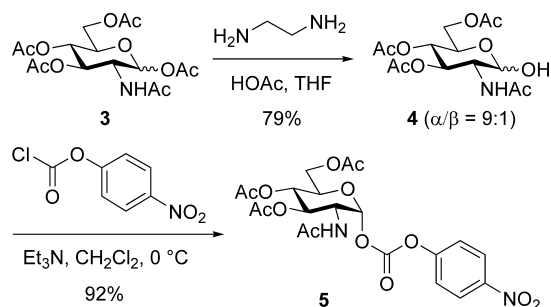
of 0.9 μ M that is 25 500-fold lower than that of GlcNAc (or 6400-fold on a valency-corrected basis). To reveal the structural basis of the enhanced WGA binding affinity of **2**, we prepared a series of mono- to trivalent GlcNAc derivatives containing the glycosyl carbamate motive found in **2** separated by different spacers. WGA binding affinities were determined by ELLA. Systematic variation of spacing identified the optimal distance between GlcNAc residues to span adjacent binding sites in WGA. The best divalent ligand has an IC_{50} value of 9.8 μ M which is even lower than that of tetravalent glycopeptide **1**. We also report a high-resolution crystal structure of the complex between a divalent ligand and the lectin. In this structure, four pairs of adjacent binding sites are bridged by four molecules of the divalent ligand. This is the first structure of WGA in which all eight binding sites are occupied. In addition, it is the first example of a crystal structure of a lectin in complex with several chelating multivalent carbohydrate ligands. Furthermore, we report the crystal structure of the complex of tetravalent neoglycopeptide **2** and WGA obtained from a soaking experiment. Comparison of the peptide conformation seen in the complex with the NMR-derived solution structure of **2** suggests an explanation for the high affinity of the glycopeptide.

Results and Discussion

Synthesis of Ligands. Glycosyl *p*-nitrophenyl carbonates are suitable precursors for bioconjugation of carbohydrates.²⁹ Their reaction with amines is fast, produces glycosyl carbamates in high yields, and proceeds under retention of configuration at the anomeric center. Thus, glycosyl *p*-nitrophenyl carbonates

(28) Maierhofer, C.; Rohmer, K.; Wittmann, V. *Bioorg. Med. Chem.* **2007**, *15*, 7661–7676.

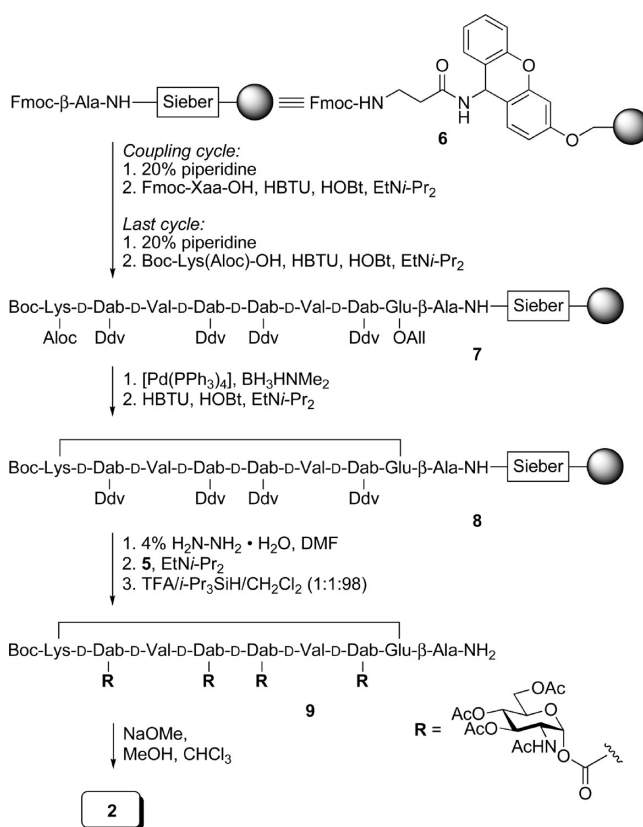
Scheme 1. Synthesis of Glycosyl Carbonate 5



are excellent precursors for the stereoselective synthesis of *O*-glycosyl carbamates, provided they are available as pure isomers. Recently, we could show that both α - and β -lactosyl *p*-nitrophenyl carbonates are accessible as pure isomers by proper choice of reaction conditions.^{29e} We now report conditions for the stereoselective synthesis of GlcNAc-derived *p*-nitrophenyl carbonate **5** in high yield (Scheme 1). Peracetylated GlcNAc **3** was selectively deprotected at the anomeric center using the method of Kováč.³⁰ Resulting **4** exists according to ¹H NMR analysis (CDCl₃) as mixture of anomers ($\alpha/\beta = 9:1$). Treatment with *p*-nitrophenyl chloroformate and triethylamine in dichloromethane exclusively led to α -glycosyl carbonate **5** in a yield of 92%. Since **5** slowly decomposes during column chromatography on silica gel, it was essential to use a short column and avoid prolonged contact of **5** with the silica gel.

To probe the effect of the length and flexibility of the spacer between the GlcNAc residues and the peptide backbone of **1** on the WGA binding affinity, we synthesized neoglycopeptide **2**. The carbohydrate-peptide spacer of **2** is shortened by five bonds and also is less flexible due to the missing butene-diol moiety. Furthermore, the GlcNAc residues are α -glycosidically linked to the peptide scaffold leading to a different spatial orientation of the sugars. It is known that WGA binds both α - and β -glycosides of GlcNAc with similar affinity.³¹ Synthesis of **2** started from TentaGel resin **6** loaded with β -alanine via the Sieber linker³² (Scheme 2). Assembly of linear peptide **7** followed the Fmoc strategy;³³ solely in the last coupling step an *N*^α-Boc-protected amino acid (Boc-Lys(Aloc)-OH) was applied. Removal of the Aloc group and cleavage of the allyl ester were achieved by treatment with [Pd(PPh₃)₄] and borane dimethylamine complex.³⁴ Subsequent addition of a mixture of 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU),³⁵ 1-hydroxybenzotriazole (HOBt), and Hünig's base (Et₃Ni-Pr₂) gave cyclic peptide **8**.

Scheme 2. Synthesis of Neoglycopeptide 2



We found that usage of resin with a low loading (0.2 mmol g⁻¹) is crucial during the cyclization step. Only in this way, it was possible to obtain high yields of cyclic peptide as determined by a test cleavage from a small sample of resin **8**. With higher resin loadings, we were not able to cleave any material from **8** by treatment with 1% TFA. A plausible explanation for this observation is the “intermolecular” reaction of peptide molecules during the cyclization step leading to a solid phase-bound peptide polymer. At lower resin loadings, cyclization of single linear peptides is preferred, an observation which is referred to as the pseudodilution phenomenon.³⁶ Cleavage of the 1-(4,4-dimethyl-2,6-dioxocyclohexylidene)isovaleryl (Ddv) groups³⁷ that were used to mask the γ -NH₂ groups of *D*-diaminobutyric acid (*D*-Dab) was effected by batch-wise treatment with a 4% solution of hydrazine hydrate which turned out^{20a} to be advantageous over the literature-recommended³⁷ 2% solution. Subsequently, the GlcNAc residues were attached by addition of glycosyl carbonate **5** (3 equiv. per free NH₂ group) in the presence of Hünig's base until Kaiser³⁸ and TNBS test³⁹ indicated the absence of free amino groups. Finally, the peptide was cleaved from the resin with 1% trifluoroacetic acid (TFA) and deacetylated to give **2**. It is important to carry out the deacetylation after cleavage from the resin because the glycosidic bond of the deacetylated sugars is not completely

(29) (a) Madec-Lougerstay, R.; Florent, J.-C.; Monneret, C. *J. Chem. Soc., Perkin Trans. 1* **1999**, 1369–1376. (b) Azoulay, M.; Escriou, V.; Florent, J.-C.; Monneret, C. *J. Carbohydr. Chem.* **2001**, *20*, 841–853. (c) Cherif, S.; Leach, M. R.; Williams, D. B.; Monneret, C. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1237–1240. (d) Cai, T. B.; Lu, D.; Tang, X.; Zhang, Y.; Landerholm, M.; Wang, P. G. *J. Org. Chem.* **2005**, *70*, 3518–3524. (e) André, S.; Specker, D.; Bovin, N. V.; Lensch, M.; Kaltner, H.; Gabius, H.-J.; Wittmann, V. *Bioconjugate Chem.* **2009**, *20*, 1716–1728.

(30) Zhang, J.; Kováč, P. *J. Carbohydr. Chem.* **1999**, *18*, 461–469.

(31) Allen, A. K.; Neuberger, A.; Sharon, N. *Biochem. J.* **1973**, *131*, 155–162.

(32) Sieber, P. *Tetrahedron Lett.* **1987**, *28*, 2107–2110.

(33) Wellings, D. A.; Atherton, E. *Methods Enzymol.* **1997**, *289*, 44–67.

(34) (a) Gomez-Martinez, P.; Dessolin, M.; Guibé, F.; Albericio, F. *J. Chem. Soc., Perkin Trans. 1* **1999**, 2871–2874. (b) Jarikote, D. V.; Köhler, O.; Socher, E.; Seitz, O. *Eur. J. Org. Chem.* **2005**, *318*, 7–3195.

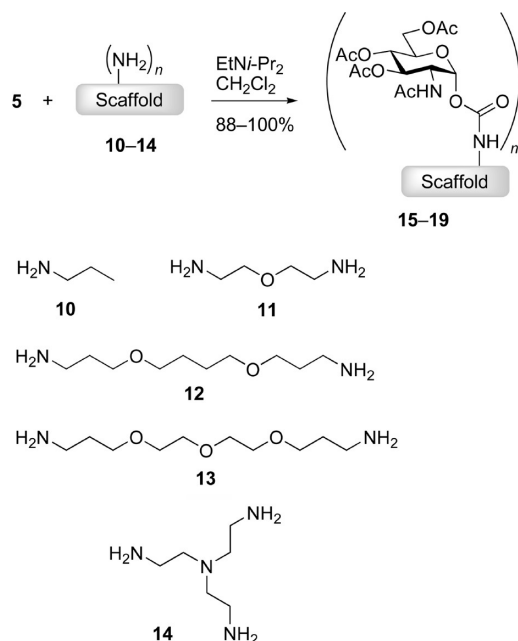
(35) Knorr, R.; Trzeciak, A.; Bannwarth, W.; Gillissen, D. *Tetrahedron Lett.* **1989**, *30*, 1927–1930.

(36) (a) Mazur, S.; Jayalekshmy, P. *J. Am. Chem. Soc.* **1979**, *101*, 677–683. (b) Albericio, F.; Hammer, R. P.; Garcia-Echeverria, C.; A., M. M.; Chang, J. L.; Munson, M. C.; Pons, M.; Giralt, E.; Barany, G. *Int. J. Pept. Protein Res.* **1991**, *37*, 402–413.

(37) Chhabra, S. R.; Hothi, B.; Evans, D. J.; White, P. D.; Bycroft, B. W.; Chan, W. C. *Tetrahedron Lett.* **1998**, *39*, 1603–1606.

(38) Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595–598.

(39) Hancock, W. S.; Battersby, J. E. *Anal. Biochem.* **1976**, *71*, 260–264.

Scheme 3. Conjugation of Glycosyl Carbonate **5** to Amines **10–14**

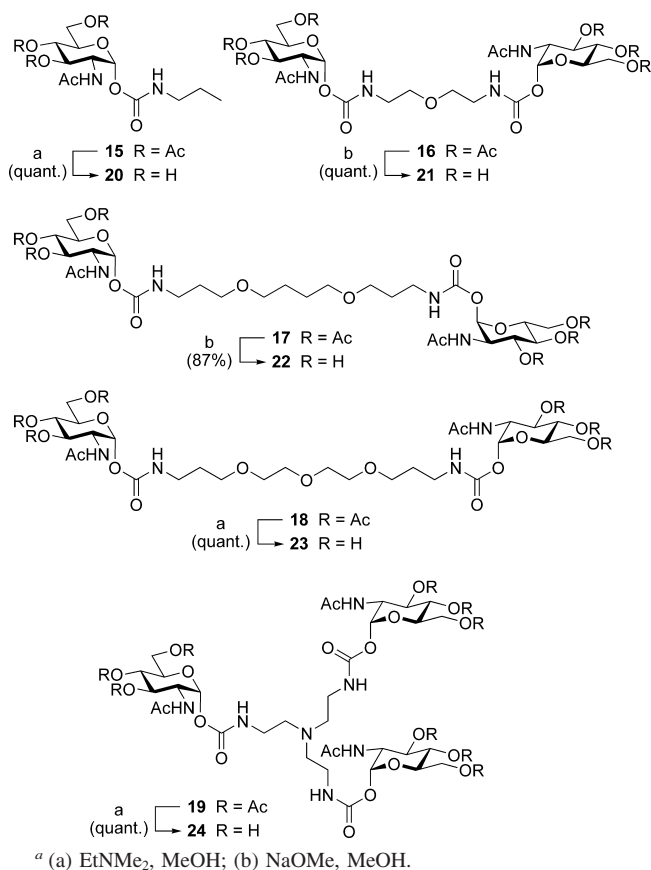
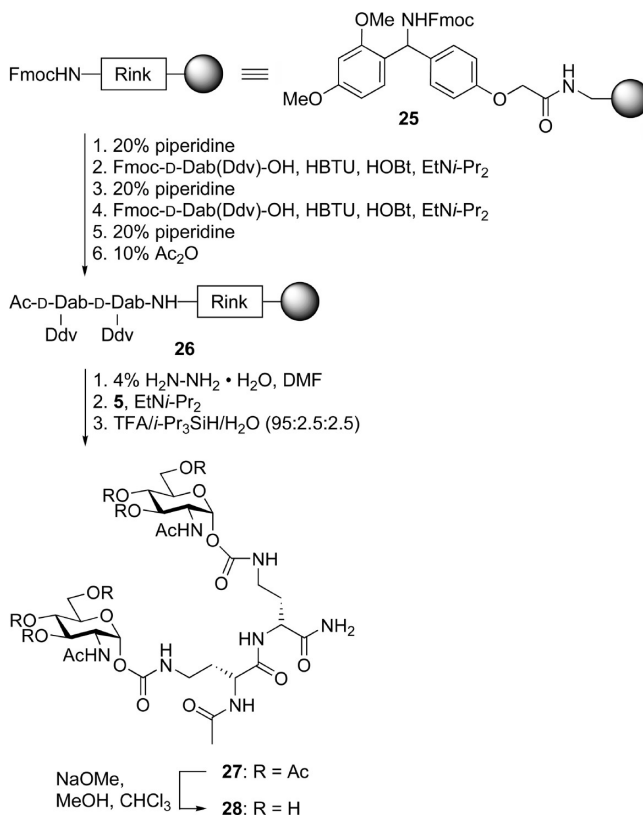
stable under the acidic conditions used for cleavage of the peptide from the resin.^{20a,40}

A series of mono- to trivalent GlcNAc derivatives containing the α -glycosyl carbamate motif found in **2** was obtained by conjugation of carbonate **5** to commercially available amines **10–14** (Scheme 3). Couplings were carried out in dichloromethane in the presence of Hünig's base to give carbamates **15–19** in very good yields. Deprotection by treatment with EtNMe_2 in MeOH or under Zemplén conditions⁴¹ yielded glycoclusters **20–24** (Scheme 4).

As detailed below, binding of neoglycopeptide **2** to WGA occurs by spanning adjacent binding sites through the two GlcNAc residues bound to D-Dab residues in position 4 and 5. Therefore, we also prepared the divalent glycopeptide **28** representing this substructure of **2** (Scheme 5). Starting from Fmoc-protected Rink Amide polystyrene resin **25**, two D-Dab residues were coupled by the Fmoc strategy followed by replacement of the N-terminal Fmoc group with an acetyl group. Ddv removal, sugar conjugation, and cleavage from the resin with concentrated TFA gave divalent glycopeptide **27** which was deacetylated to yield **28**.

Binding Assays. WGA binding potencies of the synthetic ligands were measured in terms of IC_{50} values for the inhibition of binding of horseradish peroxidase-labeled WGA to a GlcNAc derivative covalently immobilized to microtiter plates using an ELLA developed earlier.²⁸ As listed in Table 1, monovalent ligand **20** has an affinity similar to that of β -allyl glycoside **29** within the margin of error of the ELLA. Therefore, we conclude that it is not the anomeric configuration or the glycosyl carbamate structure that is responsible for the strongly increased binding affinity of glycocluster **2** compared to **1** but rather the altered spatial presentation of the sugars on the peptide scaffold and/or the reduced flexibility of **2**.

The influence of sugar spacing on the WGA affinity is clearly seen by comparison of divalent ligands **21–23**. Whereas divalent

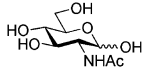
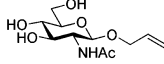
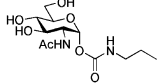
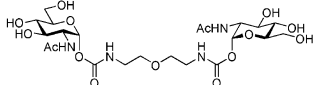
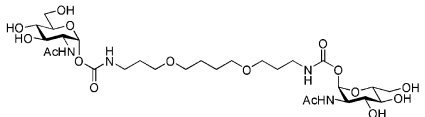
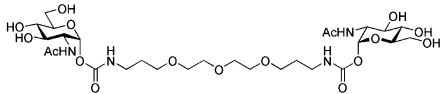
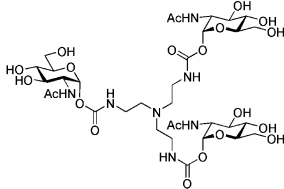
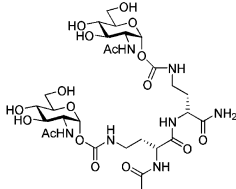
Scheme 4. Synthesis of Glycoclusters **20–24**^aScheme 5. Synthesis of Divalent Neoglycopeptide **28**

ligand **21** showed only 31-fold affinity increase over GlcNAc (11-fold increase over corresponding monovalent **20**), incorpo-

(40) Kunz, H.; Unverzagt, C. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1697–1699.

(41) Zemplén, G.; Pacsu, E. *Ber. Dtsch. Chem. Ges.* **1929**, *62*, 1613–1614.

Table 1. IC₅₀ Values of Synthetic Ligands for Inhibition of Binding of Horseradish Peroxidase-Labeled WGA to GlcNAc Residues Covalently Immobilized to Microtiter Plates from Dose-Response Curves Shown in Figures S1–S3 Determined by an ELLA²⁸

compound (valency)	IC ₅₀ / μM	rel. potency (valency-corrected) ^a	
	GlcNAc	23000 ± 2100	1
	19 (mono)	10000 ± 900	2.3
	20 (mono)	8000 ± 1000	2.9
	21 (di)	734 ± 103	31 (16)
	22 (di)	9.8 ± 2.1	2350 (1170)
	23 (di)	57 ± 13	400 (200)
	24 (tri)	812 ± 150	28 (9.4)
	28 (di)	191 ± 24	120 (60)
Cyclopeptide 1 (tetra)		16 ± 3.3	1440 (360) ²⁸
Cyclopeptide 2 (tetra)		0.9 ± 0.16	25500 (6400)

^a Relative potencies are referenced to GlcNAc.

ration of 7 additional bonds in the spacer results in **22** with a remarkable IC₅₀ value of 9.8 μM corresponding to a relative potency of 2350 over GlcNAc (816 over **20**). Further, increase of the spacer length by only one additional bond leads to a drop of relative potency to 400 over GlcNAc (140 over **20**) for ligand **23** which is, however, still unprecedentedly high for a divalent WGA ligand.^{11b} The potency of **23** is even better than that of trivalent **24** which has the same spacer length between two GlcNAc moieties as **21**. Possibly, this linker length is not sufficient to span adjacent binding sites of the lectin. To gain more insight into the structural details of multivalent binding to WGA, we determined the crystal structure of several WGA–ligand complexes. All structures were solved by molecular replacement.

Crystal Structure of WGA1 in Complex with Propyl Carbamate 20. We first determined the structure of WGA1 in complex with propyl carbamate **20** by cocrystallization (WGA1~**20**, PDB ID: 2UWG). This structure served as a reference for structures of complexes of WGA with multivalent ligands by allowing to assess whether the scaffolds of the multivalent ligands have any unfavorable impact on the sugar binding mode. The crystals belonged to space group *P3₂21* and diffracted to 1.6 Å resolution (Table 2). The structure was refined to an *R* (*R*_{free}) factor of 19.6% (23.2%). One WGA1 homodimer was found in the asymmetric unit with a conformation identical to unliganded WGA1 (PDB ID: 7WGA²⁴). All high affinity sites B1C2, C1B2, B2C1, and C2B1 of the WGA1 dimer were found occupied with one molecule of **20** (Figure

Table 2. Crystallographic Data

	WGA1~20	WGA3~23	WGA1~2
PDB ID	2UWG	2X52	2X3T
X-ray source	SLS X06SA	ESRF ID23-1	ESRF ID23-1
X-ray wavelength [Å]	1	1	1
Space group	<i>P</i> 3 ₂ 21	<i>R</i> 3	<i>P</i> 2 ₁
Cell dimensions			
<i>a</i> [Å]	63.2	101.4	44.6
<i>b</i> [Å]	63.2	101.4	93.2
<i>c</i> [Å]	153.4	145.0	92.5
α [deg]	90.0	90.0	90.0
β [deg]	90.0	90.0	98.3
γ [deg]	120.0	120.0	90.0
Resolution [Å]	44.59–1.60 (1.65–1.60) ^a	20–1.70 (1.75–1.70)	50–2.75 (2.91–2.75)
Number of measured reflections	438743 (30984)	315384 (12253)	73664 (11828)
Number of unique reflections	46032 (3729)	60413 (4459)	19531 (3124)
<i>I</i> / σ (<i>I</i>)	17.38 (3.52)	11.54 (2.06)	12.10 (3.45)
Completeness [%]	96.1 (89.9)	98.9 (87.1)	99.5 (99.2)
R _{meas} [%] ^b	6.6 (62.0)	10.2 (62.3)	10.3 (49.2)
Wilson B-factor [Å ²]	33	26	54
Twining			
Twin law		$-h-k,k,-l$	
Twin fraction		0.44	
Refinement			
Number of reflections used	43465	57320	19531
Number of protein atoms	2307	2330	4521
Number of ligand atoms	87	182	261
Number of water molecules	245	178	38
Final <i>R</i> -factor [%]	19.6	15.8	22.9
Final free <i>R</i> -factor [%]	23.2	17.1	28.5
Ramachandran plot			
Favored [%]	97.3	97.9	90.3
Outliers [%]	0.0	0.0	0.4
Rms deviation			
Bond lengths [Å]	0.012	0.006	0.002
Bond angles [deg]	1.46	1.04	0.72
Average B-factor [Å ²]			
Protein	31.4	23.7	54.7
Ligand	32.0	25.8	66.6
Water	32.4	28.4	43.7

^a All values in parentheses apply for the highest resolution shell. ^b According to Diederichs and Karplus⁴³

1A and Figure S4A). Because of a crystal contact, site A1 could not accommodate a ligand. At the low affinity sites D1A2, A2, and D2A1, blobs of $F_o - F_c$ difference electron density were found after molecular replacement, too big to be water. However, fully occupied molecules of **20** could not be placed there unambiguously. Nevertheless, in cocrystallization experiments with higher concentrations of **20** (50 mM), all sites except A1 were found occupied (data not shown). Comparison with the structures of the complexes of WGA3 with GlcNAc- β (1,4)-GlcNAc^{26b} and WGA1 with GlcNAc⁴² shows that binding of **20** to WGA1 is almost identical to GlcNAc binding (Figure 1B).

Crystal Structure of WGA3 in Complex with Divalent Ligand 23. WGA3 in complex with divalent **23** (WGA3~**23**, PDB ID: 2X52) crystallized in space group *R*3. Crystals diffracted to 1.7 Å (Table 2). The single crystal used for data collection was merohedrally twinned, with a twinning fraction of 0.44. The model could be refined to a final *R* (R_{free}) factor of 15.8% (17.1%). One WGA3 dimer was found in the asymmetric

unit in identical conformation as in WGA1~**20**, except for the minor sequence differences between the WGA isoforms 1 and 3.

All eight proposed carbohydrate binding sites of the WGA3 dimer²⁷ were found occupied by the GlcNAc-carbamate moieties of four molecules of **23** with each divalent ligand spanning pairs of adjacent binding sites (Figure 2 and Figure S4B). This shows for the first time that all sugar binding sites of the WGA dimer are simultaneously functional. The arrangement of the GlcNAc moieties in the sites is the same as in the WGA1~**20** structure. Two divalent ligands bind to the high-affinity site pairs B1C2–C2B1 and B2C1–C1B2, respectively. The conformation of the linker region of these ligands is well-defined but differs for the two ligands (Figure 3). The differences are likely due to participation of the former ligand in a crystal contact (Figure 3A). Residues Arg-139 and Gln-6 from polypeptide chain 1 of a crystallographically related WGA dimer form hydrogen bonds to carbamate oxygens at both ends of the linker region. Being constrained in this way, the linker region adopts a distinct conformation, which allows interaction with Trp-107 of chain 2 via water coordination. The linker region of the divalent ligand at site pair B2C1–C1B2 (Figure 3B) does not interact with

(42) Schwefel, D.; Wittmann, V.; Diederichs, K.; Welte, W. PDB ID: 2UVO (WGA1 in complex with GlcNAc), to be published.

(43) Diederichs, K.; Karplus, P. A. *Nat. Struct. Biol.* **1997**, *4*, 269–275.

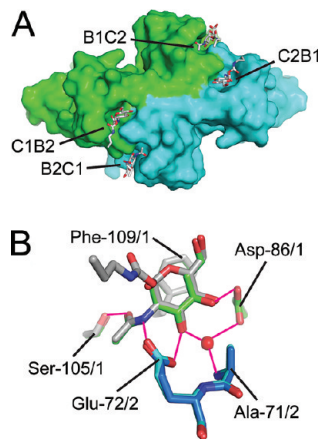


Figure 1. Crystal structure of WGA1 in complex with propyl carbamate **20** (PDB ID: 2UWG). (A) The protein is shown as surface representation. WGA1 chain 1 is colored green, chain 2 cyan. Compound **20** is shown in stick representation. Carbon atoms are colored white, oxygens red, and nitrogens blue. (B) Comparison of the binding mode of compound **20** to the GlcNAc binding mode (PDB ID: 2UVO⁴²) at site C1B2. **20** and interacting WGA chain 1 residues are colored gray; interacting WGA chain 2 residues are dark blue. GlcNAc and interacting WGA chain 1 residues are colored green; interacting WGA chain 2 residues are light blue. The digits behind the residue numbers designate the WGA chains 1 and 2. Hydrogen bonds are indicated as magenta lines.

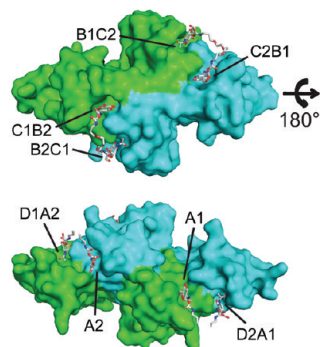


Figure 2. Crystal structure of WGA3~**23** (PDB ID: 2X52). The protein is shown as surface representation. WGA chain 1 is colored green, chain 2 is cyan. The divalent ligand is shown as stick model with carbons colored white, nitrogens colored blue, and oxygens in red.

protein residues at all. Its fixation via carbohydrate–WGA interaction is apparently sufficient to define its conformation. This conformation should, thus, be relevant for protein–ligand interaction in solution. Linker regions from ligands bridging low affinity binding sites (Ax, DxAy) show only weak electron

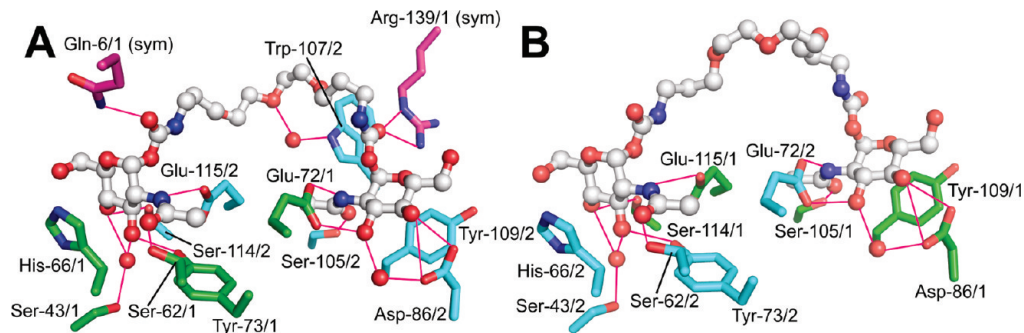


Figure 3. Detailed view of divalent ligand **23** bound at (A) binding site pair B1C2–C2B1 and (B) B2C1–C1B2 of WGA3. Compound **23** is shown in ball-and-stick representation with carbon atoms colored white, nitrogens blue, and oxygens red. Interacting protein side chains are shown in stick representation. Residues from chain 1 are colored green, those from chain 2 are cyan. Hydrogen bonds are indicated as magenta lines. The digits behind the residue numbers designate the WGA chains 1 and 2, respectively, and “sym” indicates residues from a crystallographically related WGA dimer.

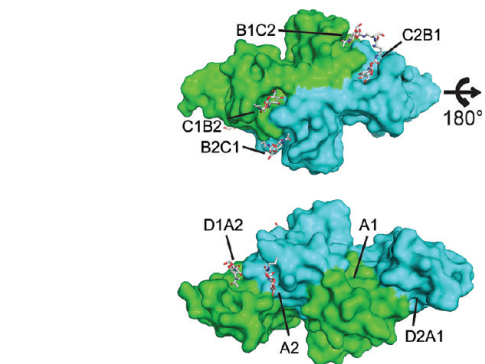


Figure 4. Portions of cyclic peptide **2** bound to WGA1 (PDB ID: 2X3T), which are visible in the crystal structure. The protein is shown as surface representation. WGA chain 1 is colored green, chain 2 is cyan. Ligands are shown as stick model with carbons colored white, nitrogens colored blue, and oxygens in red.

density. In this case, the middle part of the linker could not be modeled (Figure 2). This indicates higher static or dynamic disorder of these linker regions, possibly due to the smaller binding affinity of the GlcNAc–carbamate moieties compared to ligands at high-affinity binding sites. Nevertheless, intermolecular bridging to symmetry-related bindings sites of adjacent WGA dimers can be excluded because of the arrangement of binding sites.

Crystal Structure of WGA1 in Complex with Tetravalent Cyclic Neoglycopeptide 2 (WGA1~2**, PDB ID: 2X3T).** Glutaraldehyde-cross-linked WGA1 crystals were soaked with cyclic tetravalent glycopeptide **2**. The crystals belonged to space group $P2_1$ and diffracted to 2.75 Å. The structure was refined to a final R (R_{free}) factor of 22.9% (28.5%) (Table 2). Two WGA1 dimers were found in the asymmetric unit. The polypeptide chains of the first dimer were designated as 1 and 2, the polypeptides of the second as 3 and 4. Eleven out of 16 possible sugar binding sites of the two dimers were found occupied. The remaining five sites were obstructed by crystal contacts. At the occupied sites clear electron density extended beyond the anomeric oxygen of the GlcNAc moiety and the complete GlcNAc–carbamate moiety could be modeled. Moreover, at two pairs of the occupied sites, namely, B1C2–C2B1 and B3C4–C4B3, the complete glycosylated D-Dab residues 4 and 5 of two cyclic neoglycopeptides **2** are visible (Figure 4 and Figure S4C). These glycosylated dipeptide subunits of **2** bridge the site pairs in a similar way as does the divalent ligand in complex WGA3~**23** (Figure 5). The conformation of both of these dipeptide subunits is identical within the limits of experimental data. Further

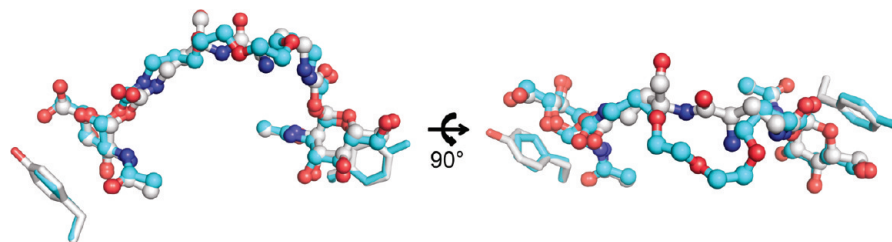


Figure 5. Comparison of binding modes of divalent ligand **23** (cyan, site pair B2C1–C1B2) and tetravalent cyclic peptide ligand **2** (white, site pair B1C2–C2B1). The ligands are shown in ball-and-stick representation, with carbon atoms colored cyan and white, respectively, nitrogens blue, and oxygens red. We compare ligands at different binding site pairs because **23** bound to site pair B1C2–C2B1 is compromised by crystallographic contacts (cf. Figure 3A). Nevertheless, this comparison is valid, because site pairs B1C2–C2B1 and B2C1–C1B2 are structurally equivalent due to pseudo-twofold rotational symmetry of the WGA dimer.

modeling of the cyclic peptides is impossible indicating that a multiplicity of conformations flattens the electron density map. The binding mode of the ligand's GlcNAc-carbamate moieties is identical to the binding mode of glycosyl carbamate **20** and the GlcNAc-carbamate moieties of divalent compound **23**. There are no interactions between the protein and the carbamate groups or the peptide part of the ligand.

Comparison of Linker Lengths in Spanning Ligands. The divalent ligands **21**, **22**, and **23** differ in their spacer length with 10, 17, and 18 single bonds, respectively, between the anomeric oxygens. The somewhat meandering appearance of the linker region seen at the high affinity binding site pairs in complex WGA3~**23** and the missing electron density for the linker at the low affinity site pairs suggest that the linker is slightly too long. This is supported by the significantly higher binding potency of compound **22**, whose linker is one bond length shorter than in compound **23** (Table 1). Superposition of the structures of the linker of bound divalent **23** and the bridging dipeptide subunit of compound **2** reveals that the dipeptide fits even better to the architecture of the binding site pair (Figure 5). This could contribute to the increased affinity of the cyclic peptide ligand compared to the divalent ligands **23** and **22**. Two further glycosylated D-Dab residues are available in positions 2 and 7 of tetravalent peptide **2**, which are separated from the WGA-binding sugar pair (residues 4 and 5) by D-valine residues. In the preformed environment of the cross-linked crystal, these cannot access binding sites. In solution, however, they may further enhance the binding affinity of **2** by interacting with additional WGA molecules. Indeed, addition of the tetravalent cyclic peptide ligand **2** to WGA solutions resulted in precipitation of the protein which indicates cross-linking of WGA dimers to higher oligomers.

According to Table 1, the shortest of the divalent ligands, **21**, apparently does not allow high-affinity chelating binding to WGA. This is supported by molecular modeling studies. Although it was possible to fit **21** in the crystallographically determined WGA structure in a mode bridging adjacent high-affinity sites without obvious conformational strain, even short molecular dynamic simulations (Tripos 5.2 forcefield⁴⁴) resulted in dissociation of one of the GlcNAc residues. Under the same conditions, the positions of the GlcNAc moieties of divalently bound **23** changed only marginally over a period of 50 ps (data not shown).

To test whether the substructure of **2** that is visible in the crystal structure (i.e., residues 4 and 5) alone is sufficient for high-affinity binding to WGA, we synthesized divalent dipeptide

28 (Scheme 5) and subjected it to the ELLA. An IC_{50} value of 191 μ M (rel. potency 120 over GlcNAc, 42 over **20**) was determined, being even worse than that of divalent ligand **23** (Table 1). Dipeptide **28** differs from tetravalent ligand **2** not only by lacking two sugar residues, but also in being more flexible. Possibly, the conformation of the cyclic peptide scaffold orients the side chains of the glycosylated D-Dab residues 4 and 5 in a way facilitating bivalent binding to WGA. To test this hypothesis, we determined the conformation of **2** in solution by NMR analysis and compared this structure with the fragment of **2** that is visible in the crystal structure.

Conformational Analysis of 2 by NMR Spectroscopy. The conformational analysis of **2** was based on a complete NMR resonance assignment which was obtained from homonuclear DQF-COSY, TOCSY, and NOESY, as well as heteronuclear 1H - ^{13}C -HSQC spectra (Table S1). TOCSY amide traces of **2** (Figure 6A) are well dispersed leading to unambiguous assignments of all backbone H^N and H^α protons as well as side chain resonances of Lys-1, Val-3, Val-6, Glu-8, and β -Ala-9. However, the side chain resonances of the D-Dab amino acids and the signals of the four GlcNAc residues are heavily overlapped such that in most cases only average values could be determined. This indicates that the sugar-bearing D-Dab side chains and the sugar moieties themselves are, in the absence of WGA, solvent accessible, largely unrestricted, and sample many conformations which leads to similar chemical shifts for all four glycosylated side chains. A similar absence of chemical shift dispersion is also observed for the side chains of Lys-1 and Glu-8 that form the lactam bridge of the cyclopeptide suggesting that this part of the molecule has to be considered flexible, as well. In contrast to this, the chemical shift dispersion of the backbone resonances of **2** allowed for extracting a number of structural restraints.

A total of 25 structurally relevant, unambiguous interproton distances were derived from analyzing cross peak volumes of 2D NOESY spectra recorded with a mixing time of 50 ms (Table S2). Four of these were long-range (i.e., nonsequential) NOE contacts. In addition, four ambiguous distance constraints could be identified. Furthermore, the 3D structure of **2** is defined by four ϕ and five χ_1 torsion angle constraints that were obtained from $^3J_{HNH\alpha}$ and $^3J_{HaH\beta}$ coupling constants (Tables S3 and S4). In the case of ϕ torsion angles, direct J -coupling restraints were employed only if the $^3J_{HNH\alpha}$ coupling was >8 Hz resulting in only two similar solutions of the Karplus equation.⁴⁵ According to the presence of, in each case, one large and one small $^3J_{HaH\beta}$ coupling constant with sums of 15–18 Hz, the χ_1 angles of the D-Dab and Glu amino acids were restrained such that the g^- rotamer was excluded in the D-Dab residues and the g^+ rotamer

(44) Clark, M.; Cramer, R. D., III; Van Opdenbosch, N. *J. Comput. Chem.* **1989**, *10*, 982–1012.

(45) Karplus, M. *J. Am. Chem. Soc.* **1963**, *85*, 2870–2871.

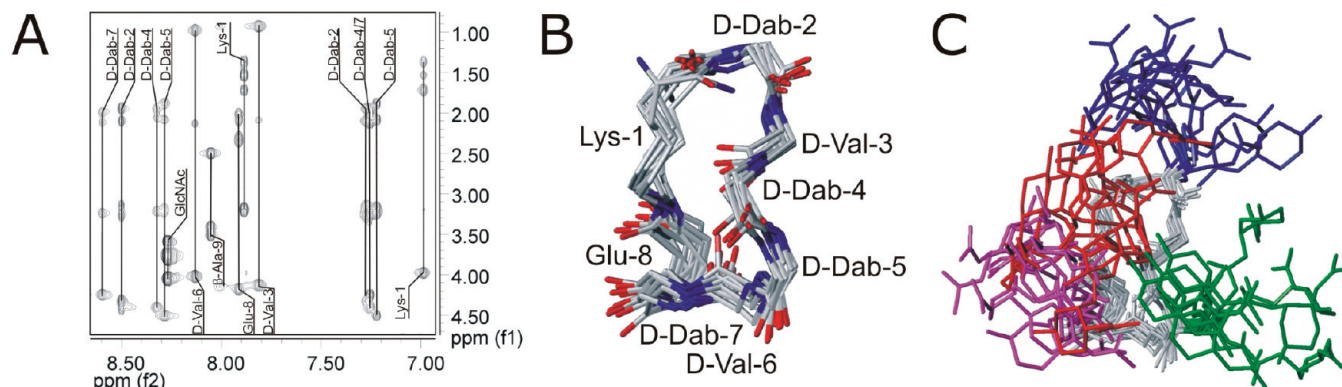


Figure 6. (A) Amide region of a 600 MHz TOCSY spectrum of **2** in H₂O/D₂O at 300 K. (B) Ensemble of the dominant family of conformers of **2**. (C) Orientations and accessible volume of sugar-substituted side chains of the ensemble (blue, D-Dab-2; red, D-Dab-4; green, D-Dab-5; magenta, D-Dab-7).

Table 3. Structure Determination Statistics of Neoglycopeptide **2**

A. NMR restraints			
Total unambiguous distance restraints ^a		25	
Sequential (<i>i,i</i> +1)		21	
Long-range (<i>l</i> _i − <i>j</i> > 1)		4	
Total ambiguous distance restraints		4	
Total dihedral angle restraints		9	
Φ		4	
χ ₁		5	
B. Ensemble statistics (20 structures)			
Violation analysis			
Maximum distance violation (Å)		0.28	
Maximum dihedral angle violation (deg)		<5	
Energies			
Mean restraint violation energy (kcal mol ^{−1})		2.534	
Mean AMBER energy (kcal mol ^{−1})		−510.18	
Mean deviation from ideal covalent geometry			
Bond lengths (Å)		0.0121	
Bond angles (deg.)		2.18	
<hr/>			
conformational family	I	II	III
Number of conformers	9	6	5
Average rmsd from the conformer with lowest restraint violation energy within each conformational family (backbone C, N, and O atoms) (Å)	0.65	1.12	0.71

^a Distance restraints that refer to nonstereospecifically assigned protons are included here.

in the Glu residue. Long range and dihedral angle constraints are summarized in Figure S5. We used the dedicated biopolymer force field of the AMBER software package⁴⁶ that provides highly versatile options for NMR-restrained structure calculations. Twenty conformers with low combined restraint violation were selected for further analysis. The statistics of the final structural ensemble is summarized in Table 3.

Description of the Structural Ensemble. The ensemble of conformers of neoglycopeptide **2** can be grouped into three conformational families consisting of 9, 6, and 5 conformers, respectively, according to backbone rmsd values with respect to conformers with lowest restraint violation energy. The conformers of the dominant conformational family superimpose with an average rmsd of 0.65 Å with the member having the lowest restraint violations (Figure 6B). All amino acids adopt allowed backbone dihedral angles when considering that the energetically favorable regions in the Ramachandran diagram

of D-amino acids are mirror images of those of L-amino acids (Figure S6). Neoglycopeptide **2** adopts a triangular shape with turns at D-Dab-2, D-Dab-5, and Glu-8. The region of highest structural similarity among all three families of conformers appears to range from D-Dab-2 to D-Val-6. Here, dihedral angle combinations are present that correspond to α_R, β, β, α_R, and β secondary structure, respectively, of peptides composed of L-amino acids. Stretches with the same secondary structure extending over more than two residues as well as turn conformations with distinct hydrogen bonding pattern are lacking.

Comparison to X-ray Structure. The X-ray structure of WGA1 in complex with neoglycopeptide **2** contains little information about the conformation of the cyclopeptide backbone. Only the fragment ranging from the amide nitrogen of D-Dab-4 to the carbonyl of D-Dab-5 is resolved in the electron density map. On the other hand, the solution structure does not report on well-defined conformations of the sugar-decorated side chains. They are conformationally unrestricted when the protein is not present (Figure 6C). Thus, the two methods provide complementary information about the structure. To find out

(46) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

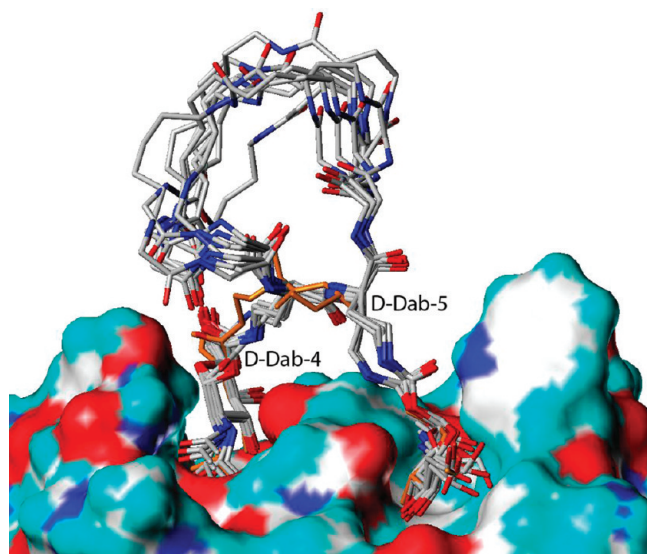


Figure 7. Superposition of the dominant conformational family of neoglycopeptide **2** determined by solution NMR spectroscopy (atom color) onto the resolved atoms of the ligand in the crystal (orange). The heavy atoms of the sugar-bearing side chains and the D-Dab-4–D-Dab-5 peptide backbone were fitted by minimizing their distances to the X-ray structure in an MD simulation (see text).

whether the solution conformation of **2**, as determined by NMR, is compatible with its WGA-bound conformation determined by X-ray crystallography, we superimposed both independent structures.

The solution structure was annealed onto the crystal structure by employing distance restraints ($1 \text{ kcal } \text{Å}^{-2}$) from all resolved C, N, and O atoms of the ligand fragment to the corresponding atoms of the NMR structure. These distances were minimized during short molecular dynamics simulations and subsequent energy minimization. The backbone conformation of the cyclopeptide was preserved during this fitting procedure by strong dihedral angle restraints (20 kcal deg^{-2}). The members of the dominant NMR ensemble achieve an average rmsd of 1.31 Å without significant conformational strain of the GlcNAc decorated side chains documenting that the decisive fragment can adopt a similar conformation in solution and in complex with WGA (Figure 7). This implies that cyclopeptide **2** is already preorganized in solution in a conformation suitable for the interaction of two GlcNAc residues with the protein thereby contributing to the high affinity of **2**.

The structures of the second family of conformers superimpose with the resolved fragment in the crystal with an average rmsd of 1.22 Å again showing natural side chain conformations. The third conformational family would also allow a similar binding mode as found in the crystal (rmsd 1.30 Å); however, one member of this ensemble would be affected by steric clashing into the protein.

Conclusions

In conclusion, we have synthesized a series of multivalent WGA ligands and studied their interaction with the protein by an enzyme-linked lectin assay and X-ray crystallography. The solution conformation of one ligand was determined by NMR spectroscopy. WGA is a plant lectin that differs from most lectins by the fact that it has several binding sites per subunit that are distributed over the whole protein. The WGA dimer accommodates eight independent binding sites (four unique

sites). Employing a GlcNAc carbamate motif with α -configuration and by systematic variation of the spacer length, we were able to identify divalent ligands with unprecedented high WGA binding potency. The best divalent ligand (**22**) has an IC_{50} value of $9.8 \mu\text{M}$ corresponding to a relative potency of 2350 (1170 on a valency-corrected basis, i.e., per mol sugar contained) compared to free GlcNAc. This is the highest inhibitory activity of any divalent WGA ligand yet reported.^{28,47} X-ray crystallography of the complex of WGA and the second best, closely related ligand **23** explains this activity. Four molecules of **23** simultaneously bind to WGA with each ligand bridging adjacent binding sites which are approximately 14 Å apart from each other (distance between the anomeric centers of bound GlcNAc moieties). To the best of our knowledge, this is the first crystal structure of a divalent lectin ligand binding in a chelating fashion. From this structure, it also becomes clear that the α -glycosidic linkage is a key feature of the spanning ligands. The axial glycosidic bonds of bound GlcNAc pairs point to each other facilitating the construction of a suitable linker. With β -glycosidic linkages, a considerably longer linker would be required. Although it has been shown that WGA can bind to simple α - and β -glycosides of GlcNAc with similar affinity (cf. Table 1 and literature³¹), our X-ray structure suggests such β -glycosidically linked divalent ligands to be less favorable to high-affinity binding for entropic reasons.

Binding to the four low-affinity (secondary) WGA binding sites has been reported to be too weak to be detectable in solution. Low occupancy of the secondary binding sites was also observed in the complex of WGA1 and monovalent propyl carbamate **20**. The high binding affinity of divalent ligand **23**, on the other hand, explains why in the complex WGA3~**23** all binding sites including the secondary ones are occupied. Thus, multivalent lectin ligands with suitable spacers can be a tool for generating crystal structures of multivalent lectins with low-affinity binding sites occupied.

With an IC_{50} value of $0.9 \mu\text{M}$, the binding potency of the tetravalent neoglycopeptide **2** is 25 500 times higher than that of GlcNAc (6400 times per contained sugar). Comparison of the X-ray structure and the solution NMR structure as well as comparison of the IC_{50} values of dipeptide **28** and divalent ligand **23** (Table 1) suggests that the backbone conformation of the fragment D-Dab-4–D-Dab-5 is already preorganized in solution to adopt a conformation suitable for bridging the high-affinity binding site pairs. However, the two additional GlcNAc residues contained in **2** allow for additional binding modes (spanning of binding sites by alternative GlcNAc pairs or cross-linking of several WGA molecules) that are likely to sum up to the observed high affinity.

The high-resolution crystal structure of WGA in complex with four molecules of a divalent ligand reveals that all binding sites of the WGA dimer are simultaneously functional. This is now the basis for the design of future high-affinity di- and higher valent WGA ligands. The structure immediately suggests that tetravalent ligands composed of two divalent units of type **22** interconnected by a linker can be envisioned that are either able to bind to four binding sites of a single WGA dimer resulting

- (47) (a) Zanini, D.; Roy, R. *Bioconjugate Chem.* **1997**, *8*, 187–192. (b) Furuike, T.; Aiba, S.; Nishimura, S.-I. *Tetrahedron* **2000**, *56*, 9909–9915. (c) Consoli, G. M. L.; Cunsolo, F.; Geraci, C.; Sgarlata, V. *Org. Lett.* **2004**, *6*, 4163–4166. (d) Yamada, A.; Hatano, K.; Koyama, T.; Matsuoka, K.; Takahashi, N.; Hidari, K. I. P. J.; Suzuki, T.; Suzuki, Y.; Terunuma, D. *Bioorg. Med. Chem.* **2007**, *15*, 1606–1614. (e) Misawa, Y.; Akimoto, T.; Amarume, S.; Murata, T.; Usui, T. *J. Biochem.* **2008**, *143*, 21–30.

in even better binding affinity or could, for example, find application as lectin cross-linker. It is also obvious that even subtle changes of the linker structure (cf. **22/23**) can have a significant impact on the binding affinity. Thus, combinatorial approaches such as our spatial screening procedure^{20b} that probe for linker length and flexibility are valuable tools for the identification of optimized multivalent ligands of proteins with unknown 3D structure.

Connecting ligands that can bind to different protein binding sites by a suitable linker is an established strategy in fragment-based drug discovery aiming at improved association constants.⁴⁸ Our findings provide clear experimental evidence for the validity of this approach and show that high-affinity binding to a protein can be achieved by carefully designed ligands that are able to bridge several pairs of binding sites.

Experimental Section

General Methods. Solvents were purified by distillation and dried by usual procedures. Thin-layer chromatography (TLC) was performed on Merck silica gel 60 F₂₅₄ aluminum sheets. Compound spots were visualized by UV light and immersion in the appropriate stain followed by heating. Preparative flash column chromatography (FC) was carried out on Merck silica gel 60 (40–63 μ m). ¹H NMR and ¹³C NMR spectra were recorded at 300 K on Bruker AC 250, Bruker Avance 400, and Bruker Avance DRX 600 instruments. ¹H chemical shifts are reported in ppm relative to the signal from residual protic solvent (CDCl₃, 7.26 ppm; D₂O, 4.67 ppm). ¹³C chemical shifts are referenced to the solvent signal (CDCl₃: 77.0 ppm) or in case of D₂O to internal standard CD₃CN (1.3 ppm). Assignments of proton and carbon signals were carried out with the aid of DQF-COSY, HMQC, and ROESY experiments. High resolution ESI mass spectra were obtained on a Bruker micrOTOF II instrument (positive mode). Elemental analyses were performed by the microanalytical facility at the Universität Konstanz.

***p*-Nitrophenoxycarbonyl 2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy- α -*D*-glucopyranoside (5).** Compound **4**^{30,49} (1.04 g, 3 mmol) was dissolved in dry CH₂Cl₂ (10 mL) and cooled to 0 °C. After addition of Et₃N (832 μ L, 6 mmol), a solution of *p*-nitrophenyl chloroformate (1.33 g, 6.6 mmol) in dry CH₂Cl₂ (1 mL) was added dropwise upon which a precipitate formed. After a reaction time of 3 h, the solvent was evaporated and the remainder dissolved in ethyl acetate and washed with 10% aq citric acid and water. The combined aqueous phases were re-extracted with ethyl acetate. The combined organic phases were dried (MgSO₄) and concentrated. Short column FC (ethyl acetate/petroleum ether 2:1, then ethyl acetate) gave **5** (1.41 g, 92%) as white solid. *R*_f = 0.51 (ethyl acetate); ¹H NMR (CDCl₃, 400.1 MHz): δ 8.34–8.32 (m, 2H, arenes), 7.48–7.46 (m, 2H, arenes), 6.20 (d, *J* = 3.6 Hz, 1H, H-1), 5.82 (d, *J* = 8.6 Hz, 1H, NH), 5.36–5.26 (m, 2H, H-3, H-4), 4.56 (ddd, *J* = 3.6, 8.5, 10.6 Hz, 1H, H-2), 4.33 (dd, *J* = 4.0, 12.5 Hz, 1H, H-6a), 4.19–4.14 (m, 2H, H-6b, H-5), 2.12 (s, 3H, C(O)CH₃), 2.10 (s, 3H, C(O)CH₃), 2.08 (s, 3H, C(O)CH₃), 2.01 (s, 3H, C(O)CH₃); ¹³C NMR (CDCl₃, 150.9 MHz): δ 171.7, 170.6, 170.2, 169.0 (4 \times C(O)CH₃), 154.9 (O–C(O)–O), 150.5 (C^{arene}), 145.6 (C^{arene}), 125.4 (HC^{arene}), 121.4 (HC^{arene}), 95.7 (C-1), 70.2, 70.1, 67.1 (C-3, C-4, C-5), 61.2 (C-6), 51.3 (C-2), 23.0 (C(O)CH₃), 20.7 (2 \times C(O)CH₃), 20.5 (C(O)CH₃); elemental analysis calcd (%) for C₂₁H₂₄N₂O₁₃ (512.42): C, 49.22; H, 4.72; N, 5.47. Found: C, 49.49; H, 4.91; N, 5.22.

Boc-Lys(Aloc)-*D*-Dab(Ddv)-*D*-Val-*D*-Dab(Ddv)-*D*-Dab(Ddv)-*D*-Val-*D*-Dab(Ddv)-Glu(OAll)- β -Ala-NH-Sieber-TG (7). Solid-phase peptide synthesis was carried out in 20 μ mol scale on

NovaSyn TG Sieber resin (Novabiochem, Darmstadt, Germany) (abbreviation: Sieber-TG) using an ABI 433A peptide synthesizer (Applied Biosystems, Foster City, CA) applying the FastMoc-Chemistry data file (“PrPK Fmoc Depro/Single”). The loading of the resin after immobilization of the first amino acid was 0.2 mmol g⁻¹. Peptide couplings were carried out in 1-methyl-2-pyrrolidone (NMP) with 5 equiv of Fmoc-amino acid and activation by HBTU (4.75 equiv), HOBt (4.75 equiv), and EtNi-Pr₂ (10 equiv). In the last coupling step, Boc-Lys(Aloc)–OH was coupled.

cyclo[Boc-Lys-*D*-Dab(Ddv)-*D*-Val-*D*-Dab(Ddv)-*D*-Dab(Ddv)-*D*-Val-*D*-Dab(Ddv)-Glu]- β -Ala-Sieber-TG (8). After having been swollen in CH₂Cl₂ for 10 min, resin **7** was shaken with a solution of [Pd(PPh₃)₄] (1.6 equiv) and BH₃·HNMe₂ (30 equiv) in CH₂Cl₂ (2 \times 20 min) and subsequently washed with DMF (5 \times 2 min), 0.5% sodium diethyldithiocarbamate solution in DMF (3 \times 1 min), DMF (10 \times 1 min), and CH₂Cl₂ (4 \times 2 min). Then, the resin was treated with 5% HOBt in DMF (3 \times 2 min) and washed with DMF (5 \times 1 min) and NMP (2 \times 1 min). Side-chain cyclization was induced by addition of a solution of HOBt (6 equiv), HBTU (4 equiv), and EtNi-Pr₂ (8 equiv) in NMP. This solution was replaced several times until Kaiser test and TNBS test indicated completion of the reaction followed by washing the resin with DMF (6 \times 2 min) and CH₂Cl₂ (3 \times 2 min).

Cyclic Neoglycopeptide (2). Resin **8** was treated with hydrazine hydrate/DMF (4:96) (6 \times 5 min) and washed with DMF (10 \times 1 min). A solution of active carbonate **5** (12 equiv) and EtNi-Pr₂ (12 equiv) in NMP was added, and after completion of the coupling step (reaction control by Kaiser test and TNBS test), the resin was washed with DMF (10 \times 1 min) and CH₂Cl₂ (5 \times 2 min). To cleave acetylated cyclic glycopeptide **9** from the solid phase, the resin was treated with TFA/*i*-Pr₃SiH/CH₂Cl₂ (1:1:98) (1 \times 10 min, 3 \times 2 min) and with TFA/*i*-Pr₃SiH/Mix (1:1:98) (Mix: CH₂Cl₂/F₃CCH₂OH 3:1) (1 \times 5 min, 2 \times 2 min) and thoroughly washed with F₃CCH₂OH, CH₂Cl₂/F₃CCH₂OH (3:1), CH₂Cl₂, MeOH, and finally CH₂Cl₂ (each 2 \times 2 min). The combined filtrates were neutralized with pyridine and concentrated under vacuum. Analytical RP-HPLC (Knauer Nucleosil C18, 4 \times 250 mm, flow: 1 mL min⁻¹, 5–100% MeCN in H₂O/0.1% TFA over 30 min) of **9**: *t*_R = 16.2 min. Acetylated glycopeptide **9** was deprotected by treatment with NaOMe (27 mM) in dry MeOH/CHCl₃ (1:1) for 3.5 h. After neutralization with weakly acidic ion-exchange resin (Amberlite IRC-86) and filtration, the solvent was removed under vacuum. The remaining ion-exchange resin was washed with MeOH and water and the combined filtrates were lyophilized and purified by RP-HPLC (Knauer Eurospher C18, 16 \times 250 mm, flow = 9.6 mL min⁻¹, 10–35% MeCN in H₂O/0.1% TFA over 30 min; *t*_R = 16.6 min) to give cyclic neoglycopeptide **2** in a yield of 15% based on initial resin loading. NMR: see Table S1; MALDI-MS (matrix: α -cyano-4-hydroxycinnamic acid): calcd for C₈₁H₁₃₅N₁₉O₄₀ *m/z* 2036.9 [*M* + Na]⁺, 2052.9 [*M* + K]⁺; found, 2036.8, 2052.8.

General Procedure 1 (GP 1): Coupling of Activated Carbonate 5 to Amino-Substituted Scaffolds. *p*-Nitrophenyl carbonate **5** (1.1 equiv per amino group) was dissolved in dry CH₂Cl₂ (8–15 mL per mmol **5**). Addition of the amine and Hünig’s Base (1 equiv base per amino group) afforded a yellow colored solution. The reaction mixture was stirred at room temperature until the reaction was complete according to TLC. After evaporation of the solvent, the residue was purified by FC.

General Procedure 2 (GP 2): De-*O*-acetylation Using MeOH/EtNMe₂. The acetylated compound was dissolved in MeOH/EtNMe₂ 5:1 and stirred at room temperature until deacetylation was complete according to TLC. Evaporation of the solvent gave the product in quantitative yield.

General Procedure 3 (GP 3): De-*O*-acetylation Using NaOMe/MeOH. The acetylated compound was dissolved in MeOH and catalytic amounts of a solution of NaOMe in MeOH were added. After stirring at room temperature until the deacetylation was complete, the reaction mixture was neutralized using acidic

(48) (a) *Fragment-based Approaches in Drug Discovery*; Jahnke, W.; Erlanson, D. A., Eds.; Wiley-VCH: Weinheim, 2006; (b) Kramer, R. H.; Karpén, J. W. *Nature* **1998**, *395*, 710–713. (c) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531–1534.

(49) Jha, R.; Davis, J. T. *Carbohydr. Res.* **1995**, *277*, 125–134.

ion-exchange resin (Amberlite IRC-86, H⁺ form). Filtration and evaporation of the solvent gave the product.

N-(2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy- α -D-glucopyranosyloxycarbonyl)-propylamine (15). According to GP 1, **5** (50 mg, 0.098 mmol), CH₂Cl₂ (1.5 mL), propylamine (7.3 μ L, 0.089 mmol), and EtN(*i*-Pr)₂ (15 μ L, 0.089 mmol) were reacted for 1.5 h. FC (EtOAc) gave **15** (34 mg, 89%) as a white solid. *R*_f = 0.16 (EtOAc); ¹H NMR (CDCl₃, 400.1 MHz): δ 6.06 (d, *J* = 3.4 Hz, 1H, H-1), 5.70 (br s, 1H, NHAc), 5.23–5.15 (m, 2H, H-3, H-4), 5.09 (br s, 1H, NH), 4.50–4.45 (m, 1H, H-2), 4.23 (dd, *J* = 12.4, 3.9 Hz, 1H, H-6a), 4.06 (d, *J* = 12.6 Hz, 1H, H-6b), 4.02–3.99 (m, 1H, H-5), 3.24–3.12 (m, 2H, CH₂CH₂CH₃), 2.06 (s, 3H, C(O)CH₃), 2.02 (s, 3H, C(O)CH₃), 2.01 (s, 3H, C(O)CH₃), 1.93 (s, 3H, C(O)CH₃), 1.56 (sext, *J* = 7.3 Hz, 2H, CH₂CH₂CH₃), 0.94 (t, *J* = 7.4 Hz, 3H, CH₂CH₂CH₃); ¹³C NMR (CDCl₃, 100.6 MHz): δ 171.5, 170.7, 170.0, 169.1 (4 \times C(O)CH₃), 153.7 (OC(O)NH), 91.4 (C-1), 70.7 (C-3 or C-4), 69.4 (C-5), 67.6 (C-4 or C-3), 61.6 (C-6), 50.9 (C-2), 42.9 (CH₂CH₂CH₃), 23.0 (CH₂CH₂CH₃), 22.9, 20.7, 20.6, 20.5 (C(O)CH₃), 11.1 (CH₂CH₂CH₃); elemental analysis calcd (%) for C₁₈H₂₈N₂O₁₀ (432.42): C, 50.00; H, 6.53; N, 6.48. Found: C, 50.08; H, 6.49; N, 6.49.

1,5-Bis-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy- α -D-glucopyranosyloxycarbonylamino)-3-oxa-pentane (16). According to GP 1, **5** (307 mg, 0.6 mmol), CH₂Cl₂ (5.5 mL), 1,5-diamino-3-oxapentane (29 μ L, 0.27 mmol), and EtN(*i*-Pr)₂ (94 μ L, 0.55 mmol) were reacted for 2 h. FC (CH₂Cl₂/MeOH 20:1) gave **16** (230 mg, 100%) as a white solid. *R*_f = 0.29 (CH₂Cl₂/MeOH 15:1); ¹H NMR (CDCl₃, 400.1 MHz): δ 6.85 (d, *J* = 6.8 Hz, 2H, NHAc), 6.15 (d, *J* = 3.5 Hz, 2H, H-1), 6.10 (t, *J* = 5.3 Hz, 2H, NH), 5.20 (t, *J* = 10.0 Hz, 2H, H-3), 5.13 (t, *J* = 9.6 Hz, 2H, H-4), 4.54 (dt, *J* = 9.7, 3.7 Hz, 2H, H-2), 4.21 (dd, *J* = 12.4, 3.5 Hz, 2H, H-6a), 4.10–4.07 (m, 2H, H-5), 4.03 (d, *J* = 12.2 Hz, 2H, H-6b), 3.58 (t, *J* = 4.4 Hz, 4H, NHCH₂CH₂O), 3.44–3.27 (m, 4H, NHCH₂CH₂O), 2.01 (s, 6H, C(O)CH₃), 1.97 (s, 12H, C(O)CH₃), 1.91 (s, 6H, C(O)CH₃); ¹³C NMR (CDCl₃, 100.6 MHz): δ 170.8, 170.6, 170.5, 169.4 (C(O)CH₃), 154.9 (OC(O)NH), 91.4 (C-1), 70.4 (C-3), 69.6 (NHCH₂CH₂O), 69.2 (C-5), 68.0 (C-4), 61.8 (C-6), 50.7 (C-2), 40.4 (NHCH₂CH₂O), 22.6, 20.6, 20.5, 20.4 (C(O)CH₃); elemental analysis calcd (%) for C₃₄H₅₀N₄O₂₁ (850.78): C, 48.00; H, 5.92; N, 6.59. Found: C, 47.62; H, 6.07; N, 6.44.

1,4-Bis-(3-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy- α -D-glucopyranosyloxycarbonylamino)-propoxy)-butane (17). According to GP 1, **5** (225 mg, 0.44 mmol), CH₂Cl₂ (5 mL), 1,4-bis-(3-aminopropoxy)-butane (43 μ L, 0.2 mmol), and EtN(*i*-Pr)₂ (70 μ L, 0.4 mmol) were reacted for 3 h. FC (CH₂Cl₂/MeOH 15:1) gave **17** (185 mg, 97%) as a white solid. *R*_f = 0.20 (CH₂Cl₂/MeOH 15:1); ¹H NMR (CDCl₃, 600.1 MHz): δ 6.08 (d, *J* = 3.5 Hz, 2H, H-1), 5.68 (d, *J* = 9.3 Hz, 2H, NHAc), 5.44 (t, *J* = 5.3 Hz, 2H, NH), 5.23–5.18 (m, 4H, H-4, H-3), 4.53–4.50 (m, 2H, H-2), 4.25 (dd, *J* = 12.4, 3.6 Hz, 2H, H-6a), 4.08 (dd, *J* = 12.3, 1.6 Hz, 2H, H-6b), 4.01–4.00 (m, 2H, H-5), 3.53–3.50 (m, 4H, CH₂CH₂CH₂), 3.47 (br s, 4H; CH₂CH₂), 3.35–3.33 (m, 4H; CH₂CH₂CH₂), 2.09 (s, 6H; C(O)CH₃), 2.06 (s, 6H; C(O)CH₃), 2.03 (s, 6H; C(O)CH₃), 1.94 (s, 6H; C(O)CH₃), 1.84–1.81 (m, 4H; CH₂CH₂CH₂), 1.67 (br s, 4H; CH₂CH₂); ¹³C NMR (CDCl₃, 150.9 MHz): δ 171.6, 170.8, 169.9, 169.1 (C(O)CH₃), 153.7 (OC(O)NH), 91.5 (C-1), 70.8 (C-3 or C-4), 70.7 (CH₂CH₂), 69.4 (C-5), 69.0 (CH₂CH₂CH₂), 67.6 (C-4 or C-3), 61.5 (C-6), 50.9 (C-2), 39.5 (CH₂CH₂CH₂), 29.4 (CH₂CH₂CH₂), 26.4 (CH₂CH₂), 23.1, 20.7, 20.7, 20.6 (C(O)CH₃); elemental analysis calcd (%) for C₄₀H₆₂N₄O₂₂ (950.93): C, 50.52; H, 6.57; N, 5.89. Found: C, 50.20; H, 6.60; N, 5.85.

N,N'-Bis-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy- α -D-glucopyranosyloxycarbonyl)-4,7,10-trioxa-1,13-tridecandiamine (18). According to GP 1, **5** (225 mg, 0.44 mmol), CH₂Cl₂ (5 mL), 4,7,10-trioxa-1,13-tridecandiamine (44 μ L, 0.2 mmol), and EtN(*i*-Pr)₂ (70 μ L, 0.4 mmol) were reacted for 2 h. FC (CH₂Cl₂/MeOH 15:1) gave **18** (192 mg, 99%) as a white solid. *R*_f = 0.20 (CH₂Cl₂/MeOH 15:1); ¹H NMR (600.1 MHz, CDCl₃): δ 6.06 (d, *J* = 3.6 Hz, 2H; H-1), 6.05 (br s, 2H; NHAc), 5.67 (t, *J* = 5.6 Hz, 2H;

NH), 5.22–5.19 (m, 4H; H-3, H-4), 4.53–4.51 (m, 2H; H-2), 4.24 (dd, *J* = 12.4, 3.7 Hz, 2H; H-6a), 4.09–4.02 (m, 4H; H-6b, H-5), 3.68–3.57 (m, 12H; CH₂CH₂, NCH₂CH₂CH₂O), 3.33–3.31 (q, *J* = 6.2 Hz, 4H; NCH₂CH₂CH₂O), 2.08 (s, 6H; C(O)CH₃), 2.03 (s, 6H; C(O)CH₃), 2.02 (s, 6H; C(O)CH₃), 1.93 (s, 6H; C(O)CH₃), 1.83–1.80 (m, 4H; NCH₂CH₂CH₂O); ¹³C NMR (CDCl₃, 150.9 MHz): δ 171.5, 170.8, 170.1, 169.1 (C(O)CH₃), 153.9 (OC(O)NH), 91.6 (C-1), 70.9 (C-3), 70.3 (CH₂CH₂), 70.0 (CH₂CH₂), 69.4 (C-5), 69.3 (NCH₂CH₂CH₂O), 67.7 (C-4), 61.6 (C-6), 50.8 (C-2), 39.3 (NCH₂CH₂CH₂O), 29.2 (NCH₂CH₂CH₂O), 22.9, 20.7, 20.7, 20.6 (C(O)CH₃); elemental analysis calcd (%) for C₄₀H₆₂N₄O₂₃ (966.93): C, 49.69; H, 6.46; N, 5.79. Found: C, 49.41; H, 6.47; N, 5.48.

Tris-(2-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy- α -D-glucopyranosyloxycarbonylamino)-ethyl)-amine (19). According to GP 1, **5** (338 mg, 0.66 mmol), CH₂Cl₂ (5 mL), tris-(2-aminoethyl)-amine (30 μ L, 0.2 mmol), and EtN(*i*-Pr)₂ (105 μ L, 0.6 mmol) were reacted for 2.5 h. FC (CH₂Cl₂/MeOH 9:1) gave **19** (220 mg, 88%) as a white solid. *R*_f = 0.26 (CH₂Cl₂/MeOH 9:1); ¹H NMR (600.1 MHz, CDCl₃): δ 6.75 (d, *J* = 8.7 Hz, 3H; NHAc), 6.25 (d, *J* = 3.1 Hz, 3H; H-1), 5.84 (br s, 3H; NH), 5.26 (t, *J* = 10.2 Hz, 3H; H-3), 5.19 (t, *J* = 9.8 Hz, 3H; H-4), 4.58–4.54 (m, 3H; H-2), 4.24 (dd, *J* = 12.3, 3.5 Hz, 3H; H-6a), 4.12–4.10 (m, 6H; H-6b, H-5), 3.28 (br s, 6H; CH₂CH₂), 2.67–2.58 (m, 6H; CH₂CH₂), 2.07 (s, 9H; C(O)CH₃), 2.04 (s, 9H; C(O)CH₃), 2.03 (s, 9H; C(O)CH₃), 1.94 (s, 9H; C(O)CH₃); ¹³C NMR (CDCl₃, 150.9 MHz): δ 171.0, 170.7, 170.5, 169.5 (C(O)CH₃), 155.1 (OC(O)NH), 91.9 (C-1), 70.4 (C-3), 69.3 (C-5), 68.2 (C-4), 61.8 (C-6), 55.8 (CH₂CH₂), 50.8 (C-2), 40.1 (CH₂CH₂), 22.8, 20.7, 20.6, 20.6 (C(O)CH₃); elemental analysis calcd (%) for C₅₁H₇₅N₇O₃₀ (1266.17): C, 48.38; H, 5.97; N, 7.74. Found: C, 48.28; H, 6.01; N, 7.30.

(2-Acetamido-2-deoxy- α -D-glucopyranosyloxycarbonyl)-propylamine (20). According to GP 2, **15** (180 mg, 0.375 mmol) was dissolved in MeOH/EtMe₂N 5:1 and reacted for 3 days. Compound **20** was obtained in quantitative yield as a white solid. *R*_f = 0.58 (MeCN/H₂O 4:1); ¹H NMR (600.1 MHz, D₂O): δ 5.87 (br s, 1H; H-1), 3.97–3.96 (m, 1H; H-2), 3.76–3.67 (m, 4H; H-3, H-5, H-6a, H-6b), 3.50 (t, *J* = 9.2 Hz, 1H; H-4), 3.04 (br s, 2H; CH₂CH₂CH₃), 1.95 (s, 3H; C(O)CH₃), 1.44 (q, *J* = 6.6 Hz, 2H; CH₂CH₂CH₃), 0.82 (t, *J* = 6.8 Hz, 3H; CH₂CH₂CH₃); ¹³C NMR (150.9 MHz, D₂O): δ 175.4 (C(O)CH₃), 157.4 (OC(O)NH), 92.3 (C-1), 74.6 (C-3 or C-5), 71.5 (C-5 or C-3), 70.3 (C-4), 61.1 (C-6), 53.5 (C-2), 43.1 (CH₂CH₂CH₃), 23.0 (CH₂CH₂CH₃), 22.6 (C(O)CH₃), 11.3 (CH₂CH₂CH₃); HR-MS (ESI-TOF) for C₁₂H₂₂N₂O₇: [M + H]⁺ calcd, 307.14998; found, 307.1494; [M + Na]⁺ calcd, 329.13192; found, 329.1312.

1,5-Bis-(2-acetamido-2-deoxy- α -D-glucopyranosyloxycarbonylamino)-3-oxa-pentane (21). According to GP 3, **16** (200 mg, 0.235 mmol) was dissolved in dry MeOH (5 mL). After addition of a solution of NaOMe in MeOH (0.5 M, 275 μ L, 138 μ mol) and stirring for 1.5 h, the reaction mixture was neutralized using acidic ion-exchange resin (Amberlite IRC-86, H⁺ form). Filtration and crystallization gave **21** (140 mg) in quantitative yield as a white solid. *R*_f = 0.19 (MeCN/H₂O 4:1); ¹H NMR (400.1 MHz, D₂O): δ 5.93 (d, *J* = 2.9 Hz, 2H; H-1), 4.03 (dd, *J* = 10.3, 3.3 Hz, 2H; H-2), 3.81–3.70 (m, 8H; H-3, H-5, H-6a, H-6b), 3.62–3.53 (m, 6H; H-4, CH₂CH₂), 3.32 (br s, 4H; CH₂CH₂), 1.99 (s, 6H; C(O)CH₃); ¹³C NMR (100.6 MHz, D₂O): δ 175.6 (C(O)CH₃), 157.5 (OC(O)NH), 92.6 (C-1), 74.8 (C-5), 71.7 (C-3), 70.4 (C-4), 70.1 (CH₂CH₂), 61.3 (C-6), 53.7 (C-2), 41.1 (CH₂CH₂), 22.8 (C(O)CH₃); HR-MS (ESI-TOF) for C₂₂H₃₈N₄O₁₅: [M + H]⁺ calcd, 599.24064; found, 599.2381; [M + Na]⁺ calcd, 621.22259; found, 621.2200.

1,4-Bis-(3-(2-acetamido-2-deoxy- α -D-glucopyranosyloxycarbonylamino)-propoxy)-butane (22). According to GP 3, **17** (141 mg, 148 μ mol) was dissolved in dry MeOH (3 mL). After addition of a solution of NaOMe in MeOH (0.54 M, 5 μ L, 27 μ mol) and stirring for 2 h, the reaction mixture was neutralized using acidic ion-exchange resin (Amberlite IRC-86, H⁺ form). Filtration and crystallization gave **22** (90 mg, 87%) as a white solid. *R*_f = 0.26 (MeCN/H₂O 4:1); ¹H NMR (600.1 MHz, D₂O): δ 5.86 (br s, 2H;

H-1), 3.98 (dd, $J = 10.5$, 2.1 Hz, 2H; H-2), 3.76–3.66 (m, 8H; H-6a, H-6b, H-5, H-3), 3.52–3.46 (m, 10H; CH₂, CH₂, H-4), 3.16 (br s, 4H; CH₂), 1.96 (s, 6H; C(O)CH₃), 1.72 (t, $J = 6.0$ Hz, 4H; CH₂), 1.55 (br s, 4H, CH₂); ¹³C NMR (150.9 MHz, D₂O): δ 175.4 (C(O)CH₃), 157.2 (OC(O)NH), 92.4 (C-1), 74.6 (C-5), 71.5 (C-3), 71.3 (CH₂), 70.2 (C-4), 68.5 (CH₂), 61.1 (C-6), 53.5 (C-2), 38.3 (CH₂), 29.5 (CH₂), 26.2 (CH₂), 22.6 (C(O)CH₃); HR-MS (ESI-TOF) for C₂₈H₅₀N₄O₁₆: [M + H]⁺ calcd, 699.32946; found, 699.3280; [M + Na]⁺ calcd, 721.3114; found, 721.3100.

N,N'-Bis-(2-acetamido-2-deoxy- α -D-glucopyranosyloxycarbonyl)-4,7,10-trioxa-1,13-tridecandiamine (23). According to GP 2, **18** (165 mg, 171 μ mol) was dissolved in MeOH/EtMe₂N 5:1 and reacted for 2 days. Compound **23** was obtained in quantitative yield as a white solid. $R_f = 0.44$ (MeCN/H₂O 2:1); ¹H NMR (600.1 MHz, D₂O): δ 5.86 (d, $J = 3.0$ Hz, 2H; H-1), 3.50 (dd, $J = 10.7$, 3.3 Hz, 2H; H-2), 3.76–3.66 (m, 8H; H-6a, H-6b, H-3, H-5), 3.61 (br s, 4H; OCH₂CH₂O), 3.60 (br s, 4H; OCH₂CH₂O), 3.52–3.48 (m, 6H; NCH₂CH₂CH₂O, H-4), 3.17–3.15 (m, 4H; NCH₂CH₂CH₂O), 1.95 (s, 6H; C(O)CH₃), 1.74–1.72 (m, 4H; NCH₂CH₂CH₂O); ¹³C NMR (150.9 MHz, D₂O): δ 175.4 (C(O)CH₃), 157.2 (OC(O)NH), 92.4 (C-1), 74.7 (C-5), 71.5 (C-3), 70.4 (C-4), 70.2 (OCH₂CH₂O), 69.0 (NCH₂CH₂CH₂O), 61.6 (C-6), 53.5 (C-2), 38.2 (NCH₂CH₂CH₂O), 29.5 (NCH₂CH₂CH₂O), 22.6 (C(O)CH₃); HR-MS (ESI-TOF) for C₂₈H₅₀N₄O₁₇: [M + H]⁺ calcd, 715.32437; found, 715.3229; [M + Na]⁺ calcd, 737.30632; found, 737.3046.

Tris-(2-(2-acetamido-2-deoxy- α -D-glucopyranosyloxycarbonylamino)-ethyl)-amine (24). According to GP 2, **19** (222 mg, 0.175 mmol) was dissolved in MeOH/EtMe₂N 5:1 and reacted for 5 days. Compound **24** was obtained in quantitative yield as a white solid. $R_f = 0.08$ (MeCN/H₂O 4:1); ¹H NMR (600.1 MHz, D₂O): δ 5.88 (br s, 3H; H-1), 3.97–3.96 (m, 3H; H-2), 3.74–3.65 (m, 12H; H-6a, H-6b, H-3, H-5), 3.50 (t, $J = 9.4$ Hz, 3H; H-4), 3.17–3.16 (m, 6H; CH₂), 2.63 (br s, 6H; CH₂), 1.94 (s, 9H; C(O)CH₃); ¹³C NMR (150.9 MHz, D₂O): δ 175.2 (C(O)CH₃), 157.2 (OC(O)NH), 92.3 (C-1), 74.5 (C-5), 71.4 (C-3), 70.1 (C-4), 60.9 (C-6), 53.7 (CH₂), 53.4 (C-2), 39.0 (CH₂), 22.5 (C(O)CH₃); HR-MS (ESI-TOF) for C₃₃H₅₇N₇O₂₁: [M + H]⁺ calcd, 888.36803; found, 888.3651; [M + Na]⁺ calcd, 910.34997; found, 910.3470.

Ac-D-Dab(Ddv)-D-Dab(Ddv)-NH-Rink-PS (26). Solid-phase peptide synthesis was carried out in 90 μ mol scale in a syringe reactor on Rink Amide polystyrene resin (Rapp Polymere, Tübingen, Germany) (abbreviation: Rink-PS, loading density 0.6 mmol g⁻¹) following standard Fmoc protocols.³³ Peptide couplings were performed in NMP using Fmoc-D-Dab(Ddv)-OH (2 equiv), HBTU (1.9 equiv), HOBT (3 equiv), and Et₃Ni-Pr₂ (4 equiv). The N-terminal Fmoc group was removed with 20% piperidine in DMF (1 \times 3 min, 1 \times 15 min) and the resin was washed with DMF (10 \times 1 min). The N-terminus was acylated with 10% Ac₂O in DMF (2 \times 15 min) followed by washing the resin with DMF (10 \times 1 min).

Neoglycopeptide (28). Resin **26** was deprotected by treatment with hydrazine hydrate/DMF (4:96) (6 \times 5 min) and washed with DMF (10 \times 1 min). After swelling for 10 min in Et₃Ni-Pr₂ (3 equiv) in DMSO/DMF (1:3), the resin was treated with active carbonate **5** (4 equiv) and Et₃Ni-Pr₂ (4 equiv) in DMSO/DMF (1:3) overnight and washed with DMF (10 \times 1 min) and CH₂Cl₂. Cleavage from the resin was achieved by treatment with TFA/*i*-Pr₃SiH/water (95:2.5:2.5) (1 \times 3 h, 1 \times 1 h). The cleavage solutions were dropped into ice-cold *tert*-butylmethylether (10 times volume) and stored at -18 °C overnight. Precipitated product was collected by centrifugation at 4 °C and 4500 rpm for 30 min. The resulting pellet was resuspended in *tert*-butylmethylether and the centrifugation procedure was repeated. The pellet was then lyophilized from water to yield crude acetylated neoglycopeptide **27**. MALDI-MS: for C₄₀H₅₉N₇O₂₃ calcd, 1028.4 [M + Na]⁺, 1044.3 [M + K]⁺; found, 1028.3, 1044.3. Acetylated **27** was deprotected by treatment with NaOMe (27 mM) in dry MeOH/CHCl₃ (1:1) for 3.5 h. The solution was neutralized with 30% acetic acid and the solvent was removed under vacuum. Purification by RP-HPLC (Knauer Eurospher C18,

16 \times 250 mm, flow = 9.6 mL min⁻¹, 1–10% MeCN in H₂O/0.1% TFA over 20 min; $t_R = 12.7$ min) gave **28** (28 mg, 41% based on initial resin loading). ¹H NMR (600.1 MHz, DMSO-*d*₆): δ 8.13–8.08 (1 H; NHAc D-Dab-1), 8.02–7.94 (1 H; NH D-Dab-2), 7.74–7.66 (2 H; NHAc GlcNAc), 7.33–7.29 (1 H; NH₂), 7.12–6.81 (3 H; 1 H from NH₂, 2 H from carbamate NH), 5.81–5.79 (2 H; H-1 GlcNAc), 4.37–4.18 (2 H; C⁶H from D-Dab-1 and D-Dab-2), 3.73–3.70 (2 H; H-2 GlcNAc), 3.58–3.43 (8 H + H₂O; H-6a, H-6b, H-3, H-5 GlcNAc), 3.35–3.21 (2 H; H-4 GlcNAc), 3.08–2.92 (4 H; C⁷H₂ from D-Dab-1 and D-Dab-2), 1.89–1.81 (11 H; 9 H from C(O)CH₃, C ^{β} H-a from D-Dab-1 and D-Dab-2), 1.69–1.58 (2 H; C ^{β} H-b from D-Dab-1 and D-Dab-2). ¹³C NMR (150.9 MHz, DMSO-*d*₆): δ 172.6, 171.0 (C(O)CH), 169.3 (C(O)CH₃), 154.4 (OC(O)NH), 90.6 (C-1 GlcNAc), 74.4 (C-5 GlcNAc), 70.1 (C-3 GlcNAc), 70.0 (C-4 GlcNAc), 60.4 (C-6 GlcNAc), 52.8 (C-2 GlcNAc), 50.4 (C ^{α} -D-Dab-1), 50.1 (C ^{α} -D-Dab-2), 36.9 (C ^{γ} -D-Dab-1 and D-Dab-2), 31.7 (C ^{β} -D-Dab-1 and D-Dab-2), 22.4, 22.3 (C(O)CH₃). Analytical RP HPLC (Knauer Eurospher C18, 4 \times 250 mm, flow: 0.9 mL min⁻¹, 1–10% MeCN in H₂O/0.1% TFA over 20 min): $t_R = 13.6$ min. HR ESI-TOF-MS: calcd for C₂₈H₄₇N₇O₁₇, 754.31012 [M + H]⁺, 776.29206 [M + Na]⁺; found, 754.3080, 776.2896.

Enzyme-Linked Lectin Assay (ELLA). Assays were carried out as previously described.²⁸ Briefly, microtiter plates with a covalently immobilized β -GlcNAc derivative [4-(2-acetamido-2-deoxy- β -D-glucopyranosyloxy)-(Z)-but-2-enyloxycarbonyl-(3-(2-(2-(3-amino-propoxy)-ethoxy)-ethoxy)propyl)-amine] as reference ligand were incubated with mixtures of horseradish peroxidase (HRP)-labeled WGA (1 μ g mL⁻¹) and the respective WGA ligand in varying concentrations. After incubation, the plates were washed and remaining labeled WGA bound to the reference ligand was quantified by an HRP-catalyzed color reaction using 2,2'-azinobis(3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt (ABTS) as substrate. From dose-response curves for inhibition of the binding of HRP-labeled WGA to immobilized GlcNAc, the concentrations that reduce the binding of labeled WGA to the microtiter plates by 50% (IC₅₀ values) were determined as a measure of potency of the synthesized inhibitors. All tests were performed in duplicate.

Purification of WGA Isoforms.⁵⁰ WGA was purchased from Sigma as lyophilized powder. About 10 mg was dissolved in 1.5 mL of buffer A (100 mM Na acetate, 20 mM NaCl, pH 3.8). The sample was applied with a flow rate of 1 mL min⁻¹ to a SP Sepharose cation exchange column equilibrated with 3 column volumes (CVs) of buffer A. The column was washed with 1 CV of buffer A. WGA isoforms were eluted with 100 mL of a linear gradient 0–100% buffer B (100 mM Na acetate, 500 mM NaCl, pH 3.8). Fractions containing either WGA isoform 1, 2, or 3 were pooled and concentrated with centrifugal filter devices (10 kDa cutoff) to a final concentration of approximately 8 mg mL⁻¹ as determined by measurement of E_{280} , using a theoretical extinction coefficient of 28 930 M⁻¹ cm⁻¹, calculated based on amino acid sequence.⁵¹

Cocrystals of WGA1 with GlcNAc-Carbamate 20. During protein concentration, buffer was exchanged against crystallization buffer (20 mM Na Acetate, 6 mM CaCl₂, 4% Ethanol at pH 4.9). Small volumes of a solution of **20** in crystallization buffer (0.1 M) were added to the protein sample at 8 mg mL⁻¹ in order to reach a final ligand concentration of 7 mM. The samples were thoroughly vortexed and incubated on ice for 1 h. Crystallization was performed in hanging drop vapor diffusion setups using 1 mL reservoir solution R (0.04 M KH₂PO₄, 20% glycerol, 14–17% PEG 8000) and a drop size of 6 μ L (4 μ L protein solution + 2 μ L reservoir solution). At 18 °C, cocrystals appeared overnight with a size of up to 0.3 mm.

Cocrystals of WGA3 with Divalent Ligand 23. During protein concentration, buffer was exchanged against buffer B. Small

(50) Rice, R. H.; Etzler, M. E. *Biochemistry* **1975**, *14*, 4093–4099.

(51) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. In *The Proteomics Protocols Handbook*; Walker, J. M., Ed.; Humana Press: Totowa, 2005; pp 571–607.

volumes of a solution of **23** in crystallization buffer (0.1 M) were added to the protein sample at 8 mg mL⁻¹ in order to reach a final ligand concentration of 1 mM. The samples were thoroughly vortexed and incubated on ice for 1 h. Crystallization was performed with hanging drop vapor diffusion setups using 1 mL of reservoir solution R and a drop size of 6 μ L (4 μ L protein solution + 2 μ L reservoir solution). At 18 °C, cocrystals appeared after 3 weeks with a size of up to 0.5 mm in the largest dimension.

Crystals of WGA1, Cross-Linking, and Soaking with Tetravalent Cyclic Peptide Ligand 2. During protein concentration of WGA1, buffer was exchanged against buffer B. Hanging drop vapor diffusion experiments were set up using 1 mL of reservoir solution R and a drop size of 6 μ L (4 μ L protein solution + 2 μ L reservoir solution). At 8 °C, crystals of WGA1 appeared overnight and had a size of approximately 0.1 mm. Soaking of WGA1 crystals with tetravalent ligand **2** was unsuccessful because the crystals cracked and dissolved upon addition of the compound. We therefore employed glutaraldehyde cross-linking. Crystals of WGA1 in complex with GlcNAc⁴² were chosen, because of the large solvent channels present in this crystal form and because native WGA1 crystals were reproducibly twinned. The crystals were extensively washed in stabilizing solution (0.04 M KH₂PO₄, 20% (w/v) glycerol, 20% (w/v) PEG 8000) to remove GlcNAc. Cross-linking was performed using 25% glutaraldehyde in a microbridge,⁵² with the exception that cross-linking was prolonged to 12 h at 4 °C. Afterward, crystals were washed again in stabilizing solution and soaked by transfer to stabilizing solution containing approximately 0.2 mM cyclic peptide ligand **2**. Crystals were soaked for 24 h at 4 °C prior to freezing.

Data Collection and Crystal Structure Analysis. Crystals were mounted in cryogenic loops (Hampton) and frozen by plunging into liquid nitrogen. Diffraction data were recorded using an ADSC Q315r detector at beamline ID23-1, ESRF, Grenoble or using a MARresearch MX-225 CCD detector at beamline X06SA, SLS, Villigen. In all cases, data were collected at 100 K. A total of 360 frames per crystal were recorded with an oscillation range of 0.5° and an exposure time of 0.5 or 1 s per frame. X-ray wavelengths are listed in Table 2. For all data sets, data reduction was carried out using the program XDS.⁵³ Data collection statistics are listed in Table 2.

Structures were solved by molecular replacement with known WGA atomic coordinates, using the program Molrep.⁵⁴ As starting model, native WGA1 coordinates (PDB ID: 7WGA²⁴) or native WGA3 coordinates (PDB ID: 1WGT^{23c}) were chosen. Ligands were only included in the model when unbiased positive $F_o - F_c$ density was visible. The structures of WGA1 bound to compound **20** and WGA3 in complex with divalent ligand **23** were refined using the program Refmac.⁵⁵ The WGA3~**23** crystal was merohedrally twinned with the twin operator $-h-k, k, -l$ and a twin fraction of 0.44, as detected by SFCHECK.⁵⁶ Ten percent of all recorded reflections were excluded for cross validation taking into account the twin operator to ensure that twin-related reflections are in the same set. Ligand restraints were generated using the CCP4i monomer sketcher,⁵⁷ Coot,⁵⁸ and the PRODRG server.⁵⁹ The structure of WGA1 complexed with the tetravalent cyclic peptide

ligand **2** was refined with PHENIX⁶⁰ applying restrained positional refinement and TLS refinement. Five percent of recorded reflections were randomly set aside for cross validation. Restraints for the visible part of the cyclic peptide ligand **2** were generated using Coot and the CCP4i monomer sketcher based on the energy-minimized NMR structure with least restraint violations. Electron density map inspection and model building were done using Coot. Geometry was validated and Ramachandran plot was calculated using the Molprobit server.⁶¹ Distances were measured using Coot. Hydrogen bonding was analyzed using UCSF Chimera.⁶² Figures containing protein X-ray structure representations were prepared using PyMol.⁶³ All molecular surface representations are solvent-excluded (Connolly) surfaces, calculated with a probe radius of 1.4 Å.

NMR Spectroscopy. NMR samples contained cyclic neoglycopeptide **2** (3–5 mM), 0.05% NaN₃, and 3-(trimethylsilyl)-propionic acid-2,2,3,3-*d*₄ (TSP) (0.5 mM) in H₂O/D₂O (95:5). The pH was adjusted to 5.0 with aqueous hydrochloric acid and sodium hydroxide solution. Final concentrations of sodium chloride were in the low millimolar range. For experiments in D₂O, the sample was first dried by evaporation and coevaporated with D₂O to ensure quantitative isotopic exchange. The residual dry substance was redissolved in D₂O. Before NMR experiments were performed, the pH was readjusted to 5.0 with NaOD. NMR spectra were recorded at 300 K on a Bruker Avance DRX 600 spectrometer equipped with an inverse 5 mm TXI-H/C/N triple resonance probe with shielded z-gradients. Assignments and calculation of distance and torsion restraints were carried out with the help of standard 2D experiments that were detected according to the States-TPPI method in a phase sensitive mode. 2D spectra were typically recorded with 4k data points in F2 (1k to 1.5k data points in ¹³C-HSQC spectra) and 256 or 512 experiments in F1, with 16–64 transients per increment. The TOCSY experiments were performed with a spinlock field of 10.0 kHz using the MLEV17 sequence with a mixing time of 100 ms. The WATERGATE scheme⁶⁴ was used to suppress the H₂O resonances in ¹H, TOCSY and NOESY spectra. In 2D COSY spectra, H₂O resonances were suppressed by pre-saturation of the H₂O resonances. The spectra were processed with simple zero filling in both dimensions and referenced on TSP using Topspin 1.3.b.20. Cara 1.5.3 was used in order to extract distance restraints from NOESY spectra that were detected with mixing times of 50 ms. Mestrec 4.6.2.3 and Topspin 1.3.b.20 were used to extract coupling constants from DQF COSY spectra. These were translated into dihedral angle restraints with the aid of a Karplus relationship ($^3J_{\text{HH}} = A \cos^2 \phi + B \cos \phi + C$, parameters: $A = 7.9$, $B = -1.55$, $C = 1.35$).⁶⁵

Structure Calculation. All molecular dynamics (MD) simulations were performed using sander of the AMBER 8 package on a computer cluster (Transtec, 8 Opteron Dual-Core processors, 32 GB RAM, CentOS Linux). The simulations were done by combining parameters for the peptide backbone from the PARM94/PARM99 force fields developed by Kollman et al.⁶⁶ with those developed by Woods and co-workers for carbohydrate structures.⁶⁷ As the structure calculations required AMBER-compatible topological descriptions of all contained structural fragments, topology

(52) Lusty, C. J. *J. Appl. Crystallogr.* **1999**, *32*, 106–112.

(53) Kabsch, W. *J. Appl. Crystallogr.* **1993**, *26*, 795–800.

(54) Vagin, A.; Teplyakov, A. *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025.

(55) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240–255.

(56) Vaguine, A. A.; Richelle, J.; Wodak, S. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 191–205.

(57) Potterton, E.; Briggs, P.; Turkmenburg, M.; Dodson, E. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 1131–1137.

(58) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.

(59) Schuettelkopf, A. W.; van Aalten, D. M. F. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1355–1363.

(60) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002**, *58*, 1948–1954.

(61) Lovell, S. C.; Davis, I. W.; Arendall, W. B., III; de Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. *Proteins: Struct., Funct., Genet.* **2003**, *50*, 437–450.

(62) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605–1612.

(63) DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific: Palo Alto, CA, 2002.

(64) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661–665.

(65) Ramachandran, G. N.; Chandrasekaran, R.; Kopple, K. D. *Biopolymers* **1971**, *10*, 2113–2131.

files were generated accordingly for several artificial residues contained in the neoglycopeptide. For Lys and Glu residues with side chains involved in an amide bond, for D-Dab, and for D-Val residues, topology files were created in the format of the amino acid library in use. AMBER-compatible topological descriptions of the GlcNAc- the Boc, and the β -Ala residues were generated with the programs antechamber and prepgen that are part of AMBER 8. Partial charges were used as given in the library all_amino94.lib or calculated using prepgen, that are both part of the AMBER 8 package. For MD simulations in vacuum, these charges were reduced to 20% of their original values.

A structurally diverse ensemble of 100 conformers of neoglycopeptide **2** was generated by 150 ps unconstrained molecular dynamics simulation at 2000 K. Subsequently, the ensemble was subjected to the following steps in the presence of NMR constraints: Simulated annealing succeeded by energy-minimization in vacuum followed by simulated annealing and subsequent energy minimization employing a generalized Born solvation model.⁶⁸

(66) (a) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197. (b) Cheatham, T. E., 3rd; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845–862.

Acknowledgment. This Article is dedicated to Professor Horst Kessler on the occasion of his 70th birthday. Financial support by the Deutsche Forschungsgemeinschaft (FOR 434 “Oligosaccharide and DNA Chips - Analysis of Secondary Gene Products”), the University of Konstanz, the Konstanz Research School Chemical Biology, and the BMBF is gratefully acknowledged. J.G.B. is grateful for a stipend of the Friedrich-Ebert-Stiftung. We thank Konrad Bergen and Philippe Reuter for their help with crystallization and the staff at the SLS and the ESRF for assistance in crystal data collection.

Supporting Information Available: Inhibition curves, electron density plots, additional data and figures on structure determination of **2**, and NMR spectra of **5**, **15–24**, and **28**. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA101646K

(67) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832–3846.

(68) Xia, B.; Tsui, V.; Case, D. A.; Dyson, H. J.; Wright, P. E. *J. Biomol. NMR* **2002**, *22*, 317–331.