Bioorganic & Medicinal Chemistry Letters 18 (2008) 6573–6575

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Rational design of novel glycomimetics: Inhibitors of concanavalin A

Karen T. Welch *, Trent A. Turner, Callie E. Preast

Department of Chemistry, 250 Forest Dr., Georgia Southern University, Statesboro, GA 30460-8064, USA

ARTICLE INFO

Article history:
Received 7 August 2008
Revised 24 September 2008
Accepted 26 September 2008
Available online 1 October 2008

Keywords:
Glycomimetic
Virtual screening
Concanavalin A

ABSTRACT

A virtual screening approach was used to identify new glycomimetics. The National Cancer Institute Diversity Set was docked into the carbohydrate binding site of the lectin concanavalin A (ConA). The resulting poses were analyzed and 19 molecules were tested for inhibition with an enzyme-linked lectin assay (ELLA). Eight of the 19 molecules inhibited ConA–carbohydrate binding. The two most potent inhibitors have IC_{50} values that are an order of magnitude smaller than the monosaccharide methyl α-D-mannopyranoside.

The carbohydrates presented on cell surfaces and those attached to proteins are assembled in patterns that comprise a carbohydrate code.[1] This carbohydrate code has a coding capacity that surpasses that of nucleic acids or proteins. The high capacity arises from the number of different ways two monosaccharides can be joined together, in linear or branched patterns, as well as the number of different monosaccharides available, the pyranose and furanose forms of sugar rings, and covalent modifications such as sulfation.[2] Protein partners are integral to the assembly and modification of carbohydrate signals, as well as their interpretation.[3] Carbohydrate chains are built by enzymes known as glycosyltransferases and they are cleaved by glycosidases. The information encoded in carbohydrate sequences is decoded by the binding of protein receptors known as lectins.

The high coding capacity of carbohydrates is put to use in many normal biological processes, as well as disease states. These include fertilization,[4] inflammation,[5] cancer,[6] bacterial infections,[7] fungal infections,[8] and viral infections.[9] The link between cell-surface interactions and disease has made carbohydrates attractive targets for chemotherapy.[10] Glycomimetics, or molecules that mimic the function of carbohydrates in the body, have also been pursued as disease therapies and are an active area of research. A recent issue of *Carbohydrate Research* was devoted to the topic of glycomimetics.[11]

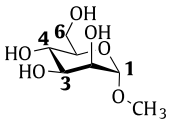
Recently, there have been reports of less polar glycomimetics that bear little or no resemblance to the carbohydrate molecules they are intended to replace. These include chitinase inhibitors[12] and selectin inhibitors.[5,13] In addition, there is much interest in polyphenolic inhibitors of dietary enzymes α-amylase and α-glucosidase.[14] Therefore, a study aimed at discovering new glycomimetics is both relevant and timely.

We are using the lectin concanavalin A (ConA) as a canonical carbohydrate-binding protein and a template for the rational design of molecules that could serve as glycomimetic lead compounds. Carbohydrate-based,[15] glycodendrimer-based,[16] and peptide-based[17] inhibitors of ConA are known. ConA is isolated from the seeds of the jack bean *Canavalia ensiformis*.[18] It exists as a dimer at acidic pH, a tetramer at pH 7, and it forms higher order aggregates at higher pH values.[19] Each ConA monomer contains one calcium ion binding site, one transition metal binding site, and one carbohydrate binding site, also referred to as the combining site.[20] The carbohydrate binding site is near the metal binding sites, but they do not overlap. ConA binds methyl α-D-mannopyranoside with higher affinity than methyl α-D-glucopyranoside, and methyl α-D-galactopyranoside does not bind (Fig. 1).[21] The equatorial substituents at carbons 3, 4, and 6 are needed for monosaccharide binding.[21]

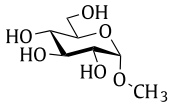
The carbohydrate binding site of ConA contains multiple distinct subsites.[22] The monosaccharide binding subsite makes a series of hydrogen bonds and ion-dipole contacts with the hydroxyl groups on carbon atoms 3, 4, and 6 of mannose or glucose (Fig. 1). This includes backbone interactions with Lue 99, Tyr100, and Arg228, as well as interactions with the sidechains of Asn14 and Asp208. A hydrophobic subsite extends from the monosaccharide binding site, formed by the sidechains of Tyr12, Leu99, Tyr100 and Ala207. An additional polar subsite allows ConA to bind the core trimannoside of asparagine-linked carbohydrates with 60-fold greater affinity than me α-man.[23] This subsite consists of the residues Pro13, Thr15, Asp16, and the sidechain of Asp228. There is a conserved water molecule in this site that is associated with Asn14, Asp16, and Arg228.[24]

* Corresponding author. Tel.: +1 912 478 1883; fax: +1 912 478 0699.
E-mail address: kwelch@georgiasouthern.edu (K.T. Welch).

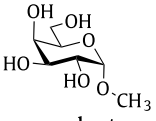
0960-894X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2008.09.095



Methyl α-D-mannopyranoside
(Me α-man)



Methyl α-D-glucopyranoside
(Me α-glc)



Methyl α-D-galactopyranoside
(Me α-gal)

Figure 1. Methyl glycosides discussed in this paper and used as reference compounds for docking. Carbon atoms 1, 3, 4, and 6 of Me α-man are labeled.

We docked the National Cancer Institute (NCI) Diversity Set,[25] a database containing 1990 molecules, into the carbohydrate binding site of ConA. The database molecules were prepared for docking with Ligprep,[26] and the docking calculations were carried out with Glide.[27] The two crystal structures of ConA used for the docking calculations were 5CNA (ConA cocrystallized with me α-man)[28] and 1GIC (ConA cocrystallized with me α-glc).[29] All of the water molecules were removed from the active site prior to docking. The poses that resulted from the docking calculations were screened using a distance criterion. Only those poses with one or more ligand atoms positioned within 3 A of the sidechain carboxylate group of Asp208 in the carbohydrate binding site of ConA were retained. This distance was chosen because two of the hydrogen bond donors required for monosaccharide binding (OH-4 and OH-6) have ion-dipole contacts with this residue. This screen removed approximately one third of the total number of poses from consideration, and it removed poses throughout the continuum of scores.

We added me α-man, me α-glc, and me α-gal to the molecules to be docked as reference compounds (Fig. 1). The three scoring functions available with Glide (Gscore,[30] EModel,[31] and Glide Energy[31]) had different selectivities for the database molecules relative to the reference compounds. Since we did not know a priori which scoring function would be most appropriate, we generated an empirical consensus score from all three scoring functions. The function used was Gscore + 2 Emodel + Glide Energy (Eq. 1). The Emodel scores were weighted more heavily because Emodel provided the largest numerical range of scores and its use was supported by precedent.[32] More detail on the results of the docking calculations may be found in the Supplementary Data.

With the consensus scores in hand, we compared the top 200 poses for 5CNA and 1GIC and identified the molecules common to both. We examined these poses manually and eliminated molecules which were not predicted to interact with Asp208, molecules with fewer than 3 predicted hydrogen bond, ion-dipole, or salt bridge interactions in the carbohydrate binding site, large molecules, and most nucleosides and carbohydrates. After these considerations, we arrived at a set of 19 molecules which were obtained from the NCI.

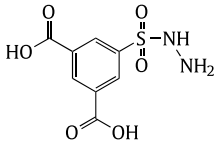
The molecules obtained from the NCI were tested in an enzyme-linked lectin assay (ELLA) which detects ConA–yeast mannan binding.[16b] In these assays, a 96-well plate is coated with yeast mannan, and the presence of ConA bound to the yeast mannan is detected by color development from the oxidation of ABTS catalyzed by a ConA–horseradish peroxidase (HRP) conjugate. DMSO was also used as a control, as it was used to dissolve all of the molecules tested in the ELLA. Inhibition by DMSO was negligible.

The ELLA results are summarized in Table 1 and the structures of the inhibitors are shown in Figure 2. Of the 19 compounds tested, a total of eight compounds inhibited ConA–mannan binding. NSC 119910 and 120634 have IC_{50} values that are an order of magnitude lower than that of me α-man. NSC 77393, 134196, and 143099 are comparable to me α-man based upon the IC_{50} data. NSC 134159 is less potent that me α-man. NSC 72234 and 99799

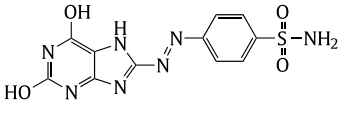
Table 1
ELLA results.

Compound	IC_{50}[a] or maximum inhibition
NSC 120634	19 μM (± 11)
NSC 119910	29 μM (± 10)
NSC 134196	234 μM (± 44)
NSC 143099	479 μM (± 121)
NSC 77393	615 μM (± 30)
NSC 134159	49% at 11 mM
NSC 72234	n.c.d.[b]
NSC 99799	n.c.d.[b]
Me α-man	350 μM (± 43)

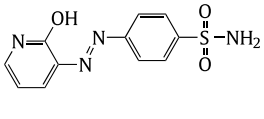
[a] Experiments were performed in triplicate and standard deviations are given in parentheses.
[b] Inhibition was not concentration-dependent.



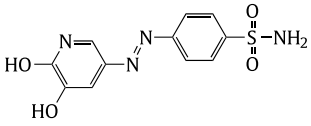
NSC 120634



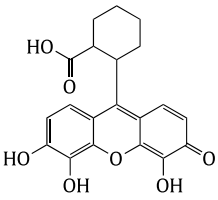
NSC 77393



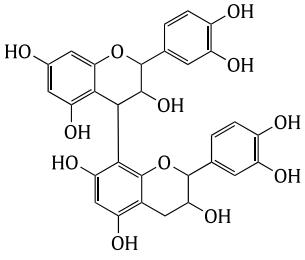
NSC 134196



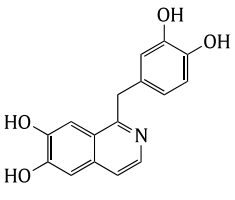
NSC 134159



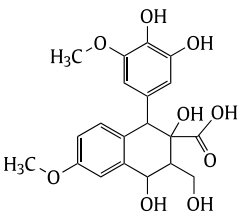
NSC 119910



NSC 143099



NSC 99799



NSC 72234

Figure 2. Structures of inhibitors from the National Cancer Institute Diversity Set.

inhibit ConA–mannan binding, but inhibition is not concentration-dependent for these compounds.

Inhibitor NSC 119910 is predicted to make several contacts in the carbohydrate binding site of Con A (Fig. 3). The catechol group interacts with Asp208. One edge of the aromatic ring system projects into the hydrophobic subsite, while the other stacks with the sidechain of Arg228. In addition, the carboxylate group is positioned near the sidechains of Asn14 and Arg228, like the crystallographically conserved water molecule.

Inhibitor NSC 120634 has NH and NH_2 groups that interact with Asp208, while the sulfonamide oxygen atoms are in close proximity to the backbone NH groups of Leu99, Tyr100, and Arg228 (Fig. 4). A hydrogen bond between a sulfonamide oxygen and the NH of Leu 99 is shown in Figure 4. One of the carboxylate groups is predicted to interact with the sidechains of Asn14 and Arg 228, like NSC 119910. The other is in close proximity to the phenolic OH groups of Tyr12 and 100.

Our results suggest several glycomimetic motifs that should be studied further. These include xanthene derivatives and polyphenolic compounds (NSC 119910 and 143099). The sulfonamide group (NSC 120634, 134196, 77393, and 134159) is also a potential glycomimetic motif. The xanthine moiety found in NSC 77393 is similar to that of chitinase inhibitors designed by van Alten and coworkers.[12] In addition, the azo linkage found in some of the inhibitors may serve as a convenient way to prepare new inhibitors in a fragment-based design approach. Moreover, the study of the interaction of NSC 143099 with ConA may shed light on the interaction of flavonoids with carbohydrate-binding proteins.

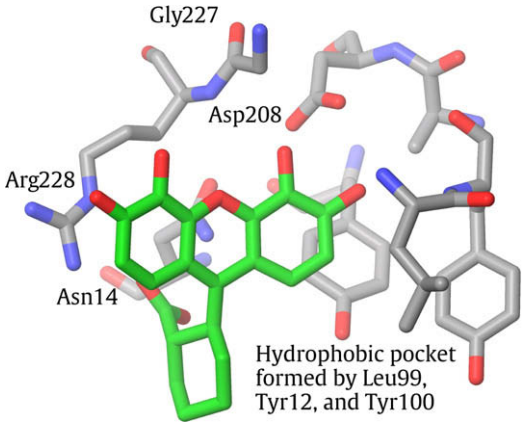


Figure 3. NSC 119910 in the active site of ConA (PDB ID 5CNA) as predicted by docking.

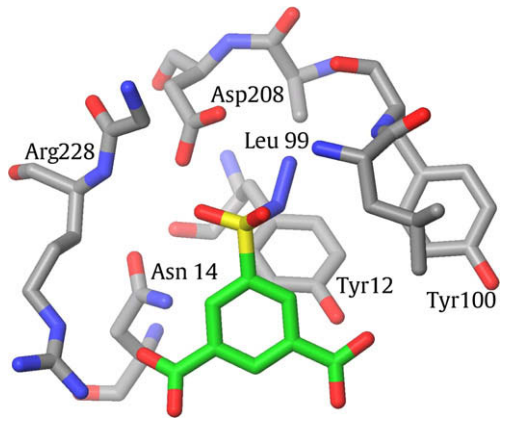


Figure 4. NSC 120634 in the active site of ConA (PDB ID 5CNA) as predicted by docking.

Work is underway in our laboratory to identify more glycomimetic molecules that inhibit ConA, to determine the modes of binding for the existing lead compounds, and to improve the potency of these compounds using information derived from molecular modeling and X-ray crystallography studies. The glycomimetic motifs from this work will also be tested with other lectins to examine their selectivities.

Acknowledgments

The authors wish to thank Dr. Michele Davis McGibony and Dr. Mary K. Boyd for their assistance with this manuscript. This work was supported by a grant from the Faculty Research Committee of Georgia Southern University.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2008.09.095.

References and notes

1. (a) Sharon, N.; Lis, H. *Science* **1989**, *246*, 4927. and references cited therein; (b) Hebert, D. N.; Garman, S. C.; Molinari, M. *Trends Cell. Biol.* **2005**, *15*, 364.
2. Gabius, H.-J.; Siebert, H.-C.; André, S.; Jiménez-Barbero, J.; Rüdiger, H. *ChemBioChem* **2004**, *5*, 740. and references cited therein..
3. Sharon, N.; Lis, H. *Glycobiology* **2004**, *14*, 53R.
4. Töpfer-Petersen, E. *Hum. Reprod. Update* **1999**, *5*, 314.
5. Kaila, N.; Thomas, B. E., IV *Med. Res. Rev.* **2002**, *22*, 566.
6. Salatino, M.; Croci, D. O.; Bianca, G. A.; Ilarregui, J. M.; Toscano, M. A.; Rabinovich, G. A. *Exp. Opin. Biol. Ther.* **2008**, *8*, 45.
7. Sharon, N. *Biochim. Biophys. Acta* **2006**, *1760*, 527.
8. Cormack, B. P.; Ghori, N.; Falkow, S. *Science* **1999**, *285*, 578.
9. Wiley, D. C.; Skehel, J . J. *Annu. Rev. Biochem.* **1987**, *56*, 365.
10. Osborn, H. M. I.; Evans, P. G.; Gemmell, N.; Osborne, S. D. *J. Pharm. Pharmacol.* **2004**, *56*, 691.
11. Jiménez-Barbero, J.; Lowary, T., Eds. *Carbohydr. Res.* **2007**, *342*.
12. (a) Schüttelkopf, A. W.; Andersen, O. A.; Rao, F. V.; Allwood, M.; Lloyd, C.; Eggleston, I. M.; van Aalten, D. M. F. *J. Biol. Chem.* **2006**, *281*, 27278; (b) Hurtado-Guerrero, R.; van Aalten, D. M. F. *Chem. Biol.* **2007**, *14*, 589.
13. (a) Slee, D. H.; Romano, S. J.; Yu, J.; Nguyen, T. N.; John, J. K.; Raheja, N. K.; Axe, F. U.; Jones, T. K.; Ripka, W. C. *J. Med. Chem.* **2001**, *44*, 2094; (b) Kranich, R.; Busemann, A. S.; Bock, D.; Schroeter-Maas, S.; Beyer, D.; Heinemann, B.; Meyer, M.; Schierhorn, K.; Zahlten, R.; Wolff, G.; Aydt, E. M. *J. Med. Chem.* **2007**, *50*, 1101.
14. (a) Gao, H.; Kawabata, J. *Bioorg. Med. Chem.* **2005**, *13*, 1661; (b) Tarling, C. A.; Woods, K.; Zhang, R.; Brastianos, H. C.; Brayer, G. D.; Andersen, R. J.; Withers, S. G. *ChemBioChem* **2008**, *9*, 433; (c) Lo Piparo, E.; Scheib, H.; Frei, N.; Williamson, G.; Grigorov, M.; Chou, C. J. *J. Med. Chem.* **2008**, *51*, 3555.
15. Mikkelsen, L. M.; Hernáis, M. J.; Martin-Pastor, M.; Skrydstrup, T.; Jiménez-Barber, J. *J. Am. Chem. Soc.* **2002**, *124*, 14941. and references cited therein.
16. (a) Page, D.; Zanini, D.; Roy, R. *Bioorg. Med. Chem.* **1996**, *4*, 1949; (b) Köhn, M.; Benito, J. M.; Mellet, C. O.; Lindhorst, T. K.; García Fernández, J. M. *ChemBioChem* **2004**, *5*, 771.
17. Scott, J. K.; Loganathan, D.; Easley, R. B.; Gong, X.; Goldstein, I. J. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5398.
18. Sumner, J. B.; Howell, S. F. *J. Bacteriol.* **1936**, *32*, 227.
19. Goldstein, I. J.; Poretz, R. D. In *The Lectins: Properties, Functions, and Applications in Biology and Medicine*; Liener, I. E., Sharon, N., Goldstein, I. J., Eds.; Academic Press, Inc.: Orlando, FL, 1986; pp 35–250.
20. Brewer, C. F.; Sternlicht, H.; Marcus, D. M.; Grollman, A. P. *Biochemistry* **1973**, *12*, 4448.
21. Goldstein, I. J. *Adv. Exp. Med. Biol.* **1975**, *55*, 35.
22. (a) Naismith, J. H.; Field, R. A. *J. Biol. Chem.* **1996**, *271*, 972; (b) Bouckaert, J.; Hamelryck, T. W.; Wyns, L.; Loris, R. *J. Biol. Chem.* **1999**, *274*, 29188.
23. Mandal, D. K.; Kishore, N.; Brewer, C. F. *Biochemistry* **1994**, *33*, 1149. and references cited therein.
24. Loris, R.; Stas, P. P. G.; Wyns, L. *J. Biol. Chem.* **1994**, *269*, 26722.
25. The NCI/DTP Open Chemical Repository, http://dtp.nci.nih.gov.
26. LigPrep, version 1.6, Schrödinger, LLC: New York, NY, 2005.
27. Glide, version 3.5, Schrödinger, LLC: New York, NY, 2005.
28. Naismith, J. H.; Emmerich, C.; Habash, J.; Harrop, S. J.; Helliwell, J. R.; Hunter, W. N.; Raftery, J.; Kalb, A. J.; Yariv, J. *Acta Crystallogr., Sect. D* **1994**, *50*, 847.
29. Harrop, S. J.; Helliwell, J. R.; Wan, T. C.; Kalb, A. J.; Tong, L.; Yariv, J. *Acta Crystallogr., Sect. D* **1996**, *52*, 143.
30. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. V. *J. Comput. Aided Mol. Des.* **1997**, *11*, 425.
31. Glide 3.5 Technical Notes; Schrödinger Press: Portland, OR, 2005.
32. Bytheway, I.; Cochran, S. J. *Med. Chem.* **2004**, *47*, 1683.