Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters
ELSEVIER
Bioorganic & Medicinal Chemistry Letters 14 (2004) 1437–1440

Interference of the galactose-dependent binding of lectins by novel pentapeptide ligands

Christopher J. Arnusch,[a] Sabine André,[b] Paola Valentini,[a] Martin Lensch,[b] Roland Russwurm,[b] Hans-Christian Siebert,[b] Marcel J. E. Fischer,[a] Hans-Joachim Gabius[b] and Roland J. Pieters[a,*]

[a]*Utrecht Institute for Pharmaceutical Sciences, Department of Medicinal Chemistry, Utrecht University, PO Box 80082, 3508 TB Utrecht, The Netherlands*
[b]*Institut für Physiologische Chemie, Tierärztliche Fakultät, Ludwig-Maximilians-Universität München, Veterinärstr. 13, 80539 München, Germany*

Received 14 October 2003; accepted 14 January 2004

Abstract—A library of pentapeptides containing the sequence –Y–X–Y– based on rational design was screened with six different lectins. Sequences were identified that modulate galectin binding to its natural carbohydrate ligand. SPR showed inhibition values 2–3 times stronger than galactose and NMR studies suggested real carbohydrate mimicry.

Protein–carbohydrate interactions are increasingly acknowledged for their role in various intra- and intercellular recognition processes.[1] Carbohydrate recognition is governed by stacking interactions and the more directional hydrogen bonds.[2] Protein–protein and peptide–protein interactions are controlled by similar intermolecular interactions. For this reason it is likely that specific peptide sequences should be able to bind to proteins with a carbohydrate recognition domain, a fundamental hypothesis with far-reaching basic and applied consequences.

Peptides capable of having strong interactions with carbohydrate binding proteins may have several applications in medicinal chemistry.[3] As enzyme inhibitors they could be useful laboratory tools and lead compounds for a new class of therapeutics that aim at modulating glycosylation. As adhesion inhibitors they could interfere with the initiation of bacterial and viral infections or undesired leukocyte accumulation. Furthermore, in cases where difficulty of immunologic protection by carbohydrate antigens is observed, the use of peptides that mimic carbohydrates may provide an alternative route for improved T-cell dependent immunization. Indeed, initial work in this area has supported the basic validity of this concept and its significant potential, and numerous examples of peptides with affinity to anti-carbohydrate antibodies have been identified.[3,4] However, only a few sequences with affinity for carbohydrate processing enzymes or other carbohydrate binding proteins, especially lectins, are known.[3c,5]

We report here on our search for short peptide sequences complementary to galactose-specific lectins. We deliberately tested two classes of lectins with separate folding patterns sharing specificity to D-galactose: the mistletoe (*Viscum album* L.) lectin with its β-trefoil motif and animal galectins.[6] Members of the galectin family of endogenous lectins are crucially involved in various physiological and pathophysiological reaction pathways such as immunomodulation, tumor growth regulation, invasion and metastasis and they also appear to interact with cytoplasmic/nuclear proteins guided by presently unknown characteristics of target site selection.[7]

A peptide library was used in search of glycomimetic substances. Tyr–Xaa–Tyr was a consensus sequence observed in a number of reports that showed these peptides to mimic different types of saccharides.[5,8]

* Corresponding author. Tel.: +31-3025-36944; fax.: +31-3025-36655; e-mail: r.j.pieters@pharm.uu.nl

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.01.029

Therefore this sequence was chosen as a basis for a rational library design. The positions of variation were chosen at aa-1[9] and aa-3,[10] while keeping aa-2 and aa-4 constant as a tyrosine, and aa-5 as a phenylalanine. With this approach, 20 mixtures of 5 peptides each were synthesized that included non-proteinogenic residues and many different classes of functional groups. Peptides in each mixture were chosen to have different molecular weights in order that electrospray mass spectrometry would indicate proper synthesis of the entire set of peptides.

The peptide mixtures were screened in a solid phase assay.[11,12] It included five galectins, the Viscum album agglutinin (VAA) and concanavalin A (ConA). The different galectin subtypes were represented, that is, three mammalian galectins from the prototype and chimera-type groups (that is, galectins-1 and -5 vs galectin-3) and the two known prototype chicken galectins (CG-14, CG-16). Biotin-labeled lectins were coincubated with the peptides (total concentration 5.6 mM) in order to test the peptides' ability to interfere with lectin binding to a glycoprotein matrix mimicking a cell surface. For the galactose-specific lectins, the matrix consisted of the glycoprotein asialofetuin (ASF), a potent ligand for galactose specific lectins, while yeast mannan was used for the ConA studies. Lectin binding to the matrices was completely dependent on the glycan but not the protein part of the glycoproteins used as ligands. From the 20 mixtures that were screened, 12 showed no effect, 6 showed an inhibitory effect with galactose specific lectin binding, and surprisingly, 2 mixtures showed an enhancement of the lectin-derived signal, that is, a stronger signal than in the absence of peptides. The strongest inhibitory mixtures brought the signal down to 10–20% residual galectin binding to the ASF coated plate. No inhibition of ConA (with yeast mannan) was seen. The 40 peptides of the mixtures of interest were resynthesized in order to evaluate their individual binding properties.[13] The results were visualized by using the TreeView program v. 1.60 developed by Eisen[14] (Fig. 1). The degree of specificity of the individual peptides for the different lectins can be easily compared by this visualization method. For example, peptide **7** binds to all the galactose-specific lectins but peptide **16** shows a preference for galectin-3 and CG-14.

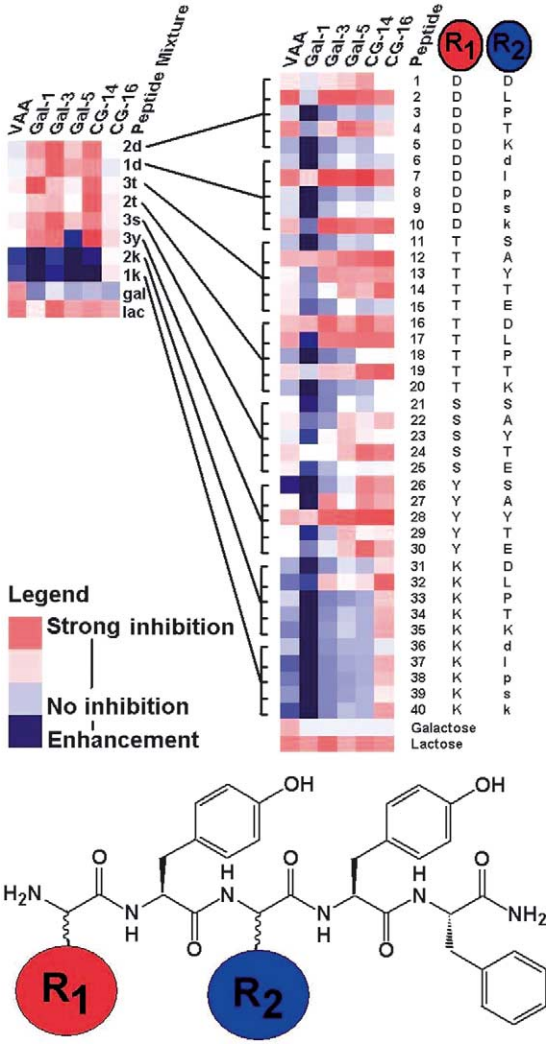

Figure 1. Summary of solid-phase screening showing mixtures of peptides (library screening) or single peptides: the gradient from blue–white–red indicates increasing lectin inhibition to ASF. Very dark blue indicates enhanced lectin binding. Shown also are data for lactose and galactose as a control for the galectins. R1 and R2 are given (lower case indicates D-form amino acid).

In order to confirm and extend the solid phase assay results three representative sequences were tested by surface plasmon resonance (SPR): a non-selective inhibitor peptide **7**, a more selective inhibitor peptide **16**, and a signal enhancing peptide **34**. They were synthesized[15] in two forms: as the free peptide and also linked to a solubility improving tail via the N-terminus (Fig. 2). Indeed, solubility was enhanced from 0.4–0.9 mM for the free peptides **7** and **16** to 7.5–10 mM for their N-extended counterparts. For the SPR experiments, a competition assay was developed. The lectins, which showed good binding to an ASF-functionalized SPR

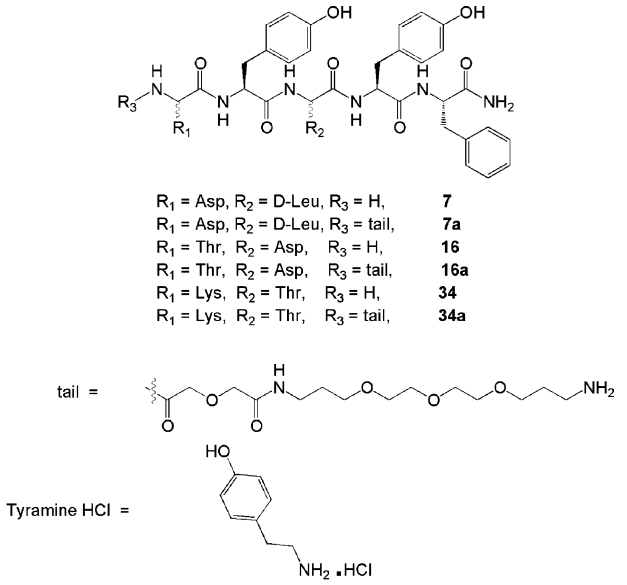

Figure 2. Structures of peptide–ligands measured with SPR in a competition assay versus Gal-1, CG-14, and VAA. Tyramine·HCl was used as a control.

chip, were exposed to different concentrations of the peptide. Inhibition was apparent from the signal reduction due to the presence of the peptides. Plotting the signal versus the concentration of peptide yielded the inhibition curve of **16a** for galectin-1 inhibition (see Fig. 3). Full inhibition was seen with peptides **7a** and **16a**. The results correlate well with the data of solid phase assay. Peptides **7** and **16** could not be subjected to SPR analysis due to poor solubility. Peptide **7a** was found to have an IC_{50} for galectin-1 of 8.0 mM and for CG-14 of 6.9 mM. For peptide **16a** these values were 5.2 and 7.5 mM, respectively. For comparison, the IC_{50} of lactose and galactose for galectin-1 inhibition were 0.2 and 12 mM. Control experiments were run with tyramine (Fig. 2) at up to 90 mM and no inhibition was observed with Gal-1. To infer binding of peptides in or near the carbohydrate-specific site, we employed the laser photo-CIDNP (chemically induced dynamic nuclear polarization) technique. Lactose weakens or even abolishes the signal of the tryptophan in the galectins' binding site, reflecting the known stacking/C-H/π-interactions.[16] Using peptide **16a** and galectin-3 the signal at 6.20 ppm originating from H6 of the indolyl ring of Trp 194, which was susceptible to lactose presence, was significantly reduced at a peptide: galectin ratio of 10:1 (6 mM and 0.6 mM), arguing in favor of the given assumption. The observation of enhancement of lectin binding in the solid phase assay was also seen in the SPR assay for peptides **34** and **34a**. In the case of galectin-1, maximum enhancement was observed in the 5 to 10 mM range for both peptides with the enhancement being around 20% of the lectin equilibrium binding signal. Similar observations were made with CG-14.

In conclusion, short peptide sequences (that is, 5-mers) that harbor the Tyr–Xaa–Tyr sequence were found to inhibit protein–carbohydrate binding of various galactose specific lectins with IC_{50} values in the 4–8 mM range. This is ~2–3 times stronger than galactose. A solid phase assay was useful for initial routine screening whereas SPR technology allowed extended analysis of selected peptides.

Regarding active sequences especially the polar residues Asp, Thr and Lys were often found. Considering the abundance of aromatic residues already present in the peptides, the polar residues likely complement the aromatic ones through electrostatic interactions and hydrogen bonds. The sequences appear to be true glycomimetics, that is, they do not merely target the vicinity of the galactose-specific site since binding was observed to both plant and animal lectins whose binding sites are known to differ.[2d,17] Furthermore specificity for the different galectins was observed in some cases and no significant inhibition was seen with the mannose-specific ConA, another argument for inherent selectivity. Besides inhibition we also observed signal enhancement by certain peptides. This was observed both in solid-phase inhibition and the SPR assay. The phenomenon of signal enhancement had been seen before with multivalent mannose ligands binding to ConA.[18] In this case it could be explained by the nature of the multivalent ligand which can cross-link several lectin molecules, a feature also implicated in galectin signaling. The ensuing result is aggregation which brings more lectin molecules to the chip than without a multivalent ligand present, hence enhancing the signal. Whether 5-mers might somehow act as cross-linkers or act by a different mechanism will be subject of further study. Efforts will also be directed towards the elucidation of the target sites and the binding mode of the defined 5-mer peptide sequences to the galactose-binding lectins using various techniques and also towards further optimization of the structure of the non-carbohydrate ligand. The finding that peptide sequences have the capacity to associate with lectins will broaden our insights into their functionality considerably, since proteins that contain these sequences may be shown to interact with mammalian lectins as well.[19]

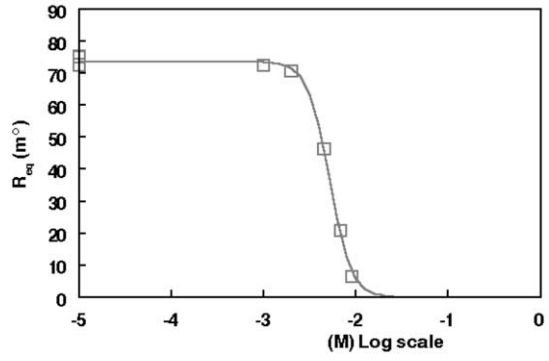

Figure 3. Inhibition of 1 μM galectin-1 binding to an ASF functionalized SPR chip. The equilibrium SPR signal is displayed versus increasing concentrations of peptide **16a**.

Acknowledgements

This work was supported by The Royal Netherlands Academy of Arts and Sciences (KNAW). Prof. R. M. J. Liskamp is gratefully acknowledged for support and advice. We thank Reshma Autar for synthesis of the solubility improving tail.

References and notes

1. (a) Reuter, G.; Gabius, H.-J. *Cell. Mol. Life Sci.* **1999**, *55*, 368. (b) Gabius, H.-J.; André, S.; Kaltner, H.; Siebert, H.-C. *Biochim. Biophys. Acta* **2002**, *1572*, 165. (c) Roseman, S. *J. Biol. Chem.* **2002**, *276*, 41527.
2. (a) Quiocho, F. A. *Pure Appl. Chem.* **1989**, *61*, 1293. (b) Rini, J. M. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 551. (c) Gabius, H.-J. *Pharmaceut. Res.* **1998**, *15*, 23. (d) Rüdiger, H.; Siebert, H.-C.; Solís, D.; Jiménez-Barbero, J.; Romero, A.; von der Lieth, C.-W.; Díaz-Mauriño, T.; Gabius, H.-J. *Curr. Med. Chem.* **2000**, *7*, 389.
3. (a) For recent reviews see Apostolopoulos, V.; Sandrin, M. S.; McKenzie, I. F. C. *J. Mol. Med.* **1999**, *77*, 427. (b) Johnson, M. A.; Pinto, B. M. *Aust. J. Chem.* **2002**, *55*, 13. (c) Monzavi-Karbassi, B.; Cunto-Amesty, G.; Luo, P.; Kieber-Emmons, T. *Trends Biotechnol.* **2002**, *20*, 207.
4. (a) Taki, T.; Ishikawa, D.; Hamasaki, H.; Handa, S. *FEBS Lett.* **1997**, *418*, 219. (b) Fukuda, M. N.; Ohyama, C.; Lowitz, K.; Matsuo, O.; Pasqualini, R.; Ruohslahti,

E.; Fukuda, M. *Cancer Res*. **2000**, *60*, 450. (c) Kieber-Emmons, I. O. T.; Otvos, L.; Blaszczyk-Thurin, M. *Biochem. Biophys. Res. Commun*. **2000**, *268*, 106. (d) Houston, D. R.; Shiomi, K.; Arai, N.; Omura, S.; Peter, M. G.; Turberg, A.; Synstad, B.; Eijsink, V. G. H.; van Aalten, D. M. F. *Proc. Natl. Acad. Sci. U.S.A*. **2002**, *99*, 9127. (e) Lee, K.-Y.; Kim, H. G.; Hwang, M. R.; Chae, J. I.; Yang, J. M.; Lee, Y. C.; Choo, Y. K.; Lee, Y. I.; Lee, S.-S.; Do, S.-I. *J. Biol. Chem*. **2002**, *277*, 49341. (f) Molenaar, T. J. M.; Appeldoorn, C. C. M.; de Haas, S. A. M.; Michon, I. N.; Bonnefoy, A.; Hoylaerts, M. F.; van Berkel, T. J. C.; Kuiper, J.; Biessen, E. A. L. *Blood* **2002**, *100*, 3570. (g) Appeldoorn, C. C. M.; Molenaar, T. J. M.; Bonnefoy, A.; van Leeuwen, S. H.; Vandervoort, P. A.; Hoylaerts, M. F.; van Berkel, T. J.; Biessen, E. A. L. *J. Biol. Chem*. **2003**, *278*, 10201.

5. (a) Oldenburg, K. R.; Loganathan, D.; Goldstein, I. J.; Schultz, P. G.; Gallop, M. A. *Proc. Natl. Acad. Sci. U.S.A*. **1992**, *89*, 5393. (b) Scott, J. K.; Loganathan, D.; Easley, R. B.; Gong, X.; Goldstein, I. J. *Proc. Natl. Acad. Sci. U.S.A*. **1992**, *89*, 5398. (c) Kaur, K. J.; Khurana, S.; Salunke, D. M. *J. Biol. Chem*. **1997**, *272*, 5539. (d) Ravishankar, R.; Thomas, C. J.; Suguna, K.; Surolia, A.; Vijayan, M. *PROTEINS* **2001**, *43*, 260.
6. (a) Lee, R. T.; Gabius, H.-J.; Lee, Y. C. *J. Biol. Chem*. **1992**, *267*, 23722. (b) Galanina, O. E.; Kaltner, H.; Khraltsova, L. S.; Bovin, N. V.; Gabius, H.-J. *J. Mol. Recogn*. **1997**, *10*, 139. (c) Gabius, H.-J.; Darro, F.; Remmelink, M.; André, S.; Kopitz, J.; Danguy, A.; Gabius, S.; Salmon, I.; Kiss, R. *Cancer Invest*. **2001**, *19*, 114.
7. (a) Kaltner, H.; Stierstorfer, B. *Acta Anat*. **1998**, *161*, 162. (b) Gabius, H.-J. *Anat. Histol. Embryol* **2001**, *30*, 3. (c) Danguy, A.; Camby, I.; Kiss, R. *Biochim. Biophys. Acta* **2002**, *1572*, 285. (d) Liu, F.-T.; Patterson, R. J.; Wang, J. L. *Biochim. Biophys. Acta* **2002**, *1572*, 274. (e) Rabinovich, G. A.; Rubinstein, N.; Toscano, M. A. *Biochim. Biophys. Acta* **2002**, *1572*, 274.
8. Westerink, M. A. J.; Giardina, P. C.; Apicella, M. A.; Kieber-Emmons, T. *Proc. Natl. Acad. Sci. U.S.A*. **1995**, *92*, 4021.
9. Amino Acid variability unit 1: Ala, Asp, Leu, Pro, Ser, Lys, Thr, Tyr, Glu.
10. Amino Acid variability unit 3: Ala, D-Ala, Asp, D-Asp, Leu, D-Leu, Pro, D-Pro, Ser, D-Ser, Lys, D-Lys, Thr, Tyr, Glu, D-Glu, D-Phe.
11. (a) André, S.; Pieters, R. J.; Vrasidas, I.; Kaltner, H.; Kuwabara, I.; Liu, F.-T.; Liskamp, R. M. J.; Gabius, H.-J. *ChemBioChem*. **2001**, *2*, 822. (b) Vrasidas, I.; André, S.; Valentini, P.; Böck, C.; Lensch, M.; Kaltner, H.; Liskamp, R. M. J.; Gabius, H.-J.; Pieters, R. J. *Org. Biomol. Chem*. **2003**, *1*, 803.
12. (a) Gabius, H.-J. *Anal. Biochem*. **1990**, *189*, 91. (b) André, S.; Unverzagt, C.; Kojima, S.; Dong, X.; Fink, C.; Kayser, K.; Gabius, H.-J. *Bioconjugate Chem*. **1997**, *8*, 845.
13. After precipitation, peptide purities were estimated as >90% using analytical HPLC. Peptides were tested at 5.6 mM. In order to reach this concentration 4% EtOH was used to enhance solubility when needed.
14. Eisen, M. B.; Spellman, P. T.; Brown, P. O.; Botstein, D. *Proc. Natl. Acad. Sci. U.S.A*. **1998**, *95*, 14863.
15. **7a**, **7a**, **34**, **34a**, **16**, **16a** were analyzed as pure peptides on HPLC. **7**: 19.65 min.; MS *m*/*z*: 719.5 $[M+H]^+$. **7a**: 19.25 min; MS *m*/*z*: 1037.7 $[M+H]^+$. **34**: 17.28 min; MS *m*/*z*: 720.8 $[M+H]^+$. **34a**: 18.02 min; MS *m*/*z*: 1038.7 $[M+H]^+$. **16**: 17.18 min; MS *m*/*z*: 707.5 $[M+H]^+$. **16a**: 17.88 min; MS *m*/*z*: 1026.5 $[M+H]^+$.
16. Siebert, H.-C.; Adar, R.; Arango, R.; Burchert, M.; Kaltner, H.; Kayser, G.; Tajkhorshid, E.; von der Lieth, C.-W.; Kaptein, R.; Sharon, N.; Vliegenthart, J. F. G.; Gabius, H.-J. *Eur. J. Biochem*. **1997**, *249*, 27.
17. Solís, D.; Romero, A.; Kaltner, H.; Gabius, H.-J.; Díaz-Mauriño, T. *J. Biol. Chem*. **1996**, *271*, 12744.
18. Burke, S. D.; Zhao, Q.; Schuster, M. C.; Kiessling, L. L. *J. Am. Chem. Soc*. **2000**, *122*, 4518.
19. For an example see: Ravindranath, R. M. H.; Tam, W.-Y.; Nguyen, P.; Fincham, A. G. *J. Biol. Chem*. **2000**, *275*, 39654.