Ligand amplification in a dynamic combinatorial glycopeptide library†

Tom Hotchkiss, Holger B. Kramer, Katie J. Doores, David P. Gamblin, Neil J. Oldham and Benjamin G. Davis*

Received (in Cambridge, UK) 2nd June 2005, Accepted 28th June 2005
First published as an Advance Article on the web 9th August 2005
DOI: 10.1039/b507674a

***N*-acetyl glucosamine binding protein amplifies the concentration of one member in a dynamic combinatorial glycopeptide library based on exchanging disulfides.**

Dynamic combinatorial chemistry represents an evolutionary approach to molecular design and therefore has the potential to offer a very efficient route to specific inhibitors, catalysts or receptor ligands.[1] In order to achieve full potential, a library of building blocks should exchange efficiently to form its constituents by reversible reactions. Introduction of a template to the thermodynamic equilibrium mixture results in binding of suitable ligands. Stabilization of the bound states leads to a shift in equilibrium favouring the respective library members and results in their amplification. Other members, which do not bind well but contain common building blocks, are at the same time reduced in concentration. Determination of the changes in composition of the library requires a means of concurrently monitoring the concentrations of all library members.

In this communication we wish to report our efforts towards a dynamic combinatorial glycopeptide library based on 1-thiosugars and cysteine containing dipeptides. These studies are the first results in a programme aimed at using DCL approaches to influence and detect dynamic protein glycosylation (Fig. 1).‡ Glycopeptides **C** and **G** (Fig. 2) were synthesized (Scheme 1) starting from suitably protected amino acids using solution phase peptide coupling methods, followed by reduction of the resulting disulfides with Bu_3P and formation of the disulfide-linked glycopeptides using Glyco-SeS methodology.[2] It should be noted that the minimal use of protecting groups in peptide syntheses and the ability to employ entirely unprotected sugars in a convergent strategy allowed syntheses in overall yields of 12–13% from parent amino acids and carbohydrates.

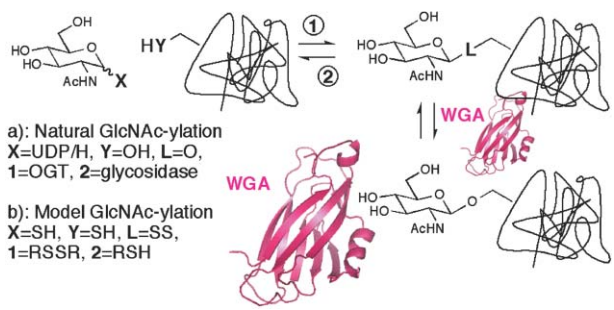


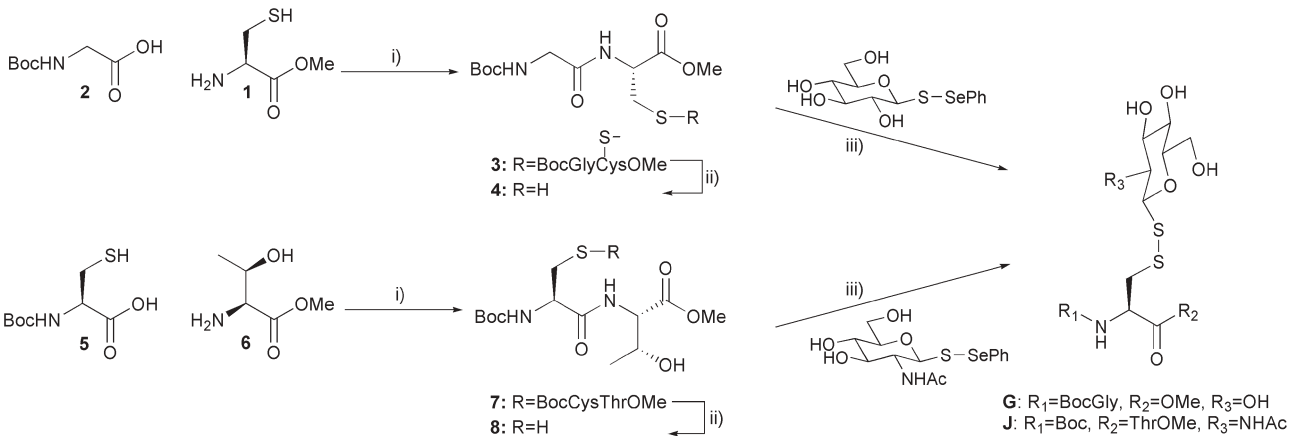
Fig. 1 Dynamic Protein GlcNAc-ylation may be considered to be an example of a natural DCL.

Department of Chemistry, University of Oxford, Mansfield Road, Oxford, UK OX1 3TA. E-mail: Ben.Davis@chem.ox.ac.uk; Fax: +44 (0)1865 285002; Tel: +44 (0)1865 275652
† Electronic supplementary information (ESI) available: Experimental details. See http://dx.doi.org/10.1039/b507674a

The conditions for the formation of a dynamic combinatorial library based on these well-defined starting glycopeptides were monitored using electrospray ionization mass spectrometry (ESI-MS) in positive ion mode and established as follows. Glycopeptides containing D-glucose (Glc) and *N*-acetyl-D-glucosamine (GlcNAc), GlyCys(Glc) **C** and Cys(GlcNAc)Thr **G**, respectively, were each stirred in NH_4OAc buffer (pH 7.5), containing catalytic dithiothreitol (DTT, ~0.2 equiv.) for efficient disulfide exchange, for 48 h. Under equivalent conditions at pH 4.5 (4% formic acid) no disulfide exchange was observed after 48 h or even after 14 days. Therefore, reaction mixtures were quenched and disulfide exchange halted by addition of formic acid prior to analysis. Simultaneous observation of the peaks of all members of the resulting DCL in a single spectrum required calibration of cone voltage and capillary voltage of the mass spectrometer. Gratifyingly, starting from just **C** and **G**, nine out of ten possible disulfide exchange products **A–M** were identified in the library under the chosen conditions (Fig. 2).

Having established suitable conditions for the development and analysis of libraries, we turned our attention towards the concentration dependence of the ion count obtained. While a clear dependence of ion count on the concentration of glycopeptides **C** and **G** could be observed, these measurements do not provide a reliable measure of *absolute* concentration of the respective compound. Therefore the ratio of peak height of the observed compound to the sum of peak heights of the remaining members of the library was taken as a measure of *relative* concentration. This internal calibration allows observation of changes in relative proportions of compounds in the library to be made. Time-dependant measurements of DCLs derived from **A–M** at room temperature demonstrated that disulfide exchange, in fact, led to formation of mixtures containing the expected disulfide exchange products within 10 h at pH 7.5.

After successfully establishing formation and monitoring of the glycopeptide libraries, their use as probes of binding was investigated. The dynamic addition (so-called 'GlcNAc-ylation') and removal of GlcNAc to and from Ser/Thr is a process of emerging importance in regulating the activity of proteins ranging from nucleoporins to transcription factors and implicated in gibberellin signalling and diabetes.[3] While elegant methods for tracking the existence of GlcNAc-modified amino acids at a given point in time in proteins have been described[4] to the best of our knowledge no models of the dynamic addition and removal have been proposed. With the longer term aim of creating and understanding this process ultimately in proteins, we first wished



Scheme 1 (i) EDDQ, Et_3N, DCM, **3**: 84%, **7**: 60%, (ii) $^nBu_3P$, THF–MeOH, **4**: 95%, **8**: 99%, (iii) Et_3N, MeOH, **C**: 43%, **G**: 91%.

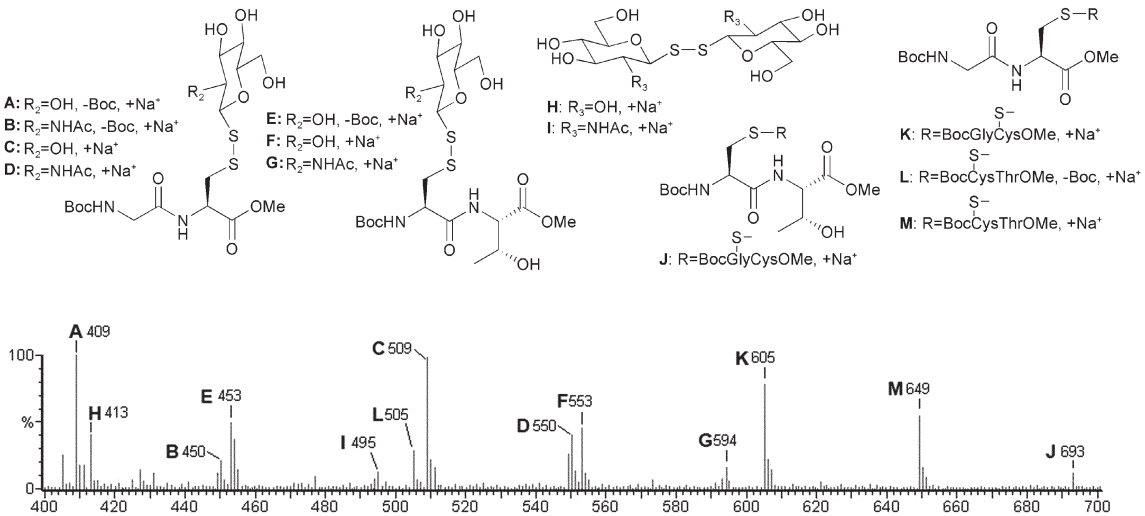


Fig. 2 Glycopeptide DCL members and their observation by ESI-MS.

to create a chemical analogue relying on the reversibility of disulfide bond formation and that could be interrogated by one of the same sugar-binding proteins used to detect natural GlcNAc-ylation, wheat germ agglutinin (WGA).[5] WGA (14.7 nmol) was added to glycopeptide DCLs (~206 and 168 nmol of starting components **C** and **G** respectively) containing both dynamically glucosylated and GlcNAc-ylated peptides that had been allowed to exchange for 78 h; significant changes of the relative concentrations of select members of the library were observed. Specifically, an increase in the concentration of GlcNAc-ylated glycopeptide **B**, and a less significant increase in the concentration of GlcNAc-ylated glycopeptide **G** were observed, while at the same time the glucosylated glycopeptide **F** showed a significant decrease in concentration (Fig. 3).

This is an important result which is coherent with the known carbohydrate specificity of WGA.[6] The observed amplification of **B** can be understood in terms of relative stabilisation of the bound species and consequent alteration in the reequilibration of the reaction mixture. Upon quenching of the samples prior to electrospray mass spectrometric determination, the lectin–glycopeptide complex is expected to break down and the determined concentration of the selected glycopeptide is larger

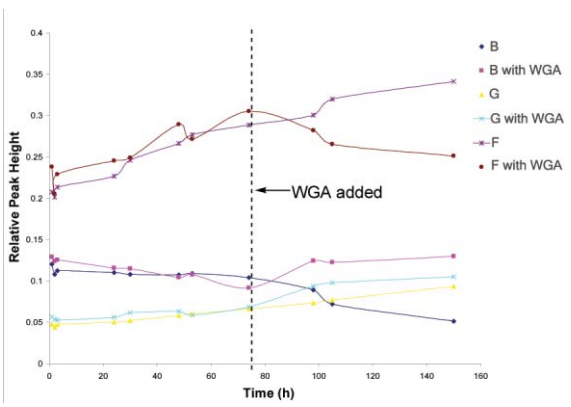


Fig. 3 Change in the relative concentration of **B**, **F** and **G** on addition of WGA.

than the corresponding system prior to WGA addition. Interestingly, the much stronger amplification of **B**, rather than the other GlcNAc-containing species in the DCL, suggests that this glycopeptide with a Gly–Cys backbone is bound more strongly than GlcNAc-ylated glycopeptide **G** with a Cys–Thr

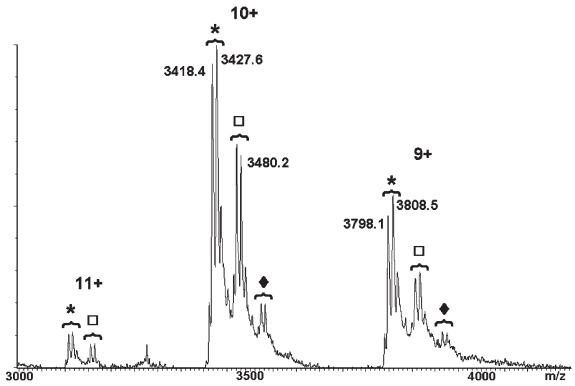


Fig. 4 ESI-MS analysis of WGA lectin following exposure to the DCL. * = WGA dimer; □ = WGA dimer + glycopeptide **B**; ◆ = WGA dimer + 2 × glycopeptide **B**. Peak multiplicity is due to the heterogeneous nature of the two WGA chains.

backbone or even 1,1-linked diGlcNAc species **I**. This illustrates that WGA is not only GlcNAc specific but in the context of GlcNAc-ylation the peptide backbone may have a strong additional influence.

We sought to confirm directly the selection and binding of GlcNAc-ylated glycopeptide **B** by WGA. The lectin-substrate complexes $(WGA)_2$ + **B** and $(WGA)_2$ + $\mathbf{B}_2$ were directly observed by native mass spectrometry[7] at approximately 526 ± 2 Da higher than the free lectin (Fig. 4). This striking result highlights the strong potential of native MS in DCL and GlcNAc-ylation investigation, confirms the observed amplification of glycopeptide **B** (M_r 527) and highlights the multiplicity of binding of WGA that further enhances its role as an interrogating protein.

While the detailed reasons for the clear selection of glycopeptide **B** in the presence of other GlcNAc-ylated ligand possibilities remains to be explored, the success of the developed methodology opens clear possibilities to explore further not only GlcNAc-ylation processes using such techniques but the influence of the peptide backbone in glycopeptides on binding to other lectins. Our results confirm the usefulness of a dynamic combinatorial approach to the investigation of binding preferences in lectins[8] and in detection of peptide glycosylation.

In conclusion, we have established conditions for the successful formation of a glycopeptide library based on disulfide exchange of 1-thiosugars with cysteine containing peptides and monitoring by ESI-MS. Selective amplification of a glycopeptide component by lectin was observed for the first time. The lectin–ligand complex was identified directly by native mass spectrometry. The potential of this method is currently being used to interrogate chemically and naturally GlcNAc-ylated protein systems.

Notes and references

‡ First details of this work were described at the GlycoTrain network Seville meeting, 12 December 2003. For an example of mixed disulfide library formation that shows component binding but not amplification or selection see: S. Sando, A. Narita and Y. Aoyama, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 2835. Unfortunately, in this prior work none of the individual members were individually synthesised; in our hands this mixture-based approach carries ambiguity in results.

1 G. R. L. Cousins, S.-A. Poulsen and J. K. M. Sander, *Curr. Opin. Chem. Biol.*, 2000, **4**, 270.
2 D. P. Gamblin, P. Garnier, S. van Kasteren, N. J. Oldham, A. J. Fairbanks and B. G. Davis, *Angew. Chem., Int. Ed.*, 2004, **43**, 828.
3 (*a*) K. Vosseller, K. Sakabe, L. Wells and G. W. Hart, *Curr. Opin. Chem. Biol.*, 2002, **6**, 851–857; (*b*) G. Hart, K. Greis, L. Dong, M. Blomberg, T. Chou, M. Jiang, E. Roquemore, D. Snow, L. Kreppel and R. Cole, *Adv. Exp. Med. Biol.*, 1995, **376**, 115; (*c*) W. Lubas and J. Hanover, *J. Biol. Chem.*, 2000, **275**, 10983.
4 (*a*) D. J. Vocadlo, H. C. Hang, E.-J. Kim, J. A. Hanover and C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 9116; (*b*) N. Khidekel, S. Arndt, N. Lamarre-Vincent, A. Lippert, K. G. Poulin-Kerstien, B. Ramakrishnan, P. K. Qasba and L. Hsieh-Wilson, *J. Am. Chem. Soc.*, 2003, **125**, 16162–16163.
5 C. S. Arnold, G. V. W. Johnson, R. N. Cole, D. L.-Y. Dong, M. Lee and G. W. Hart, *J. Biol. Chem.*, 1996, **271**, 28741.
6 M. M. Burger and A. R. Goldberg, *Proc. Natl. Acad. Sci. USA*, 1967, **57**, 359; A. K. Allen, A. Neuberger and N. Sharon, *Biochem. J.*, 1973, **131**, 155; I. J. Goldstein, S. Hammarstrom and G. Sundblad, *Biochim. Biophys. Acta*, 1975, **405**, 63.
7 R. H. H. van den Heuvel and A. J. R. Heck, *Curr. Opin. Chem. Biol.*, 2004, **8**, 519.
8 For a leading example of lectin binding in DCLs see O. Ramstrom and J. M. Lehn, *ChemBioChem*, 2000, **1**, 41.

