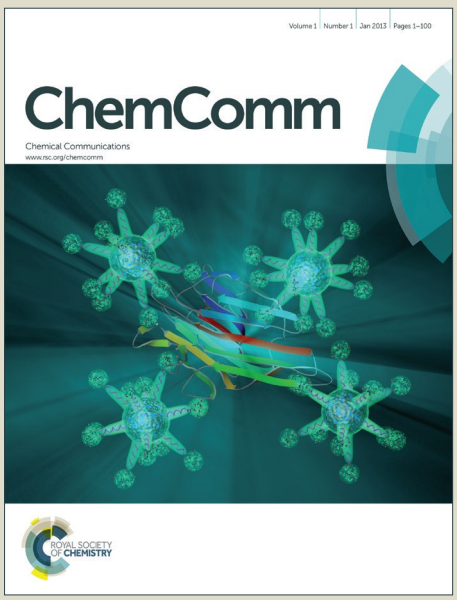
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

A rigid lanthanide binding tag to aid NMR studies of a 70 kDa homodimeric coat protein of human norovirus

Alvaro Mallagaray,[a] Gema Domínguez,[b] Thomas Peters*[a] and Javier Pérez-Castells*[b]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

Attachment of human noroviruses to histo blood group antigens is thought to be essential for infection of host cells. Molecular details of the attachment process can be studied *in vitro* using a variety of NMR experiments. The use of protein NMR based experiments requires assignments of backbone NMR signals. Using uniformly $^{2}H,^{15}N$-labeled protruding domains (P-dimers) of a prevalent epidemic human norovirus strain (GII.4 Saga) we have studied the potential of α-L-fucose covalently linked to a rigid lanthanide binding tag to aid backbone assignments using the paramagnetic properties of lanthanide ions. The synthesis of tagged α-L-fucose is reported. Notably, the metal chelating unit connects to the carbohydrate via a triazole linker constructed using click chemistry.

Norovirus is a non-enveloped (+)-single stranded RNA virus belonging to the family of *Caliciviridae*. The virus is highly contagious, and norovirus infection constitutes the most common cause of acute non-bacterial gastroenteritis in humans.[1] Norovirus is also a prominent example for carbohydrate-mediated host cell attachment.[2-4] Details of attachment and subsequent cell-entry are largely unknown as is the identity of cells susceptible to infection.[5] It is thought that attachment of noroviruses occurs via interaction with Histo Blood Group Antigens (HBGAs)[3, 6] located on the host cell surface. Previous work has shown that α-L-fucose is the minimal binding motif for genotype II (GII) noroviruses.[7] In these studies STD NMR experiments delivered binding epitopes at atomic resolution utilizing virus like particles (VLPs) as targets.[8]

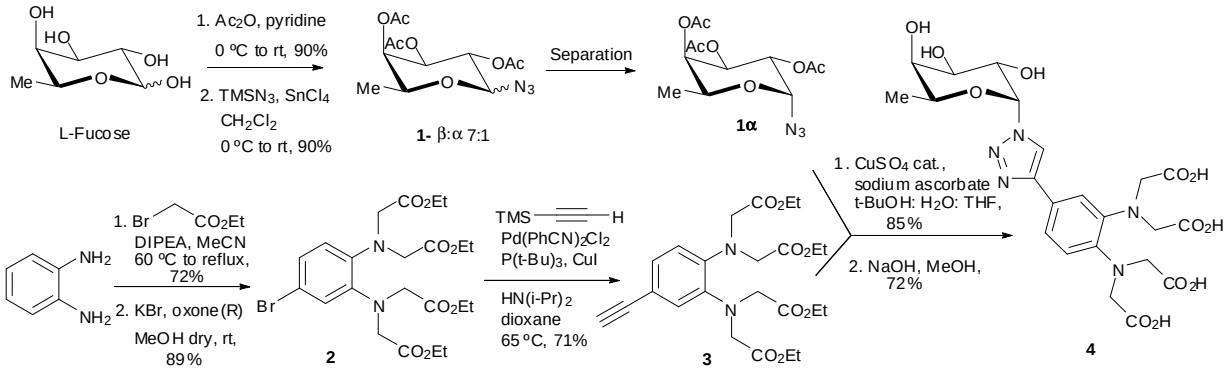
Expression of the protruding domain (P-domain) of human norovirus in *E.coli* yields so-called P-dimers, homodimers with a molecular weight of ca. 70 kDa. It has been shown that P-dimers represent the minimal functional entity for HBGA recognition.[9] Therefore, P-dimers have frequently been employed for studies into norovirus-HBGA interactions. We have prepared uniformly $^{2}H,^{15}N$ labeled P-dimers of a currently prevalent GII.4 Saga strain to allow for protein NMR experiments such as chemical shift perturbation experiments.[10] Employing e.g. $^{1}H,^{15}N$ TROSY HSQC experiments[11] chemical shift perturbation discloses binding site topologies of protein receptors. In fact, such experiments have been instrumental for the recent discovery of a novel paradigm for norovirus-HBGA recognition. Chemical shift perturbation experiments in combination with STD NMR experiments and native mass spectrometry have shown that binding of HBGAs to noroviruses is a complex cooperative process with at least four binding pockets for HBGAs per P-dimer.[10] It appears that the virus has developed a tunable carbohydrate recognition system, and it will be very interesting to understand the underlying molecular principles.

Currently, we explore the potential of protein NMR spectroscopy to gain insight into the mechanistic principles of norovirus-carbohydrate recognition. For this we will need at least a partial backbone assignment of the P-dimers. At a molecular weight of 70 kDa this is a significant challenge and will certainly require application of novel assignment strategies based on e.g. sparse labeling or paramagnetic tags. The focus of this study is on exploring paramagnetic tags tailored for applications targeting norovirus P-dimers.

Paramagnetic metal ions cause enhanced nuclear spin relaxation (PRE), pseudocontact shifts (PCS), and molecular alignment in the presence of magnetic fields resulting in residual dipolar couplings (RDCs).[12-17] When tagged to a protein, a broad range of paramagnetic properties of lanthanides offers opportunities for structure determination, alternative assignment strategies, binding studies and insights into protein dynamics. Two recent reviews summarize these options.[18] Here, we follow a different strategy for paramagnetic tagging where the paramagnetic label is covalently attached to a carbohydrate ligand. This approach has been shown to deliver valuable information on the conformation of carbohydrates.[19-22] Importantly, in combination with protein NMR spectroscopic approaches paramagnetically tagged carbohydrates give access to information on the binding site not easily available from other experimental techniques.[23] In the following we describe the synthesis of an α-L-fucose derivative carrying a chelating unit for lanthanide ions and serving as a non-covalent paramagnetic probe to assist assignment of norovirus P-dimers, and to pave the way for novel NMR experiments to unravel the complex recognition of HBGAs.

Design and synthesis of a suitable carbohydrate ligand has been guided by two thoughts. First, fucose is the minimal binding moiety recognized by norovirus P-dimers, and thus is ideally suited as a "minimal ligand" targeting the HBGA binding sites of P-dimers. Second, lanthanide binding tags (LBT) covalently attached to carbohydrates have been explored by us before,[20] and we utilize structural elements proven to be beneficial for paramagnetic NMR applications. In our prior work the metal ion

was chelated either to an EDTA or PhDTA moiety. In order to avoid reduction of PCS due to motional averaging, the metal was separated from the biomolecule by a rigid linker. For this project we envisioned the use of a PhDTA chelating unit connected to L-fucose through a triazole ring that is accessible by click chemistry. Thus, L-fucose azide was prepared using literature methods.[24] A separable mixture of anomers was obtained. The complementary ethynyl derivative **3** was prepared using a Sonogashira coupling reaction.[25] As synthesis of the parent iododerivative proved inefficient we tried coupling with bromine derivative **2**, which was readily prepared from phenylenediamine in two steps.[20] As arylbromides are less reactive in carbon-carbon coupling reactions we elevated the reaction temperature and extended the reaction time to obtain the trimethylsilyl intermediate. This intermediate was deprotected without purification furnishing **3** in 71% yield after the two steps. The target L-fucose derivative **4** was synthesized through a click reaction between α-L-fucose azide **1α** and the ethynyl-phDTA protected chelating unit **3,** followed by hydrolysis of the esters. All condensations and hydrolysis steps proceeded with high yields (Scheme 1).



Scheme 1 Synthesis of tagged L-fucose **4**

In order to survey its suitability as a tool for the advancement of backbone assignments, fucose derivative **4** was subjected to NMR experiments in the absence and presence of P-dimers.

Monitoring a titration of **4** with $DyCl_3$ by ^1H NMR spectra shows that exchange of the lanthanide ion is in slow exchange on the chemical shift time scale as two sets of ^1H NMR signals are observed (Fig. S1). This is due to the high binding affinity of the chelating PhDTA moiety for lanthanides and is a prerequisite for any paramagnetic NMR experiment. Noteworthy, large PCS (ca. 4.8 ppm for the anomeric ^1H) were observed on the L-Fucose moiety when the chelating unit was loaded with Dy^{3+}, indicating a highly rigid structure (Fig. S2).

Next, STD NMR titrations[26, 27] were performed to estimate the dissociation constants of **4** binding to P-dimers in the presence and in the absence of any lanthanide ions. From simple ^1H NMR spectra it is obvious that **4** without any lanthanide ions chelated forms aggregates via the chelating moieties at concentrations above ca. 600 μM (Fig. S3). Therefore, no binding isotherms from STD NMR titrations were evaluated. When $LaCl_3$ is chelated by the PhDTA moiety (**4**/La^{3+}) ^1H NMR spectra give no indication of aggregation up to concentrations of **4**/La^{3+} over 2.5 mM (Fig. S4). Above this concentration, aggregation outweighs chelation and aggregation occurs with loss of the metal ion. In parallel, the protein precipitates at these ligand concentrations. In fact, STD NMR experiments indicate a partial aggregation at concentrations over c.a. 900 μM leading to direct irradiation artifacts (data not shown). The titrations yielded only an estimate for the dissociation constant K_D of 0.68 ± 0.12 mM (Fig. S5).

An epitope mapping of **4**/La^{3+} at 900 μM using STD NMR experiments shows that the fucose moiety is mediating the key contacts with the binding pocket (Fig. S6).

As shown in our previous study[10] and as supported by a subsequent crystal structure analysis[6] P-dimers have at least four non-equivalent binding sites exhibiting cooperative binding, which are occupied in a step-wise manner. Aggregation of **4** at concentrations at which this complex binding behavior becomes detectable prohibits further studies into cooperative binding with STD NMR.

In order to accurately calculate the K_D for the first binding event and to estimate whether a single binding pocket could be selectively saturated using **4** loaded with paramagnetic metals, a uniformly ^2H,^{15}N labeled sample of P-dimers was prepared. First, a chemical shift titration with **4**/La^{3+} was conducted employing ^1H,^{15}N TROSY HSQC experiments. As mentioned above precipitation of P-dimers at ligand concentrations above 2.5 mM prevents use of larger ligand concentrations as used in our initial binding study with methyl α-L-fucopyranoside. Therefore, chemical shift perturbations were monitored up to a ligand concentration of 2.5 mM. These experiments showed that the first binding site from the P dimers is exclusively occupied up to 1.75 mM of **4**/La^{3+}. Beyond this threshold a second binding event is observed as expected.[10] Therefore, 1.75 mM was selected as the maximum ligand concentration, which results in selective binding of **4**/La^{3+} to the first HBGA binding pocket. Interestingly, these curves indicate cooperative binding behavior for binding to the first site, and yield a dissociation constant K_D of 0.94 ± 0.04 mM.

It is of note, that the same cross peaks are affected upon titration with **4**/La^{3+} as observed for methyl α-L-fucopyranoside. This strongly suggests binding to the same site (for details see Fig. S10). Chemical shift perturbations with **4**/La^{3+} are larger as with methyl α-L-fucopyranoside, and exhibit a different direction due to the presence of the aromatic rings attached to the reducing end.

To summarize, NMR binding experiments show that **4**/La^{3+} at concentrations up to 1.75 mM specifically and exclusively binds

to the primary HBGA binding pocket of P-dimers. This ensures the presence of a single metal atom per P-dimer at a known averaged position during the binding event, easing the calculations required for backbone assignment based on paramagnetic effects.

In order to estimate the positioning of the paramagnetic tag relative to the protein we built a simple model based on crystal structures of GII.4 Saga P dimers complexed with HBGAs[6], and of PhDTA complexed with Fe^{3+}.[28] This model shows that the rigid paramagnetic tag is pointing away from the binding pocket. Rotation around the bond between the triazole ring and the benzene moiety is possible and allows for different positions of the metal ion (Fig. S11). This will have to be taken into account for PCS-based assignment strategies.

Next, PCS detection was tested with concentrations of **4**/Dy^{3+} of 1.75 mM. Both positive and negative PCS were observed (Fig. 1). It is noteworthy that peaks broadened beyond detection became observable at lower **4**/Dy^{3+} concentrations permitting measurement of PCS (for an example see supporting information, Fig. S12). A simple titration using concentrations of **4**/Dy^{3+} ranging from 0.25 mM to 1.75 mM yielded PCS even for nuclei very close to the paramagnetic center. This is possible because the tag is not covalently attached to the protein. The average metal position in the bound state remains constant throughout the titration. Therefore, PCS and PRE effects should deliver assignments based on geometry.

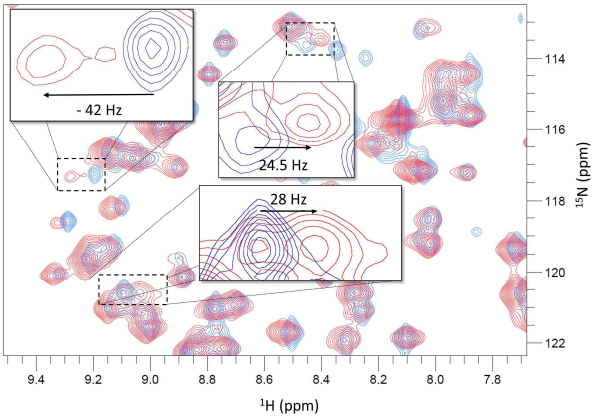


Figure 1: Superposition of $^{1}H,^{15}N$ TROSY HSQC spectra of P-dimers in the presence of **4** (1.75 mM) and a) La^{3+} (blue, diamagnetic reference), or b) Dy^{3+} (red paramagnetic sample). Inserts highlight selected cross peaks and reflect the size of the PCS.

In the presence of 0.1 mM of **4**/Gd^{3+} we observe broadening of numerous cross peaks (Fig. 2) matching chemical shift perturbations as observed before for a complex of P-dimers and methyl α-L-fucopyranoside. These signals likely belong to amino acids located in the binding pocket. In addition, we observe broadening of peaks that were not disturbed upon addition of methyl α-L-fucopyranoside. These peaks likely belong to amino acids of loop structures located above the binding pocket as this is shown in the insert of Fig. 2. This shows how groups of peaks may be assigned to regions in the protein based on simple qualitative arguments.

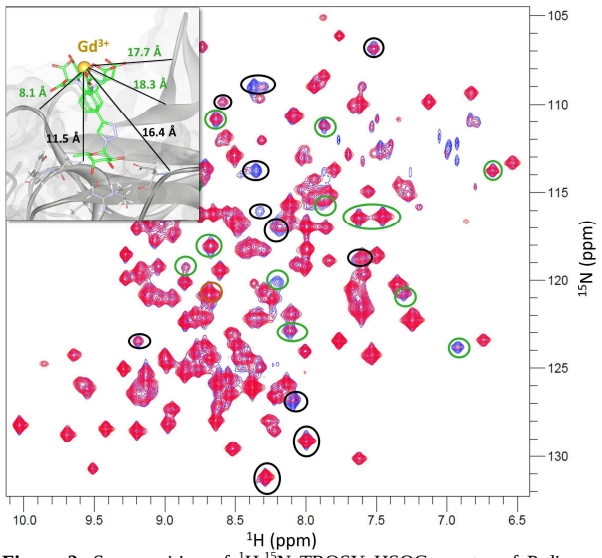


Figure 2: Superposition of $^{1}H,^{15}N$ TROSY HSQC spectra of P-dimers containing La^{3+} as a diamagnetic reference (blue) and Gd^{3+} sample (red) (0.1 mM **4** + metal). Cross peaks in circles show line broadening upon titration with **4**/Gd^{3+}. Cross peaks in black circles match cross peaks that undergo chemical shift perturbation in the presence of methyl α-L-fucopyranoside[10]. Therefore, it is assumed that the corresponding amino acids are located in the fucose binding pocket. For cross peaks in green circles no such corresponding chemical shift perturbations are detected in the presence of methyl α-L-fucopyranoside. This suggests amino acids in adjacent loops that can come close to the paramagnetic center. The insert shows a model of **4**/Gd^{3+} modeled into the fucose binding site, and highlights distances between the paramagnetic ion and closest amino acid residues located either in the fucose binding pocket (black numbers) or belonging to adjacent loop structures (green numbers).

Our current efforts aim at using PRE and PCS data to aid backbone assignment[29, 30]. This requires additional experimental data sets using different paramagnetic ions, and synthesis of ligands with improved binding affinity.

Conclusions

The use of lanthanide tags non-covalently binding to viral coat proteins such as P-dimers offers a strategy that complements classical assignment approaches based on $^{2}H,^{13}C,^{15}N$ labeling and corresponding triple resonance 3D NMR experiments. This approach may be instrumental in obtaining a backbone assignment of this demanding 70 kDa protein target in the future.

Acknowledgements

A.M. is recipient of a Marie Curie fellowship within the 7th Framework Programme of the EU (Grant Agreement No. 329485, NoroCarb). T.P. acknowledges continuing support by the University of Lubeck. Funding of this project by Spanish MINECO (No. CTQ2012-31063/BQU) and FUSP-CEU (PC17/14) is acknowledged.

Notes and references

1. S. M. Ahmed, A. J. Hall, A. E. Robinson, L. Verhoef, P. Premkumar, U. D. Parashar, M. Koopmans and B. A. Lopman, *The Lancet Infectious Diseases*, 2014, **14**, 725-730.

2. L. C. Lindesmith, E. F. Donaldson, A. D. Lobue, J. L. Cannon, D. P. Zheng, J. Vinje and R. S. Baric, *PLoS Med*, 2008, **5**, e31.
3. S. Yazawa, T. Yokobori, G. Ueta, M. Ide, B. Altan, A. Thongprachum, T. Nishimura, T. Nakajima, Y. Kominato, T. Asao, A. R. Saniabadi, K. Furukawa, H. Kuwano, J. Le Pendu and H. Ushijima, *PLoS One*, 2014, **9**, e89071.
4. P. Huang, T. Farkas, S. Marionneau, W. Zhong, N. Ruvoen-Clouet, A. L. Morrow, M. Altaye, L. K. Pickering, D. S. Newburg, J. LePendu and X. Jiang, *J Infect Dis*, 2003, **188**, 19-31.
5. M. K. Jones, M. Watanabe, S. Zhu, C. L. Graves, L. R. Keyes, K. R. Grau, M. B. Gonzalez-Hernandez, N. M. Iovine, C. E. Wobus, J. Vinje, S. A. Tibbetts, S. M. Wallet and S. M. Karst, *Science*, 2014, **346**, 755-759.
6. B. K. Singh, M. M. Leuthold and G. S. Hansman, *J Virol*, 2015, **89**, 2024-2040.
7. B. Fiege, C. Rademacher, J. Cartmell, P. I. Kitov, F. Parra and T. Peters, *Angewandte Chemie*, 2012, **51**, 928-932.
8. C. Rademacher, N. R. Krishna, M. Palcic, F. Parra and T. Peters, *J Am Chem Soc*, 2008, **130**, 3669-3675.
9. M. Tan, R. S. Hegde and X. Jiang, *J Virol*, 2004, **78**, 6233-6242.
10. A. Mallagaray, J. Lockhauserbäumer, G. S. Hansman, C. Uetrecht and T. Peters, *Angew Chem Int Ed*, 2015, **54**, 12014-12019.
11. K. Pervushin, R. Riek, G. Wider and K. Wuthrich, *Proc Natl Acad Sci U S A*, 1997, **94**, 12366-12371.
12. G. Pintacuda, M. John, X.-C. Su and G. Otting, *Acc Chem Res*, 2007, **40**, 206-212.
13. G. Bouvignies, P. R. Markwick and M. Blackledge, *Chemphyschem*, 2007, **8**, 1901-1909.
14. J. R. Tolman and K. Ruan, *Chem Rev*, 2006, **106**, 1720-1736.
15. A. Bax and A. Grishaev, *Curr Opin Struct Biol*, 2005, **15**, 563-570.
16. N. A. Lakomek, T. Carlomagno, S. Becker, C. Griesinger and J. Meiler, *J Biomol NMR*, 2006, **34**, 101-115.
17. T. Zhuang, H. S. Lee, B. Imperiali and J. H. Prestegard, *Protein Sci*, 2008, **17**, 1220-1231.
18. G. Otting, *Annu Rev Biophys*, 2010, **39**, 387-405.
19. A. Canales, A. Mallagaray, J. Pérez-Castells, I. Boos, C. Unverzagt, S. André, H.-J. Gabius, F. J. Cañada and J. Jiménez-Barbero, *Angewandte Chemie International Edition*, 2013, n/a-n/a.
20. A. Mallagaray, A. Canales, G. Dominguez, J. Jimenez-Barbero and J. Perez-Castells, *Chemical communications*, 2011, **47**, 7179-7181.
21. S. Yamamoto, T. Yamaguchi, M. Erdelyi, C. Griesinger and K. Kato, *Chemistry*, 2011, **17**, 9280-9282.
22. M. Erdelyi, E. d'Auvergne, A. Navarro-Vazquez, A. Leonov and C. Griesinger, *Chemistry*, 2011, **17**, 9368-9376.
23. A. Canales, A. Mallagaray, M. A. Berbis, A. Navarro-Vazquez, G. Dominguez, F. J. Canada, S. Andre, H. J. Gabius, J. Perez-Castells and J. Jimenez-Barbero, *J Am Chem Soc*, 2014, **136**, 8011-8017.
24. C. Palomo, J. M. Aizpurua, E. Balentova, I. Azcune, J. I. Santos, J. Jimenez-Barbero, J. Canada and J. I. Miranda, *Organic Letters*, 2008, **10**, 2227-2230.
25. T. Hundertmark, A. F. Littke, S. L. Buchwald and G. C. Fu, *Organic Letters*, 2000, **2**, 1729-1731.
26. M. Mayer and B. Meyer, *Journal of the American Chemical Society*, 2001, **123**, 6108-6117.
27. M. Mayer and B. Meyer, *Angew Chem Int Ed*, 1999, **38**, 1784-1787.
28. M. Mizuno, S. Funahahi, N. Nkasaku and M. Tanaka, *Inorg. Chem.*, 1991, **30**, 1550-1553.
29. S. P. Skinner, M. Moshev, M. A. S. Hass and M. Ubbink, *Journal of Biomolecular NMR*, 2013, **55**, 379-389.
30. G. Pintacuda, M. A. Keniry, T. Huber, A. Y. Park, N. E. Dixon and G. Otting, *Journal of the American Chemical Society*, 2004, **126**, 2963-2970.

[a] *Institute of Chemistry, Center for Structural and Cell Biology in Medicine (CSCM) University of Luebeck, Ratzeburger Allee 160, 23538 Luebeck, Germany.*

[b] *Facultad de Farmacia, Dpto. Química y Bioquímica, Universidad San Pablo CEU, Urb. Montepríncipe, ctra. Boadilla km 5,300 Boadilla del Monte, 28668 Madrid.E-mail: jpercas@ceu.es*

† Electronic

Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/