



## **A post-translational modification of human Norovirus capsid protein attenuates glycan binding**

Mallagaray, Alvaro; Creutzmacher, Robert; Dülfer, Jasmin; Mayer, Philipp H O; Grimm, Lena Lisbeth; Orduña, Jose Maria; Trabjerg, Esben; Stehle, Thilo; Rand, Kasper D; Blaum, Bärbel S; Uetrecht, Charlotte; Peters, Thomas

*Published in:*  
Nature Communications

*DOI:*  
[10.1038/s41467-019-09251-5](https://doi.org/10.1038/s41467-019-09251-5)

*Publication date:*  
2019

*Document version*  
Publisher's PDF, also known as Version of record

*Document license:*  
[CC BY](#)

*Citation for published version (APA):*  
Mallagaray, A., Creutzmacher, R., Dülfer, J., Mayer, P. H. O., Grimm, L. L., Orduña, J. M., ... Peters, T. (2019). A post-translational modification of human Norovirus capsid protein attenuates glycan binding. *Nature Communications*, 10(1), [1320]. <https://doi.org/10.1038/s41467-019-09251-5>

ARTICLE

<https://doi.org/10.1038/s41467-019-09251-5>

OPEN

# A post-translational modification of human Norovirus capsid protein attenuates glycan binding

Alvaro Mallagaray<sup>1</sup>, Robert Creutzmacher<sup>1</sup>, Jasmin Dülfer<sup>2</sup>, Philipp H.O. Mayer<sup>3</sup>, Lena Lisbeth Grimm<sup>1</sup>, Jose Maria Orduña<sup>1,6</sup>, Esben Trøjbjerg<sup>4,7</sup>, Thilo Stehle<sup>3</sup>, Kasper D. Rand<sup>4</sup>, Bärbel S. Blaum<sup>3</sup>, Charlotte Uetrecht<sup>2,5</sup> & Thomas Peters<sup>1</sup>

Attachment of human noroviruses to histo blood group antigens (HBGAs) is essential for infection, but how this binding event promotes the infection of host cells is unknown. Here, we employ protein NMR experiments supported by mass spectrometry and crystallography to study HBGA binding to the P-domain of a prevalent virus strain (GII.4). We report a highly selective transformation of asparagine 373, located in an antigenic loop adjoining the HBGA binding site, into an iso-aspartate residue. This spontaneous post-translational modification (PTM) proceeds with an estimated half-life of a few days at physiological temperatures, independent of the presence of HBGAs but dramatically affecting HBGA recognition. Sequence conservation and the surface-exposed position of this PTM suggest an important role in infection and immune recognition for many norovirus strains.

<sup>1</sup>Institute of Chemistry and Metabolomics, Center of Structural and Cell Biology in Medicine (CSCM), University of Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany. <sup>2</sup>Heinrich Pette Institute, Leibniz Institute for Experimental Virology, Martinistrasse 52, 20251 Hamburg, Germany. <sup>3</sup>Interfaculty Institute of Biochemistry University of Tübingen, University of Tübingen, Hoppe-Seyler-Strasse 4, 72076 Tübingen, Germany. <sup>4</sup>Department of Pharmacy, University of Copenhagen, Universitetsparken 2, Copenhagen 2100, Denmark. <sup>5</sup>European XFEL GmbH, Holzkoppel 4, 22869 Schenefeld, Germany. <sup>6</sup>Present address: Faculty of Pharmacy, Institute of Chemistry, San Pablo CEU University, Urb. Montepríncipe, 28668 Boadilla del Monte, Spain. <sup>7</sup>Present address: Institute of Molecular Systems Biology, ETH Zürich, Zürich, Switzerland. Correspondence and requests for materials should be addressed to T.P. (email: [thomas.peters@uni-luebeck.de](mailto:thomas.peters@uni-luebeck.de))

Infection with human norovirus is the leading cause of acute gastroenteritis worldwide. Attempts to provide antivirals or vaccines have not been successful so far. This is not surprising as noroviruses undergo fast epochal evolution with immune escape variants emerging every 2–5 years<sup>1–4</sup> making it very difficult to design broadly active vaccines or to identify sites on the viral coat protein as targets for entry inhibition. So far, only few studies have focused on the development of such inhibitors<sup>5–8</sup>, and the lack of human norovirus cell culture systems has been a major obstacle in proceeding beyond the stage of hit discovery. During recent years two cell culture systems have been developed<sup>9–11</sup> and, in the future, will allow testing of entry inhibitor candidates in cell culture-based infection assays. Binding of human norovirus to histo blood group antigens (HBGAs), a critical step preceding host-cell entry, has been the focus of a substantial number of biophysical studies involving nuclear magnetic resonance (NMR) experiments, native mass spectrometry (MS), surface plasmon resonance, or crystallography as this has been reviewed recently<sup>12–15</sup>. The highly conserved HBGA-binding sites are located in the P2 subdomain of the P-domain of the norovirus capsid protein VP1<sup>16–18</sup>. Expression of the P-domain in *Escherichia coli* provides homodimeric P-domain species, so called P-dimers<sup>18</sup> that are amenable to HBGA-binding studies. Recently, we observed that HBGA binding to Saga *GII.4* P-dimers yielded discontinuous binding isotherms strongly suggesting the presence of allosteric effects<sup>19</sup>. Initially, we had linked this complex binding behavior to sequential binding of HBGAs to the canonical L-fucose-binding sites and to two additional L-fucose sites that had been identified by crystallography<sup>20</sup> for *GII.10* P-dimers. However, docking experiments<sup>21</sup> and saturation transfer difference (STD) NMR titrations using virus-like particles of *GII.4* and *GII.10* human noroviruses<sup>22</sup> are inconsistent with this interpretation. Therefore, other factors must cause the observed discontinuous binding isotherms.

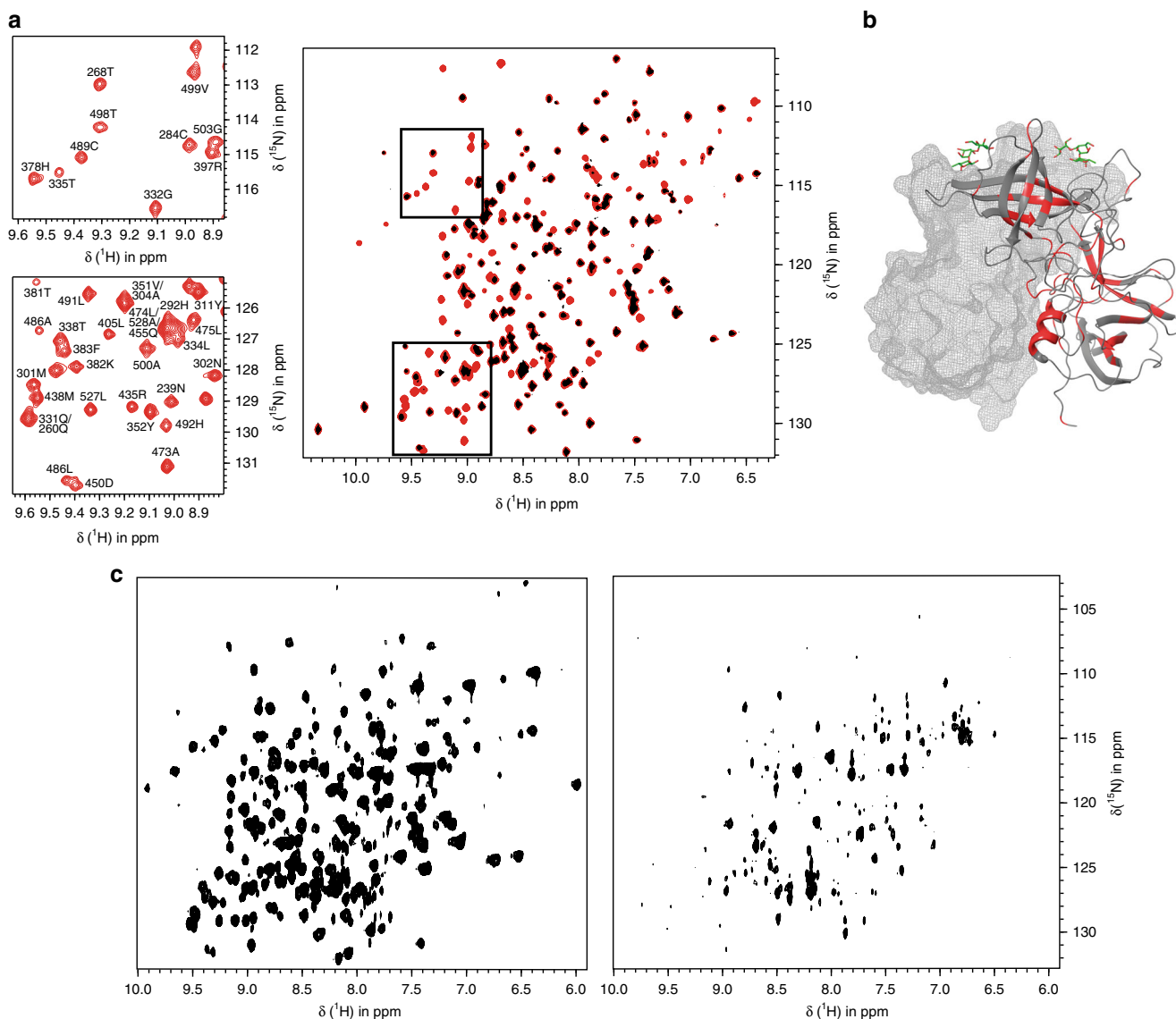
To shine more light on HBGA binding to human noroviruses we have applied chemical shift perturbation (CSP) experiments<sup>23</sup> to Saga *GII.4* P-dimers. Measurement of CSPs of backbone NH resonances upon ligand binding readily identifies amino acids directly involved in ligand interactions. At the same time, CSP experiments can provide long-range effects that may indicate allosteric networks. CSP titration experiments also provide dissociation constants, and can in favorable cases yield access to binding kinetics. However, to take full advantage of CSP experiments two requirements must be met. First, a high-resolution structure of the target protein of interest must be available, and second, an NMR assignment of all or at least most backbone NH resonances must exist. While high-resolution crystal structures of Saga *GII.4* P-dimers were available<sup>17</sup> the assignment of backbone NH resonances for the 73 kDa P-dimer was a substantial challenge<sup>24</sup>. Our previous protein-based CSP study made use of unassigned NH signals undergoing changes upon ligand titration<sup>19</sup> to determine dissociation constants  $K_D$  for L-fucose, the minimal ligand for *GII.4* norovirus. For lack of an assignment this study provided no information about the amino acids involved in binding. Here we present an almost complete assignment of NH backbone resonances of *GII.4* Saga P-dimers. Unexpectedly, the NMR assignment exposed a highly specific deamidation of amino acid N373, resulting in an isopeptide linkage and causing a dramatic decrease of HBGA-binding affinity. These NMR results were supported by crystallography and hydrogen/deuterium exchange MS (HDX MS). HDX MS measures the exchange of protein backbone hydrogens to deuterium in solution. Alterations in the hydrogen bonding network introduced by variable conditions, such as PTMs or ligand interactions, result in altered H/D exchange rates, which makes HDX MS a valuable tool for mapping protein–ligand interaction surfaces and changes in protein

dynamics<sup>25,26</sup>. Sequence analysis of *GII.4* strains suggests that the modification is highly abundant, implying that previously published HBGA binding data may require careful revision. Obviously, such a dramatic change in HBGA binding will have implications for the infection process and, due to the antigenicity of the affected region<sup>27–29</sup>, very likely also for the immune response.

The present study builds on our previous work<sup>19</sup> where we monitored binding of methyl- $\alpha$ -L-fucopyranoside to uniformly <sup>2</sup>H,<sup>15</sup>N-labeled *GII.4* Saga P-dimers. CSP titration experiments based on unassigned <sup>1</sup>H,<sup>15</sup>N Transverse Relaxation Optimized Spectroscopy (TROSY) Heteronuclear Single Quantum Coherence (HSQC) spectra had yielded discontinuous binding isotherms suggesting cooperative binding. Here, we present an almost complete backbone assignment of the P-dimers and uncover spontaneous deamidation of asparagine 373, influencing HBGA binding. In light of this evidence a simple one-site binding model is plausible.

## Results

**Backbone assignment of *GII.4* Saga P-dimers.** At a molecular weight of 73 kDa for the homodimer backbone assignment using triple-resonance three-dimensional (3D) NMR experiments was challenging. It turned out that sample preparation was the most critical factor for acquiring high-quality 3D NMR spectra of [<sup>U</sup>-<sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N] labeled P-dimer samples. Long delays in respective pulse sequences are incompatible with short transverse relaxation times ( $T_2$ ) of backbone NH signals. Therefore, increasing protein concentrations improve signal-to-noise (S/N) ratios in triple-resonance 3D NMR spectra only to the point where protein aggregation and associated slower molecular tumbling become dominant. In the case of Saga P-dimers balancing between high protein concentration for better S/N and sufficiently long  $T_2$  values was rather delicate (Supplementary Table 5 and Supplementary Figs. 7, 8). To better understand P-dimer aggregation and precipitation, we determined molecular correlation times  $\tau_c$  employing <sup>1</sup>H,<sup>15</sup>N-TRACT experiments<sup>30</sup>. With increasing P-dimer concentrations we observed an increase of  $\tau_c$  and concomitantly a decrease of cumulative  $H^N$   $T_2$  values, strongly suggesting formation of aggregates at protein concentrations higher than ca. 200  $\mu$ M (for details cf. Supplementary Note 2, Supplementary Figs. 5–8, and results from Dynamic Light Scattering (DLS) measurements, Supplementary Fig. 9). <sup>1</sup>H,<sup>15</sup>N TROSY HSQC spectra experiments also showed that a substantial fraction (ca. 30%) of NH backbone signals was missing due to very slow exchange of backbone deuterons  $D^N$  to  $H^N$ . Therefore, to achieve a full backbone assignment an unfolding-refolding protocol allowing for complete  $D^N/H^N$  exchange was required (for details cf. Supplementary Note 1, Supplementary Figs. 1–4, and Supplementary Tables 1–4). Following a refolding-screening scheme proposed previously<sup>31</sup>, we established an optimized unfolding-refolding procedure that provides excellent yields of refolded P-dimers. <sup>1</sup>H,<sup>15</sup>N TROSY HSQC spectra of refolded [<sup>U</sup>-<sup>2</sup>H,<sup>15</sup>N]-labeled samples demonstrated complete  $D^N/H^N$  exchange (Fig. 1a, b). Notably, Fig. 1b and Supplementary Fig. 4 show that slowly exchanging  $H^N$  are not only located in  $\beta$ -sheets or  $\alpha$ -helices, where they would be expected. For example, slow  $D^N/H^N$  exchange is also observed for the stretch of amino acids G264 to S279, forming an extended loop region, which is stabilized by a network of intra-loop hydrogen bonds as seen from the crystal structure (pdb: 4X06). Backbone  $H^N$  involved in respective intra-loop hydrogen bonds are invisible or have significantly reduced intensities in non-refolded samples, reflecting protection and likely reduced flexibility of this region of the protein.



**Fig. 1** Exchange of backbone NH protons and protein concentration affect spectral quality. **a**  $^1\text{H}$ ,  $^{15}\text{N}$  TROSY HSQC spectra of  $100\ \mu\text{M}$  [ $U\text{-}^2\text{H}$ ,  $^{15}\text{N}$ ] labeled *GII.4* Saga P-dimers before (black) and after (red) protein refolding recorded on a Bruker AVIII 500 MHz NMR spectrometer. Samples were prepared in buffer containing 75 mM sodium phosphate, 100 mM NaCl, and  $100\ \mu\text{M}$  DSS- $d_6$  in 10%  $\text{D}_2\text{O}$  with the  $\text{pH}^*$  adjusted to 7.30 and with the temperature set at 298 K. **b** Non-exchangeable NH protons (red color) mapped onto the surface of *GII.4* Saga P-dimers (pdb: 4X06). **c** Positive  $^1\text{H}$ - $^{15}\text{N}$  projection of all planes along the  $^{13}\text{C}$  dimension of 3D TROSY HN(CO)CACB spectra acquired under identical conditions (in sodium phosphate 20 mM,  $\text{pH}^*$  7.30) for [ $U\text{-}^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ] *GII.4* Saga P-dimers at concentrations of  $180\ \mu\text{M}$  (left) and of  $300\ \mu\text{M}$  (right). Source data are provided as a Source Data file

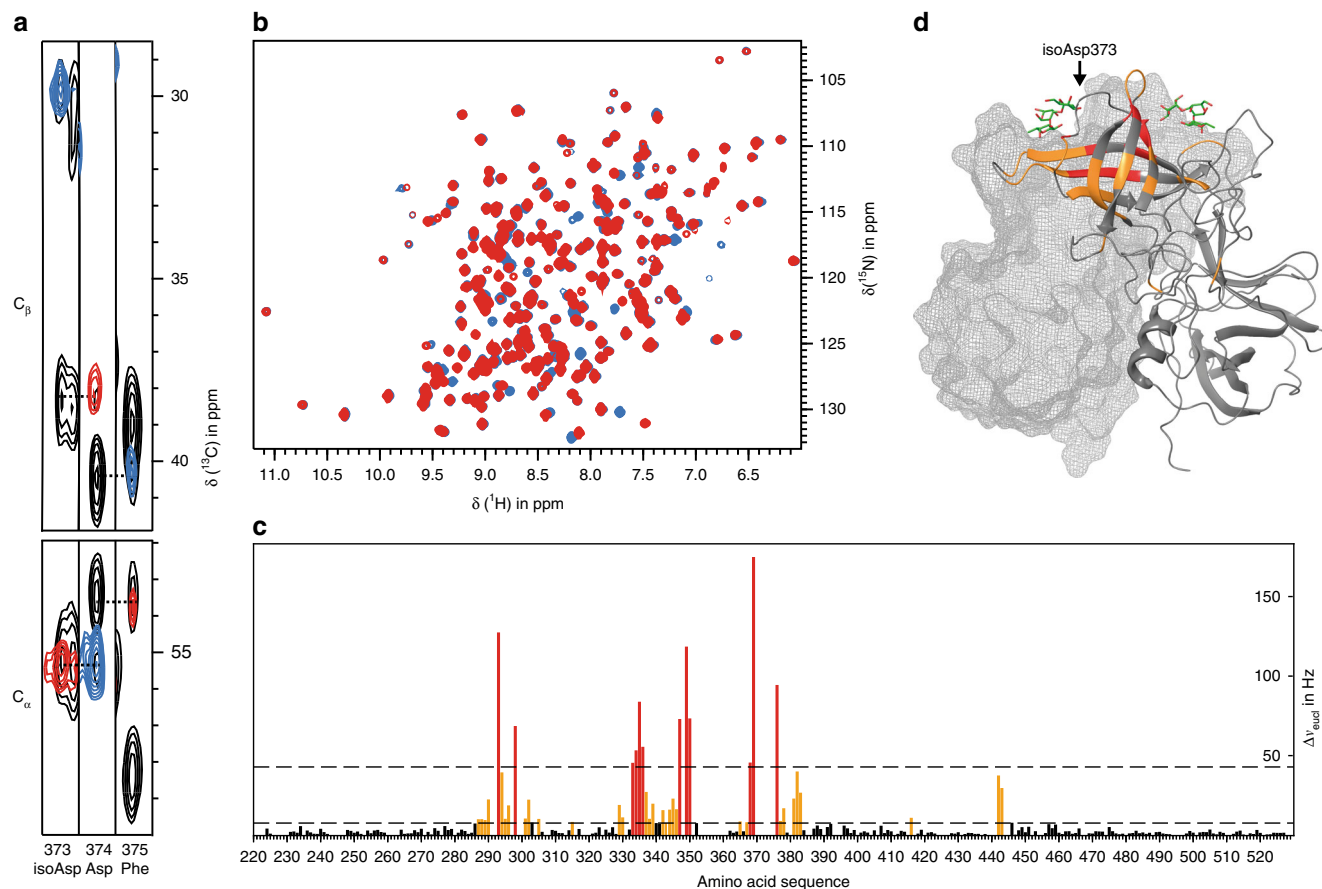
Misadjusting experimental conditions has a dramatic effect on spectral quality. This is shown in Fig. 1c comparing the  $^1\text{H}$ - $^{15}\text{N}$  projection of all planes along the  $^{13}\text{C}$  dimension of a HN(CO)CACB experiment for two concentrations of refolded P-dimers. An optimal concentration of  $180\ \mu\text{M}$  is compared to a slightly increased concentration of  $300\ \mu\text{M}$ .

For the backbone assignment we followed established strategies employing triple-resonance 3D NMR experiments<sup>32</sup>. Using a [ $U\text{-}^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labeled sample of refolded P-dimers we were able to assign 86% of all NH backbone resonances. Except for one case all missing assignments are within surface-exposed loops suggesting local conformational interconversion on the intermediate chemical shift timescale. A BMRB<sup>33</sup> query yielded only three outliers of  $^{13}\text{C}$  chemical shifts of  $\text{C}^\alpha$  and  $\text{C}^\beta$  atoms from the expected distribution, i.e. the deviation from the mean chemical shift was  $>3\sigma$ . In each case ( $\text{C}^\beta$  of Q336,  $\text{C}^\beta$  of A349, and  $\text{C}^\beta$  of P400) the

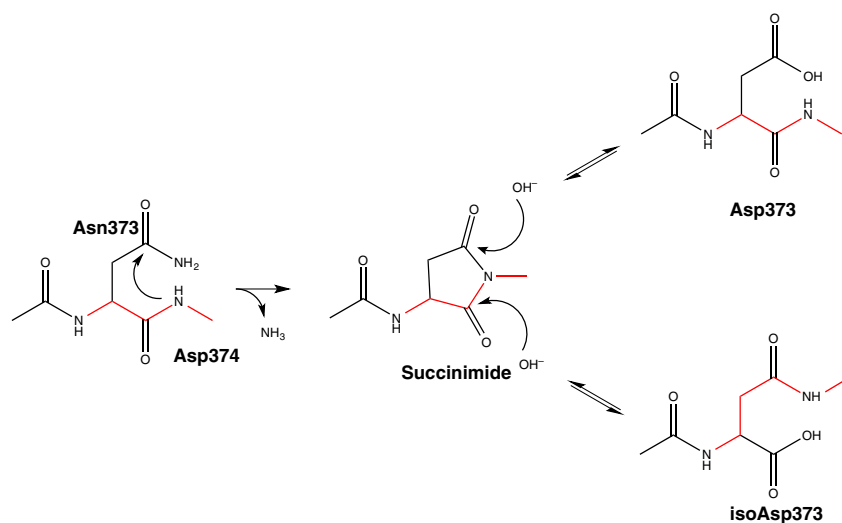
deviation can be explained by the presence of neighboring aromatic side chains. Validation of the chemical shift assignment was accomplished by predicting secondary structure elements exclusively based on the observed chemical shifts using the TALOS-N algorithm<sup>34</sup>. Except for very few positions, likely due to missing assignments or signal overlap, predicted secondary structure elements match very well with those observed in the crystal structure (see Supplementary Fig. 10).

#### NMR reveals an irreversible post-translational modification.

During the assignment process we noticed an irreversible transformation of P-dimers into a second species. Over time, new cross peaks appeared and other peaks disappeared in  $^1\text{H}$ ,  $^{15}\text{N}$  TROSY HSQC spectra. Complete conversion was observed independent of the presence or absence of HBGAs but was significantly affected by the temperature. It is well established that peptides



**Fig. 2** Three-dimensional NMR spectra reveal deamidation of N373 and formation of isoD373. **a** Slices from HNCACB (black) and HN(CO)CACB spectra (red and blue) at the  $^1\text{H}^{\text{N}}$  and  $^{15}\text{N}$  chemical shifts of isoD373, D374, and F375. The cross peak at the  $^{13}\text{C}^{\alpha}$  chemical shift of D374 is negative (blue) showing that the peptide bond is now via the side chain of the deamidated Asn residue in position 373 yielding an iso-Asp (isoD373). **b**  $^1\text{H}$ ,  $^{15}\text{N}$  TROSY HSQC of native Saga P-dimers (red) and deamidated, post-translationally modified Saga P-dimers (blue). **c** Chemical shift perturbations (CSPs) induced by irreversible deamidation of N373 with dashed lines at  $2\sigma$  (red color) and at the experimentally determined significance threshold of 7.9 Hz (orange color, see Supplementary Note 2 and Supplementary Fig. 11 for details). **d** CSPs mapped onto the crystal structure of P-dimers complexed with B trisaccharide (pdb: 4X06). N.B.: for the native P-dimer containing N373, an assignment of the loop comprising amino acids 370–375 is missing. Source data are provided as a Source Data file



**Fig. 3** Deamidation of Asn373 and formation of isoD373. Nucleophilic attack of the side-chain carboxylic acid carbon of Asn373 by the backbone nitrogen of Asp374 leads to formation of a reactive succinimide derivative via a nucleophilic acyl substitution. The unstable succinimide reacts with water molecules to yield either isoAsp373 (isopeptide bond) or Asp373 (normal peptide bond). Red bonds highlight the peptide backbone

and proteins undergo a variety of post-translational modifications, in particular covalent attachments of glycans or phosphate groups to amino-acid side chains. It is less noticed that PTMs can also alter the protein backbone, although very prominent examples exist<sup>35</sup>. A common but often overlooked backbone PTM is the formation of isopeptide bonds via deamidation and subsequent rearrangement of asparagine residues. Following this line of thoughts, we inspected sequential connections in 3D TROSY HNCACB and HN(CO)CACB spectra, which directly report on the formation of isopeptide linkages<sup>36,37</sup> by a phase change of cross peaks sequentially connecting amino acids. Indeed, a tell-tale phase change was observed in converted samples for the cross peaks connecting N373 to D374. The cross peak at the  $^1\text{H}^{\text{N}}$  and  $^1\text{H}^{\text{H}}$  chemical shifts of D374 and the  $^{13}\text{C}^{\alpha}$  chemical shift of the *i*-1 amino acid, N373 in the native form of the P-dimer, has a negative sign in the HN(CO)CACB spectrum (blue color in Fig. 2a) but should be positive for a conventional peptide bond. This phase change of the cross peak is due to formation of an isopeptide bond connecting the backbone NH of D374 to the side chain  $\text{C}^{\gamma}$  of N373 reversing the roles of  $\text{C}^{\alpha}$  and  $\text{C}^{\beta}$  in the HN(CO)CACB experiment. Therefore, the observed phase change is unambiguous evidence for rearrangement of N373 into an isoaspartate (isoD373), as was described previously for malate synthase<sup>37</sup>. Of note, the backbone NH resonances of P-dimers containing isoD373 instead of N373 have been assigned, too. In fact, some assignments were only made for the isoD373 form (cf. Supplementary Table 7).

Conversion of N373 to isoD373 follows the mechanism shown in Fig. 3. Nucleophilic attack of the backbone NH of D374 leads to deamidation of N373, yielding a succinimide as the reactive intermediate. Generally, subsequent nucleophilic acyl substitution with water furnishes a mixture of aspartate and iso-aspartate, with iso-aspartate being the main product. As  $^1\text{H},^{15}\text{N}$  TROSY HSQC spectra of Saga P-dimers do not indicate the presence of D373 isomers formation of an isopeptide linkage in this case seems to be significantly preferred (>90%) over formation of a regular peptide bond. The intermediate succinimide was not detected, which is likely due to its short half-life.

**N373 rapidly converts into isoD373.** For a more quantitative assessment of the conversion of N373 to isoD373 we established an improved purification protocol. We have used ion exchange (IEX) chromatography for separation of isoD373 P-dimers from the native protein since conversion of N373 to isoD373 generates one additional negative charge. The IEX chromatograms exhibit three major peaks, one for the native form with N373 present in both monomeric units, from now on called NN P-dimers. Two additional peaks correspond to the completely converted form, carrying an isoD373 in both monomeric units (iDiD P-dimers), and to the asymmetric form, carrying isoD373 only in one of the monomeric units (iDN P-dimers). Respective IEX chromatograms are shown in Supplementary Fig. 20. A comparison of  $^1\text{H},^{15}\text{N}$  TROSY HSQC spectra of native and deamidated P-dimers (Fig. 2b) shows that the largest CSPs are found in the vicinity of the site of transformation (Fig. 2c, d) but there are also notable long-range CSPs. The regions with large CSPs are in line with deuterium uptake differences seen in the HDX MS data (see results below). It turned out that pure native NN P-dimers are not available from synthesis in bacteria because deamidation and isopeptide bond formation already occur during protein expression. NMR experiments and IEX revealed that freshly expressed samples always contain the isoD373 form in varying amounts, depending on the expression protocol and, in particular, on the temperature used during bacterial protein synthesis.

The IEX protocol was used to monitor conversion of purified native NN P-dimers into iDiD P-dimers as a function of time yielding information on the half-life of NN P-dimers at different temperatures (Supplementary Fig. 20). At room temperature and at temperatures used for sample storage (277 K) deamidation and subsequent formation of the isoD373 is rather slow with a half-life of several days to weeks. However, at body temperature (310 K) conversion is significantly faster with an estimated half-life of 1.6 days (Supplementary Fig. 21).

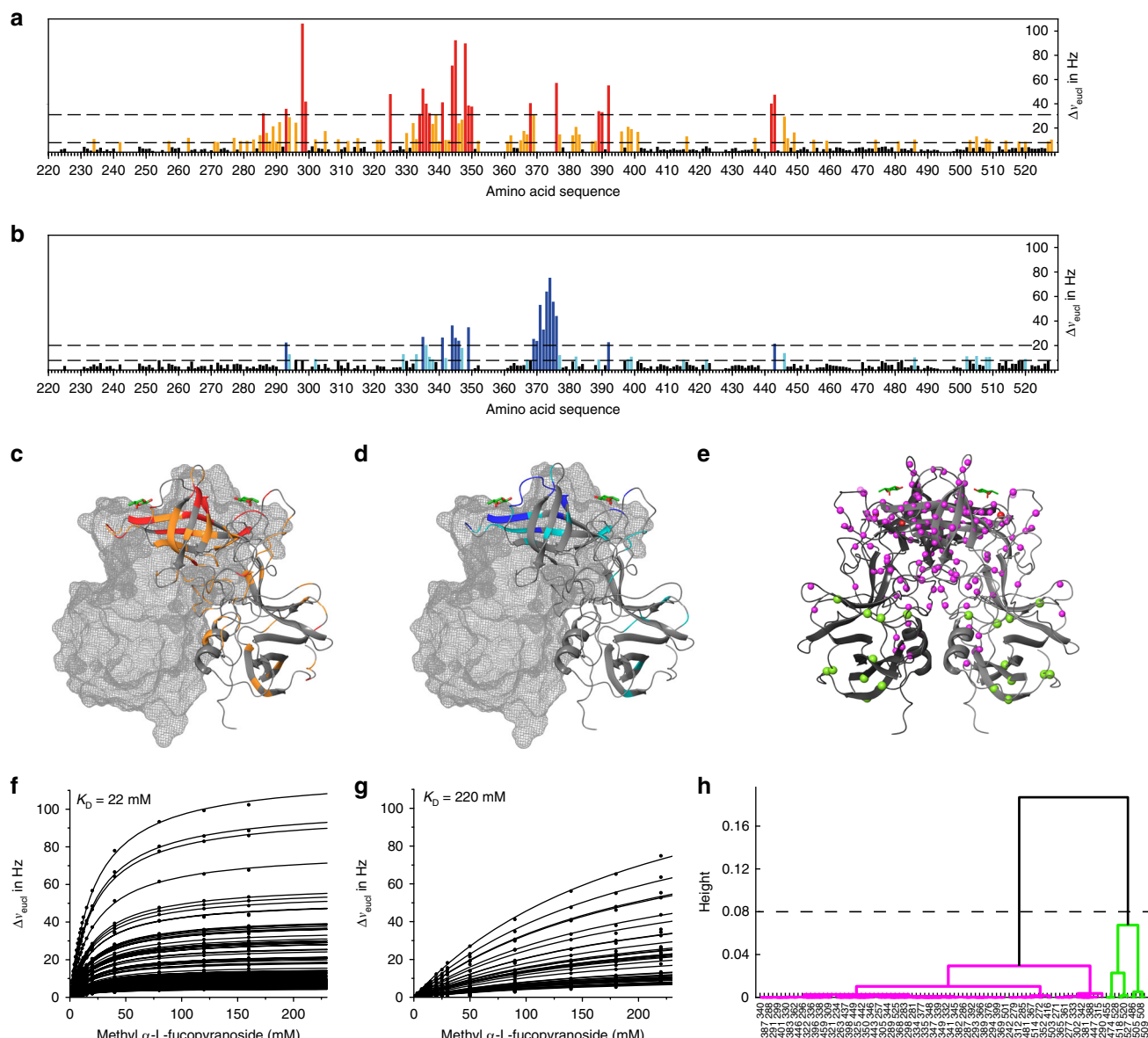
The IEX separated native and deamidated *GII.4* Saga P-dimers were also subjected to peptic cleavage followed by liquid chromatography (LC) separation and peptide identification using tandem MS (MS/MS). Peptides containing the potential deamidation site show peaks with clearly distinguishable retention times in the chromatogram and could be assigned to either the pure native or deamidated form, which differ in mass by +0.98402 Da (Supplementary Fig. 24). Analysis of peptides covering the entire protein sequence revealed no other deamidation sites apart from N373 (Supplementary Fig. 26).

We selected P-dimers of three other norovirus strains, *GII.4* MI001<sup>38</sup>, *GII.10* Vietnam 026<sup>39</sup>, and *GII.17* Kawasaki 308<sup>40</sup> to test whether deamidation of N373 is also relevant for strains other than *GII.4* Saga, and to explore how site-specific this transformation is. *GII.4* MI001 contains an asparagine at position 373, whereas the other two strains lack an asparagine residue at the equivalent position. Samples were incubated at elevated temperatures and IEX chromatograms were obtained before and after incubation as explained in detail in Supplementary Note 4. We observed deamidation for MI001 (Supplementary Fig. 22) but not for the Kawasaki 308 or Vietnam 026 strains.

We also subjected these samples to peptic cleavage followed by LC-MS. For *GII.4* MI001 deamidation of N373 was detected along with minor deamidation of another asparagine residue (either N446 or N448) after incubation at 37 °C (cf. IEX chromatograms in Supplementary Fig. 22 and MS data in Supplementary Table 12 and Supplementary Fig. 25). The other two strains showed no indication of deamidation in the entire protein confirming the results from IEX chromatography.

Our results suggest that asparagine residues preceding the aspartate that is critical for L-fucose recognition (D374 in *GII.4* Saga) will also deamidate (see Supplementary Note 4) if the respective strain displays a similar loop conformation of the amino acids connecting two antiparallel  $\beta$ -sheets underneath the HBGA-binding site (in *GII.4* Saga this is the loop consisting of residues D370-N380, cf. Supplementary Figs. 18 and 19 and Supplementary Table 10).

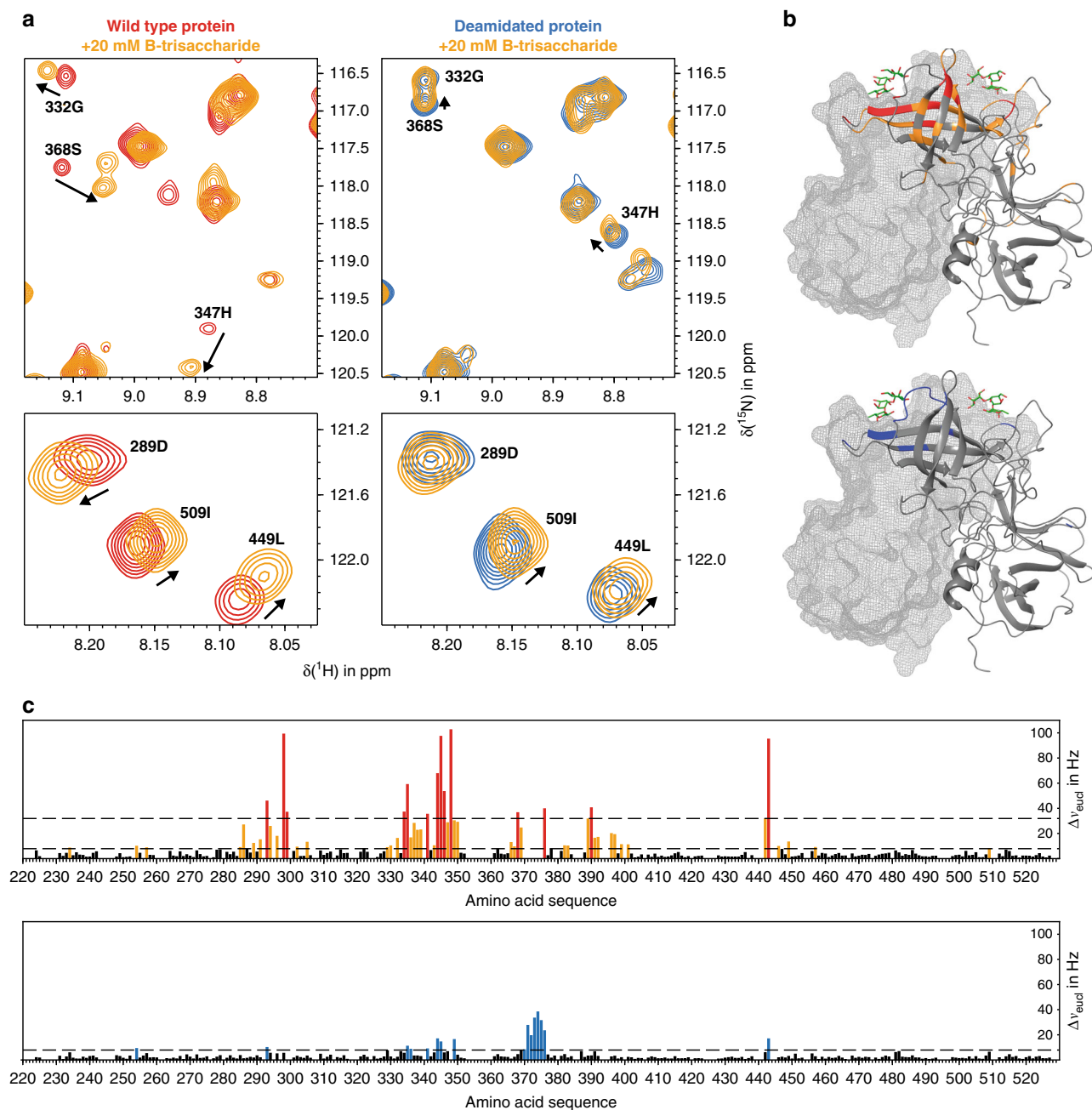
**NMR experiments demonstrate altered glycan binding.** In order to evaluate the influence of the observed PTM on HBGA recognition we performed CSP NMR experiments using methyl  $\alpha$ -L-fucopyranoside and blood group B trisaccharide as representative HBGA ligands. In our previous work<sup>19</sup> we had already studied binding of methyl  $\alpha$ -L-fucopyranoside to *GII.4* Saga P-dimers employing chemical shift titrations. Since at that time the backbone PTM had not been discovered yet the analysis was based on a mixture of NN, iDN, and iDiD P-dimers and could not provide pure dissociation constants for either species. Therefore, we repeated a chemical shift titration using methyl  $\alpha$ -L-fucopyranoside as a minimal HBGA ligand employing NN and iDiD P-dimers. Chemical shift titrations employing freshly prepared NN P-dimers were performed within <4 days to keep conversion at a minimum. iDiD P-dimers were stable, showing no signs of conversion during titration. As many of the backbone NH resonances are highly sensitive to even very small pH changes (unpublished data) we used imidazole as an internal pH



**Fig. 4** Binding of methyl  $\alpha$ -L-fucopyranoside to *GII.4* Saga P-dimers. **a, b** Chemical shift perturbations (CSPs) of NH backbone signals of *GII.4* Saga NN P-dimers (**a**) and iDiD P-dimers (**b**) in the presence of 160 and 220 mM methyl  $\alpha$ -L-fucopyranoside, respectively. Dashed lines are at  $2\sigma$  (red, blue) and at the experimentally determined significance threshold of 7.9 Hz (orange, light blue). **c, d** CSPs for NN (**c**) and iDiD P-dimers (**d**) mapped on the crystal structure (pdb: 4X06) of *GII.4* Saga P-dimers. Color coding as in **a** and **b**. **e** Clusters of amino acids (N backbone atoms are shown as balls) of NN P-dimers reflecting two distinct types of binding isotherms (cf. Supplementary Note 2 and Supplementary Fig. 13). Color coding is as in **h**. **f, g** Global fitting of the law of mass action to chemical shift titration curves of NN P-dimers for amino acids belonging to the magenta cluster (**f**) and for iDiD P-dimers (**g**). For details of the curve fitting see Supplementary Note 2, Supplementary Equation 11, and Supplementary Table 8. The curves reflect one-site binding, and global fitting yields dissociation constants of  $K_D = 22$  mM and of  $K_D = 220$  mM for NN and iDiD P-dimers, respectively. **h** Complete linkage clustering separating amino acids of NN P-dimers based on distinct shapes of binding isotherms into two clusters (magenta and green). Source data are provided as a Source Data file

standard<sup>41</sup>. Data were denoised using singular value decomposition (SVD) as described before<sup>19</sup>, and CSPs were calculated as Euclidean distances. As the resulting curves exhibited visibly different shapes the data were subjected to complete linkage clustering employing the Pearson algorithm (cf. Supplementary Note 2, and Supplementary Figs. 12 and 13 for details), yielding two distinct clusters of binding isotherms for NN P-dimers. Mapping the two clusters onto the surface of the crystal structure (Fig. 4e, h) separates residues in the upper part, which includes the HBGA-binding site, from residues in the bottom part. The binding isotherms of the cluster comprising the binding site residues perfectly match a one-site binding model, and global

fitting of Supplementary Equation 11 yielded a dissociation constant  $K_D$  of 22 mM for binding of methyl  $\alpha$ -L-fucopyranoside to NN P-dimers (Fig. 4f and Supplementary Table 8). This value is considerably higher than dissociation constants reported in previous work<sup>19,42</sup>. Binding isotherms belonging to the other main cluster were precluded from further analysis at this point. For iDiD P-dimers, a cluster analysis was impossible due to much weaker CSPs and to the inability to reach saturation. Although saturation could not be reached, we estimated the dissociation constant by global fitting of iDiD titration curves as 220 mM, demonstrating a dramatic decrease in affinity upon conversion of N373 into isoD373 (Fig. 4g).



**Fig. 5** Binding of blood group B trisaccharide to *GII.4* Saga P-dimers. **a**  $^1\text{H},^{15}\text{N}$ , TROSY HSQC spectra of native (red, left) and deamidated (blue, right) Saga P-dimers in their free forms and in the presence of 20 mM B trisaccharide (orange). Perturbations induced by ligand binding are indicated by arrows for selected amino acids. **b** Mapping of chemical shift perturbations (CSPs) for both forms onto the crystal structure of P-dimers complexed with B trisaccharide (pdb: 4X06). **c** Complete chemical shift mapping of the respective P-dimer species with experimentally determined 7.9 Hz significance thresholds and additional  $2\sigma$  threshold to indicate largest CSPs in native P-dimers. Source data are provided as a Source Data file

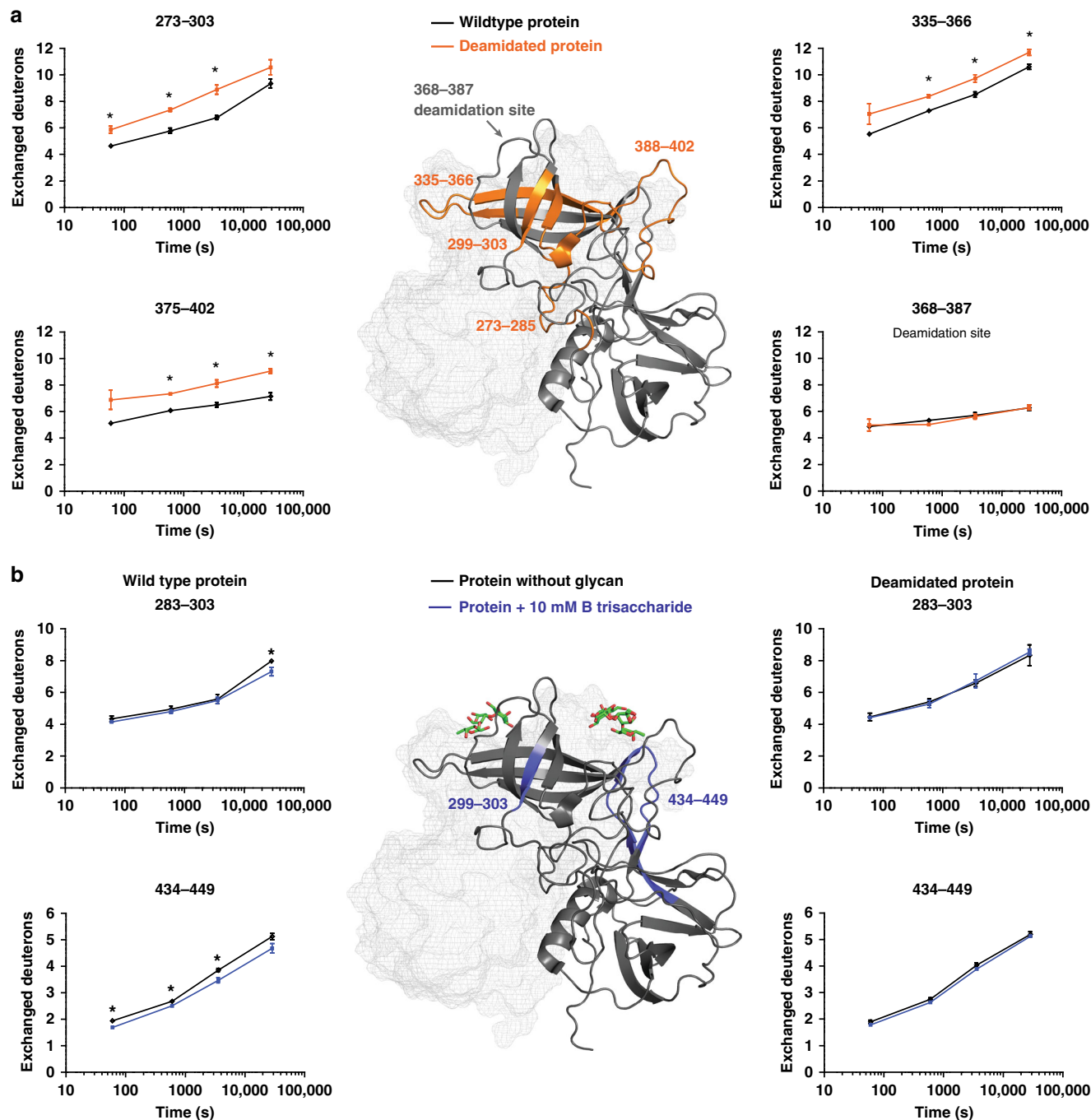
The deamidation reaction also allowed assignment of the N373 side chain NH as this signal disappeared upon conversion into an iso-aspartate. Not surprisingly, this signal showed by far the largest CSP upon addition of methyl  $\alpha$ -L-fucopyranoside, and separate fitting yielded an identical dissociation constant for binding to NN P-dimers (Supplementary Fig. 14).

In order to demonstrate that the loss of HBGA-binding affinity upon transformation of N373 into isoD373 is not confined to methyl  $\alpha$ -L-fucopyranoside we present qualitative CSP data on the binding of blood group B trisaccharide as a representative HBGA ligand to NN and iDiD P-dimers. Blood group B trisaccharide was added at a concentration of 20 mM to NN P-dimers and to

iDiD P-dimers. CSPs were observed for iDiD P-dimers demonstrating B trisaccharide binding. A comparison to CSPs observed for NN P-dimers reveals significant differences (Fig. 5). At the same B trisaccharide concentration CSPs are notably larger for NN P-dimers than for iDiD P-dimers indicating a reduced binding affinity for iDiD P-dimers.

Finally, we reanalyzed the published chemical shift titration dataset<sup>19</sup> for binding of methyl  $\alpha$ -L-fucopyranoside in the light of our findings (for details cf. Supplementary Note 2, and Supplementary Figs. 15 and 16). Since the published binding isotherms reflect binding to a mixture of NN, iDN, and iDiD Saga P-dimers, we had to exclude all cross peaks resulting from



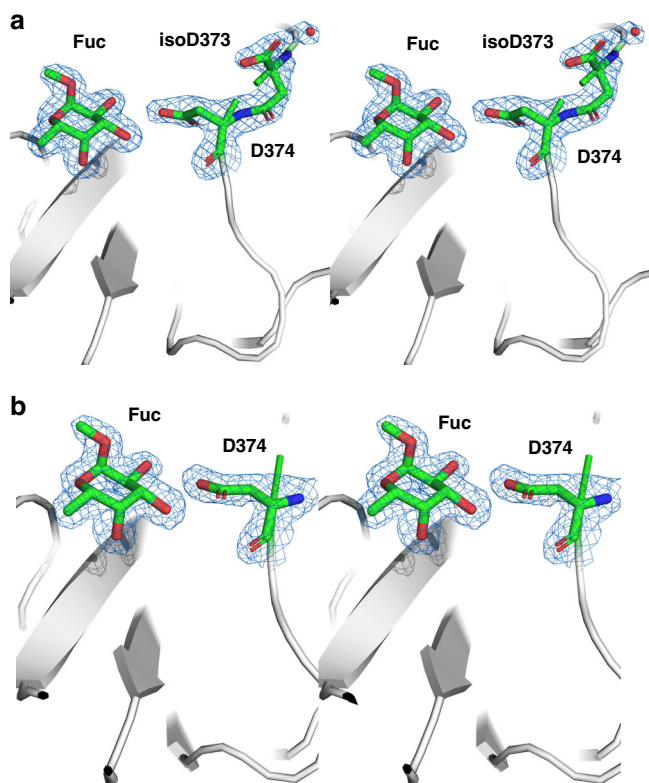


**Fig. 6** Hydrogen/deuterium exchange (HDX) MS experiments reveal changes in protein dynamics upon deamidation. **a** Deuterium uptake differences of the wild-type and deamidated *GII.4* Saga P-dimer mapped to the crystal structure (pdb: 4 × 06). **b** Deuterium uptake differences of the wild-type and deamidated P-dimer in presence and absence of 10 mM B trisaccharide. Significant ( $*p < 0.01$ , Student's *T*-test) protection is detected in peptides covering the canonical binding site (G443, Y444), whereas this effect is absent in the deamidated P-dimer. The reduction in HDX observed in peptide 283–303 can be narrowed down to residues Y299–L303, as a peptide covering residues I283–N298 shows no differences in exchange at the 8 h time point (see Supplementary Data 1–4 for details). All measurements were performed in triplicate, and error bars indicate the standard deviation. A complete collection of deuterium uptake plots is found in Supplementary Data 1–4

deamidated species leaving us with binding isotherms reflecting binding to NN P-dimers. Data were denoised using SVD, and Euclidean distances were calculated. Curves with maximal CSPs larger than  $2\sigma$  (Supplementary Fig. 16) match a single-site binding model well, and are associated with amino acids located in the vicinity of the fucose-binding site. Global fitting of the law of mass action to the titration curves yields a dissociation constant  $K_D$  of 16 mM very similar to the value obtained from the

current titration, with the slightly higher affinity likely due to the use of a different buffer system.

**HDX MS reveals localized alterations in protein flexibility.** To substantiate our findings from NMR experiments, HDX MS was used to map changes in the protein dynamics caused by protein–glycan interactions. Notably, the deamidated P-dimer



**Fig. 7** Binding of methyl  $\alpha$ -L-fucopyranoside to deamidated *GII.4* Saga P-dimers. Wall-eyed stereo representation of the crystal structure showing a non-canonical binding mode of methyl  $\alpha$ -L-fucopyranoside in the modified binding pocket. The  $2F_o - F_c$  map is shown ( $\sigma$  level = 1.0). **a** For one chain the isopeptide bond (isoD373) and the ligand are clearly visible. **b** For the other chain the loop bordering the binding pocket is unstructured N-terminally of D374, and a canonical binding mode is observed. A corresponding ligand interaction diagram and data collection and refinement statistics are found in Supplementary Note 5 (Supplementary Fig. 23 and Supplementary Table 11)

showed a higher H/D exchange rate in peptides covering large parts of the top P2 subdomain irrespective of the presence of glycans. This indicates higher flexibility of respective regions of the P2 subdomain of the deamidated protein (Fig. 6a and Supplementary Data 3). However, this difference was not detected in the peptide containing the deamidation site. In the presence of 10 mM B trisaccharide the wild-type protein was significantly protected from H/D exchange in peptide 434–449 along the canonical binding site (G443, Y444), indicating occupation of the binding pocket. As seen in the NMR experiments binding is highly reduced in the deamidated protein showing essentially no protection (Fig. 6b). In addition, we detect protection of some residues of a  $\beta$ -sheet in the center of the binding cleft (Y286–L303, Fig. 6b). Moreover, this region contains N298, which displays the largest CSP of all assigned backbone NH resonances (cf. Figs. 4a and 5c), supporting a long-range effect of glycan binding into this region of the protein. Apart from overlapping peptides covering the protein regions presented in Fig. 6 no other peptides of either the wild type (Supplementary Data 1) or the deamidated (Supplementary Data 2) P-dimer show significant deuteration differences, indicating that other subtle long-range effects observed by CSPs, which are extremely sensitive to changes in the chemical environment, are beyond the detection limits of the HDX MS experiments performed here. Binding of 100 mM methyl  $\alpha$ -L-fucopyranoside to the wild-type P-dimer results in similar

deuteration differences as observed for 10 mM B trisaccharide suggesting similar interactions. In contrast, no differences can be detected for the negative control, 100 mM D-galactose (Supplementary Data 4).

### Crystallography reveals distinct fucose-binding geometries.

The deamidated form of Saga P-dimers was then subjected to crystal structure analysis in the presence of methyl  $\alpha$ -L-fucopyranoside. Well-diffracting apo crystals were obtained belonging to the space group  $P2_1$ , different from  $C2_1$  as observed previously for the native form of Saga P-dimers<sup>17</sup>. Upon incubation of deamidated P-dimer apo crystals with 600 mM methyl  $\alpha$ -L-fucopyranoside we obtained a co-crystal structure at 1.52 Å resolution, in which the isoD373 isopeptide bond is clearly visible in one of the two protein chains that comprise the asymmetric unit (i.e. in one half of the P-dimer) (Fig. 7 and Supplementary Table 11). We also observed binding of methyl  $\alpha$ -L-fucopyranoside to the two canonical HBGA-binding sites in the Saga P-dimer. However, no binding to additional sites (sites three and four) as observed for *GII.10* P-dimers<sup>20</sup> was detectable, even at this drastic ligand concentration. Whereas the overall P-dimer structure is almost identical to previously published Saga P-dimer structures<sup>16</sup> binding of methyl  $\alpha$ -L-fucopyranoside to the modified binding pocket shows some deviation from the native P-dimer. In the chain with unambiguous density for the isopeptide bond the side chain of D374 is rotated such that one oxygen of the  $\gamma$ -carboxy group forms a single hydrogen bond to OH-4 of methyl  $\alpha$ -L-fucopyranoside (Fig. 7, left). The previously observed binding mode differs somewhat in that a bidentate hydrogen bond involving both oxygen atoms of the  $\gamma$ -carboxy group is formed between the D374 side chain and fucose<sup>16</sup> (corresponding ligand interaction diagrams are shown in Supplementary Fig. 23). This latter orientation is also observed in the second chain in our co-crystal (Fig. 7, right). In this chain, no unambiguous density is observed for residue 373 and the preceding amino acids, and so the density does not allow us to determine whether this chain contains an N373 or an isoD373 residue. Because the crystallization preceded the development of the abovementioned IEX separation protocol, it is possible that the crystallized protein sample contained a certain amount of N373, giving rise to ambiguous density in one of the two chains (where ligand binding is, however, clearly observable).

### Discussion

Our previous studies suggest that binding of human noroviruses to HBGAs involves cooperativity<sup>19,22,43</sup>. Therefore, we set out to better understand underlying molecular principles by studying HBGA P-dimer interactions using protein NMR spectroscopy, HDX MS, and crystallography. We succeeded in obtaining an almost complete NH backbone assignment of the *GII.4* Saga P-dimers using standard 3D triple-resonance NMR experiments, per-deuteration, and an optimized unfolding-refolding protocol for back-exchange of amide protons. This assignment provides an excellent basis for detailed studies into binding of HBGAs as well as other ligands under near-physiological conditions. In the assignment process we discovered that N373, located in a loop directly adjoining the HBGA-binding site, transforms into an iso-aspartate, isoD373, within 1–2 days at body temperature. Experimental proof comes from NMR reflecting the altered connectivity between D374 and N373 via an isopeptide bond by swapping the phases of  $^{13}\text{C}^\alpha$  and  $^{13}\text{C}^\beta$  cross peaks of D374 in 3D HN(CO)CACB spectra. This phenomenon was previously described for malate synthase<sup>37</sup> and was applied more recently for the identification of an isopeptide bond connecting an Asn side chain and a Lys side chain of a bacterial surface protein<sup>36</sup>. MS

confirms the mass shift of +0.98402 Da in peptides containing N373/isoD373 supporting the NMR findings.

Since residue 373 is located in a loop critically involved in fucose recognition it is maybe not surprising that the transformation of N373 into an iso-Asp affects HBGA recognition as demonstrated by CSP NMR experiments using methyl  $\alpha$ -L-fucopyranoside and blood group B trisaccharide as exemplary ligands (Figs. 4 and 5). In fact, loss of binding affinity due to iso-Asp formation has been described for monoclonal antibodies<sup>44</sup>. Chemical shift titrations show for the case of methyl  $\alpha$ -L-fucopyranoside a dramatic loss of affinity of about one order of magnitude upon formation of the isopeptide linkage (Fig. 4). This observation is qualitatively supported by HDX MS experiments showing significantly higher protection from H/D exchange for NN P-dimers as compared to iDiD P-dimers (Fig. 6b). Crystal structure analysis provides an explanation for the observed loss of binding affinity. In our crystal structure of the deamidated P-dimer complexed with methyl  $\alpha$ -L-fucopyranoside two binding geometries can be observed at the same time, i.e. one in each crystallographically distinct half of the P-dimer. Fucose binds to the modified binding pocket in much the same place and orientation as observed for the native protein, but the hydrogen bonding to the side chain carboxy group of D374 appears altered as shown in Fig. 7 and Supplementary Fig. 23. The altered D374 side-chain orientation and the loss of one hydrogen bond can be explained by altered conformational dynamics of the loop. In addition, HDX MS reveals alterations of the P2 subdomain flexibility upon formation of the isopeptide linkage. Together, the altered dynamics and consequently less favorable rotamer distribution of the D374 side chain may explain the loss of affinity of the modified binding pocket that is detectable by CSP NMR experiments (Fig. 5). The hypothesis that the isopeptide formation alters the conformational dynamics of the respective loop is based on the observation that NH cross peaks from residues 370–375 of this loop are missing in the native P-dimer TROSY HSQC spectra but are visible for the deamidated form (Supplementary Table 7). Except for the loop carrying the isopeptide bond, the overall structure of the P-dimers is very similar to the structures published before<sup>17</sup>. However, numerous long-range CSPs and deuterium uptake differences measured by HDX MS are observed upon deamidation (Figs. 2d and 6a, respectively). This implies that effects of deamidation are not confined to the vicinity of the deamidation site suggesting differences in overall protein dynamics.

Chemical shift titration analysis shows that binding of methyl  $\alpha$ -L-fucopyranoside to NN as well as iDiD P-dimers can be entirely described using a single-site binding model (Fig. 4f, g). Binding isotherms resulting from CSPs of residues belonging to a minor cluster (Fig. 4e–h) are likely secondary effects of glycan binding. Therefore, we reanalyzed our old chemical shift titration dataset<sup>19</sup> to test whether it may match these findings. In fact, global fitting of binding isotherms originating from amino-acid residues in the vicinity of the HBGA-binding site can be analyzed using a single-site binding model, yielding a similar dissociation constant for the binding of methyl  $\alpha$ -L-fucopyranoside to NN P-dimers (cf. Supplementary Note 2 and Supplementary Fig. 16). The discontinuities observed in our previous study can now be assigned to amino acids located at remote positions relative to the fucose-binding site, exhibiting weaker CSPs and thus are more sensitive to small variations of experimental conditions. In the light of our data the discontinuities have to be ascribed to variations in sample conditions due to ongoing deamidation (Fig. 3) and the use of a protein concentration of 300  $\mu$ M, well above the critical limit for aggregation determined in the present study (cf. Fig. 1c). It should be noted that this does not explain the steps observed in STD NMR titration curves<sup>19,22</sup>, which

apparently have a different origin. Further work is required to clarify this issue.

Interestingly, recalculation of the dissociation constant  $K_D$  for blood group B trisaccharide based on data from a competitive STD NMR titration<sup>19</sup> and using a  $K_D$  value of 22 mM instead of 2.4 mM as a reference for methyl  $\alpha$ -L-fucopyranoside yields a  $K_D$  value of 5.5 mM for blood group B trisaccharide. In the light of our findings we must conclude that our original interpretation and the resulting working hypothesis were incorrect and that HBGA binding is not a cooperative process. However, binding of HBGAs causes many long-range CSPs across distances larger than 20–30 Å. Although the biological implications of these long-range effects are not known yet their presence indicates a rather subtle cross-talk between the HBGA-binding site and other sites in the P-dimer.

What could be the biological significance of irreversible deamidation of N373? Deamidation and isomerization of asparagine residues as well as isomerization of aspartate residues (Fig. 3) are spontaneous processes generating PTMs associated e.g. with protein aging. Therefore, these modifications serve as molecular clocks controlling the life time of proteins<sup>45,46</sup>. On the other hand, a cellular enzymatic repair system exists, reconverting iso-aspartate into aspartate and restoring the conventional protein backbone<sup>47–49</sup> through the action of protein-L-iso-aspartyl O-methyltransferase. Identification and characterization of these PTMs, which are linked to a single unit change in molecular mass, is a challenging analytical problem<sup>50–54</sup>. Therefore, the number of studies describing such modifications and elucidating their biological function is sparse. Nevertheless, there is growing evidence that deamidation of asparagine and isomerization of aspartate are in fact critically involved in a variety of biological processes. For instance, such PTMs may play a regulatory role in apoptosis<sup>55,56</sup> and confer thermal stability in enzymes present in hyperthermophilic bacteria<sup>57</sup>. One study suggested that iso-Asp formation in integrin ligands leads to an isoDGR motif mimicking the canonical RGD integrin-binding motif<sup>58</sup>. In general, there are a number of PTMs modifying the protein backbone and thus modulating protein functions<sup>35</sup>. In the Saga *GII.4* P-dimer the site of iso-aspartate formation is in a loop that engages in HBGA binding, and sequence analysis shows that in *GII.4* strains 66% of all P-domains carry an asparagine residue in position 373. In fact, almost all *GII.4* strains, for which crystal structures of P-dimers have been obtained carry this motif (for details cf. Supplementary Note 3, Supplementary Fig. 17, and Supplementary Table 9). Given a half-life of N373 of the order of only 1–2 days this PTM may play an important role during infection and may also modulate recognition by monoclonal antibodies binding to this region of the P-domain<sup>27–29</sup>. Importantly, vaccines based on subviral complexes presenting P2 subdomains on the surface should be designed such that no deamidation can occur<sup>59</sup>. Correspondingly, screening for antivirals based on inhibition of HBGA binding and subsequent optimization of binding affinities to generate lead compounds will require employing the non-deamidated form of the protein in respective binding assays. Finally, it will be important to clarify whether deamidation adversely affects the infectivity of noroviruses in cell culture systems and thus reduces virus titers.

Which parameters determine the high regioselectivity of the observed PTM? For peptides it is known that the  $i + 1$  residue is the main factor controlling the deamidation rate. According to these studies Asn residues followed by a Gly have the shortest half-life<sup>45,49</sup>. It has been recognized early on that for proteins other factors are dominant<sup>60,61</sup>. In particular, the conformation of the sequence containing the deamidation site plays a crucial role. A recent study has addressed this problem by computing free energy changes upon deamidation using a hybrid QM/MM

approach<sup>62</sup>. One important finding is that the deamidation motif requires a defined local conformation of the Asn deamidation site  $i$  and the  $i + 1$  residue that positions the backbone NH of the  $i + 1$  residue at a short distance to the carboxylic acid amide carbon of the side chain of Asn to facilitate nucleophilic attack. Two reactive conformations were identified corresponding to specific combinations of the backbone dihedral angle  $\psi$  ( $N_i-C_i^\alpha-C_i'-N_{i+1}$ ) and the side-chain dihedral angle  $\chi$  ( $C_i'-C_i^\alpha-C_i^\beta-C_i^\gamma$ ) of  $\psi/\chi$  ca.  $180^\circ/60^\circ$  and of  $\psi/\chi$  ca.  $0^\circ/-60^\circ$ . Consequently, steric hindrance toward accessing such reactive conformations and solvation of the corresponding transition state are the two single most important factors influencing the rate of deamidation<sup>62</sup>. Analysis of the crystal structures of *GII.4* Saga P-dimers (see Supplementary Table 10) shows that N373 in both monomers matches the  $\psi/\chi$  ca.  $0^\circ/-60^\circ$  conformation. This reactive conformation is also found in known crystal structures of other *GII.4* P-dimers carrying the N373-D374 motif (Supplementary Fig. 18 and Supplementary Table 10). However, another Asn residue in Saga P-dimers, N512, also matches this conformation closely and is followed by a Gly residue, suggesting this site being at high risk for deamidation. In fact, no deamidation is observed for N512 even at prolonged storage at room temperature underscoring that prediction of deamidation sites in proteins based on simple sequence motifs and on local backbone and side-chain conformations is still error prone. This observation may also indicate that the side-chain conformational equilibrium of N512 is altered in solution as compared to the crystal.

Superposition of the loops containing the N373-D374 motif in published crystal structures shows that the conformation of this loop is highly conserved (Supplementary Figs. 18 and 19, and Supplementary Table 10) and, therefore, maybe critical for the formation of a special local environment promoting deamidation. We hypothesize that incorporation of the N373-D374 motif into this loop is necessary and sufficient to induce site-specific fast deamidation. This hypothesis is supported by the observation of deamidation for *GII.4* MI001 but not for *GII.10* Vietnam 026 and *GII.17* Kawasaki 308. The latter two strains display the proper loop conformation but lack an asparagine in position 373. For MI001 the critical loop has the sequence STDTSND<sup>374</sup>, which is almost identical to the one found in Saga P-dimers (STDTE<sup>374</sup>). Although there is no crystal structure available for MI001, this suggests that the conformation of the critical loop is identical to the conformation found for *GII.4* Saga.

By revealing a hitherto overlooked backbone PTM in the ligand-binding site of norovirus *GII.4* P-dimers this study offers insights into HBGA binding to noroviruses. Our present work demonstrates that binding of methyl  $\alpha$ -L-fucopyranoside to P-dimers follows a simple single-site binding model with much higher dissociation constants than reported previously. Formerly observed discontinuities of binding isotherms must be linked to progressive deamidation and protein aggregation. However, binding of methyl  $\alpha$ -L-fucopyranoside and of B trisaccharide to P-dimers cause long-range CSPs of amino-acid residues more than 30 Å away from the binding site, indicating the presence of subtle allosteric cross-talk. Future studies are needed to analyze these long-range effects and to explore their biological implications.

An important question arising from our studies is how general isopeptide formation is among different norovirus strains. Simple sequence analysis shows that in *GII.4* strains there is a high degree of conservation of N373 of ca. 66%, whereas D374, which is essential for fucose recognition, is 100% conserved. In addition, available crystal structures suggest that the conformation of the loop containing the N373-D374 is highly conserved. Therefore, it is likely that all strains carrying the N373-D374 motif are undergoing the described backbone PTM, which would concern a large fraction of all strains including the predominant ones.

Employing IEX chromatography and MS, it should be possible to systematically test the prevalence of this PTM among different strains. The availability of cell culture systems will then allow infection studies with corresponding strains under controlled conditions and comparative binding to known monoclonal antibodies will inform on the effects of this PTM on immune recognition and immune evasion.

## Methods

**Expression and purification of P-dimers.** *GII.4* Saga 2006 (residues 225–530), *GII.4* MI001 (residues 225–530), *GII.10* Vietnam 026 (residues 224–538), and *GII.17* Kawasaki 308 2015 (residues 225–530) P-domains, with GenBank accession numbers AB447457, KC631814, AF504671, and LC037415, respectively, were synthesized and purified using the following protocol. *E. coli* BL21(DE3) were transformed with a pMal-c2x expression vector encoding the genes for ampicillin resistance, a fusion protein of maltose-binding protein (MBP), two His-tags, a HRV3C cleavage domain, and the P-domain. Due to the cloning strategy, the sequences from *GII.4* Saga 2006 and *GII.17* Kawasaki 308 2015 P-domains contain an extra GPGS sequence preceding K225, whereas *GII.10* Vietnam 026 contains a GPG sequence preceding S224. For synthesis of unlabeled proteins, transformed cells were grown for 3 h at 37 °C in modified terrific broth medium (12 g tryptone, 24 g yeast extract, and 40 ml glycerol per liter culture) supplemented with M9 minimal medium components (0.5 g of NaCl, 3.3 g of  $KH_2PO_4$ , 16.6 g of  $Na_2HPO_4 \cdot 12 H_2O$ , 1 g of  $NH_4Cl$ , 1 ml of 1 M  $MgSO_4$ , 1 ml of 0.1 M  $CaCl_2$ , and 0.2 % glucose per liter culture), 0.4 % casamino acids, and 100  $\mu g ml^{-1}$  ampicillin. Overexpression was induced with 1 mM isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG) at an  $OD_{600}$  value of 1.5. Incubation was continued at 16 °C for 48 h. [ $U$ - $^2H$ , $^{15}N$ ]- and [ $U$ - $^2H$ , $^{13}C$ , $^{15}N$ ]-labeled *GII.4* Saga 2006 P-dimers were synthesized following a modified protocol. Cells were grown in 40 ml of terrific broth medium at 37 °C in an orbital shaker operating at 200 rpm for 16 h. Cells were harvested and resuspended in 40 ml of  $D_2O$ -based M9+ minimal medium with a starting  $OD_{600}$  of 0.1. M9+ medium contained 3  $g l^{-1}$   $^{15}NH_4Cl$  (Deutero), 3  $g l^{-1}$  deuterated glucose (1,2,3,4,5,6,6- $d_7$ , Deutero) for [ $U$ - $^2H$ , $^{15}N$ ] labeling, or deuterated  $^{13}C_6$ -glucose (Deutero) for [ $U$ - $^2H$ , $^{13}C$ , $^{15}N$ ] labeling as sole nitrogen and carbon sources, respectively. Culture media further contained 13 g  $Na_2HPO_4 \cdot 2 H_2O$ , 3.6 g  $KH_2PO_4$ , 1 g NaCl, 465 mg  $MgSO_4$ , 200 mg  $MgCl_2$ , and 14.2 mg  $CaCl_2$  per l, 100  $\mu g ml^{-1}$  ampicillin, and 24 % *E. coli* OD2 DN medium (Silantes), and were supplemented with vitamins (20 mg vitamin B1, 0.1 mg riboflavin, and 1 mg D-biotin, choline chloride, folic acid, nicotinamide, D-pantothenic acid, pyridoxal hydrochloride, and cobalamine, respectively). The culture was incubated at 37 °C until an  $OD_{600}$  of 0.5 was reached. Cells were again harvested and resuspended in the final culture volume of fresh M9+ medium. Temperature was lowered to 16 °C at an  $OD_{600}$  of 0.8 and expression was induced with 1 mM IPTG. Cells were harvested after an  $OD_{600}$  of 5 was reached, resuspended in phosphate-buffered saline buffer, and then lysed using a high-pressure homogenizer (Thermo). The lysate was clarified by centrifugation, and the fusion protein was purified using a Ni-NTA resin (Qiagen). MBP and the His-tag were cleaved from the P-domain using HRV 3C protease (Novagen). Cleaved P-domain protein eluted from Ni-NTA resin and was further purified by size-exclusion chromatography using a Superdex 26/600 75  $\mu g$  column (GE Healthcare) in 20 mM sodium phosphate buffer (pH 7.3). Protein purity and dimer concentration were monitored by SDS-polyacrylamide electrophoresis and ultraviolet absorption ( $\epsilon_{280}$  70,820  $M^{-1} cm^{-1}$ ), respectively. Separation of fully, partially, and non-deamidated P-dimer species was achieved by cation exchange chromatography using a 6 ml Resource S column (GE Healthcare) at 6 °C. Protein samples were prepared in 20 mM sodium acetate buffer (pH 4.9) and eluted using a 78.4 ml linear salt gradient up to 195 mM NaCl with a flow rate of 1  $ml min^{-1}$ . The deamidation reaction rate at 310 K of *GII.4* Saga P-dimers was determined by incubation of 125  $\mu l$  P-dimer aliquots with a protein concentration of 2.3  $mg ml^{-1}$  in 75 mM sodium phosphate buffer, 100 mM NaCl (pH 7.3) for up to 48 h and subsequent analytical IEX chromatography. Baseline correction and quantification of peak integrals of the native P-dimer species was performed using Unicorn 7 software (GE Healthcare). *GII.4* MI001 and *GII.10* Vietnam 026 P-dimers were incubated in 25 mM TRIS, 300 mM NaCl (pH 7.3) at 25 °C for 3 weeks. *GII.17* Kawasaki 308 P-dimers were incubated for 48 h at 37 °C in 25 mM Tris and 500 mM NaCl (pH 7.3). All samples were tested for deamidation by IEX chromatography and MS.

**Protein unfolding and refolding.** Conditions for refolding of P-dimers were screened (cf. Supplementary Note 1), and a protocol for  $D^N/H^N$  exchange was developed. Every step was performed at 6 °C, and materials and solutions used in the process were pre-cooled prior to use. Protein samples were prepared in 20 mM sodium phosphate buffer (pH 7.3) at a protein concentration of 7  $mg ml^{-1}$ . One milliliter of the protein solution was transferred into 35 ml of unfolding buffer (4 M guanidine hydrochloride, 0.5 M Tris, 0.3 M NaCl, and 10 mM 2-mercaptoethanol (BME), pH 7.3) and incubated for 4 h. Then, refolding was started by addition of 35 ml of stabilization buffer (3 M guanidine hydrochloride, 0.5 M Tris, 1 M L-proline, 0.4 M D-sucrose, and 10 mM BME, pH 7.3) to the reaction mixture and subsequent dialysis against 2 l of refolding buffer (0.5 M Tris, 0.5 M L-proline,

0.2 M D-sucrose, and 10 mM BME, pH 7.3) overnight. The protein solution was then dialyzed stepwise against 2 l of diluted refolding buffer (0.25 M Tris, 0.25 M L-proline, 0.1 M D-sucrose, and 5 mM BME, pH 7.3) for 8 h, 2 l of gel-filtration buffer (50 mM Tris, 0.45 M NaCl, and 5 mM BME, pH 7.3) once for 8 h and twice for 24 h. Note, that the volume of the protein mixture doubled during dialysis due to osmotic pressure. Refolded P-dimers were again purified by size-exclusion chromatography with an overall refolding yield of 80%.

**NMR spectroscopy.** All NMR experiments were recorded at 298 K. Spectra were processed using TopSpin 3.5 (Bruker) and analyzed using CCPNMR 2.4.2<sup>63</sup>. NMR spectra for protein backbone assignment were acquired on 900 and 500 MHz spectrometers (Bruker) equipped with a TCI cryogenic probe using TROSY versions of standard 3D pulse programs (for experimental details cf. Supplementary Table 6). Final assignment spectra were acquired with a 180  $\mu$ M [ $U$ - $^2$ H, $^{13}$ C, $^{15}$ N] protein sample in 20 mM sodium phosphate buffer, 10 % D<sub>2</sub>O, and 200  $\mu$ M DSS- $d_6$  (pH\* 7.3). Carbon chemical shift outliers were identified by comparison with the BRMB amino-acid chemical shift database (queried in March 2018) and assignment data were deposited under the accession code 27445. Secondary structure prediction was performed using the TALOS-N webserver<sup>34</sup> with chemical shift corrections for deuterium isotope effects. Predicted structural elements shorter than three amino acids have not been considered for comparison with crystal structure data.

Optimization of sample conditions and ligand titrations were performed on a Bruker AVIII 500 MHz spectrometer equipped with a TCI cryogenic probe. Rotational correlation times, apparent molecular weights, and cumulative  $T_2$  relaxation times were estimated employing  $^1$ H, $^{15}$ N-TRACT and spin-echo experiments under a variety of sample conditions (for details cf. Supplementary Figs. 6–8).  $^1$ H, $^{15}$ N TROSY HSQC experiments were acquired with samples in 75 mM sodium phosphate buffer (pH\* 7.3), 100 mM NaCl, 100  $\mu$ M DSS- $d_6$ , 8–10 % D<sub>2</sub>O, and 300–400  $\mu$ M imidazole and [ $U$ - $^2$ H, $^{15}$ N] P-dimer concentrations between 85 and 150  $\mu$ M unless stated otherwise.

**Ligands used for CSP experiments.** Methyl  $\alpha$ -L-fucopyranoside was purchased from Carbosynth. Blood group B trisaccharide  $\alpha$ -L-Fuc-(1,2)-[ $\alpha$ -D-Gal-(1,3)-] $\alpha$ -D-Gal-(1,N)-N<sub>3</sub> was a gift from Dr. Hanne Peters in our laboratory and had been obtained via enzymatic synthesis from  $\alpha$ -L-Fuc-(1,2)- $\alpha$ -D-Gal-(1,N)-N<sub>3</sub>, which was a gift from Prof. Javier Pérez Castells (CEU San Pablo, Madrid). To reproduce the CSP experiments described here, commercially available blood group B trisaccharide may equally well be used as a ligand.

**HDX MS and peptide identification.** Wild-type and deamidated P-dimer (50 pmol) were mixed with glycans at 10-fold of the final concentration and directly diluted 1:9 in 99% deuterated 20 mM Tris buffer (pH 7.4, 150 mM NaCl, 25 °C) to start the exchange reaction. After various time points the exchange reaction was quenched by 1:1 addition of ice-cold quench buffer (300 mM phosphate buffer, pH 2.3, 6 M urea), which decreased the pH to 2.3, and frozen in liquid nitrogen. The wild-type and deamidated datasets with B trisaccharide were acquired on two consecutive days and all time points were performed in triplicate. The wild-type dataset with B trisaccharide, methyl  $\alpha$ -L-fucopyranoside, and galactose contains single measurements performed on the same day.

The samples were thawed and injected onto a cooled (0 °C) HPLC System (Agilent Infinity 1260, Agilent Technologies) equipped with a home packed pepsin column (IDEX guard column with an internal volume of 60  $\mu$ L, Porozyme Immobilized Pepsin beads, Thermo Scientific) in a column oven (25 °C), a peptide trap column (OPTI-TRAP for peptides, Optimize Technologies), and a reversed-phase analytical column (PLRP-S for Biomolecules, Agilent Technologies). Pepsin digestion was performed online at a flow rate of 75  $\mu$ l min<sup>-1</sup> (0.23% formic acid in water) and peptides were trapped in the trap column. Peptides were eluted and separated on the analytical column using a 7 min gradient of 8–40% solvent B (solvent A: 0.23% formic acid in water, solvent B: 0.23% formic acid in acetonitrile) at 150  $\mu$ l min<sup>-1</sup>. MS was performed using an Orbitrap Fusion Tribrid in positive electron spray ionization MS only mode (Orbitrap resolution 120 K, 4 microscans).

Peptide and PTM identification was performed on non-deuterated samples using a longer elution gradient (27 min, 8–40 % solvent B) in data-dependent MS/MS acquisition mode (Orbitrap resolution 120 K, 1 microscan, HCD 30 with dynamic exclusion of Top 20N). A collection of the data identifying deamidation sites is found in Supplementary Note 6 (Supplementary Table 12 and Supplementary Figs. 24–25). Peptide coverage maps for wild-type and deamidated protein are shown in Supplementary Fig. 26.

**Peptide mapping and HDX MS data analysis.** Precursor and fragment ions were searched and matched against a local protein database in MaxQuant<sup>64</sup> with a minimum score of 20 for unmodified and 40 for modified peptides. For peptides carrying the deamidation site spectra were checked manually and chromatographic peak areas were calculated in Xcalibur (Thermo Scientific) giving a wild-type/deamidated peptide ratio.

DeutEx software (obtained from peterslab.org) was used to determine the deuterium uptake. Excel (Microsoft) and GraphPad Prism software (GraphPad Software, Inc.) were used to create uptake plots and perform statistical analysis. For

comparison of triplicate data a Student's  $T$ -test was used with the  $\alpha$ -value set to 0.01. In addition, a peptide was only considered to have a significant HDX difference if overlapping peptides also showed a difference with  $p < 0.01$  at the respective time point. For comparison of the wild-type and deamidated protein an additional cutoff of  $\Delta\Delta > 0.64$  (99% percentile calculated according to ref. <sup>65</sup>) was used to account for possible day to day variation in the experimental conditions. The peptide coverage map was plotted with MS Tools<sup>66</sup>. To allow access to the HDX data of this study, the HDX Data Summary Table (Supplementary Table 13) is according to the community-based recommendations (Masson, G. R. et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX MS) experiments; manuscript submitted). HDX data tables as well as MS raw data have been deposited to the ProteomeXchange Consortium<sup>67</sup> via the PRIDE<sup>68</sup> partner repository (dataset identifier PXD011914). Deuterium uptake plots are shown in Supplementary Data 1–4 and HDX data are summarized according to community-based recommendations in Supplementary Note 6 (Supplementary Table 13 and Supplementary Data 5).

**Crystallography.** Crystals of deamidated *GII.4* Saga P-dimers were grown at 293 K by hanging drop vapor diffusion. The protein solution was buffered with 25 mM Tris (pH 7.3) and 25 mM NaCl, and the reservoir solution contained 0.2 M Mg formate (pH 5.9) and 20% PEG (w/v) 3350. Crystals grew in a 1:1 mix of protein and reservoir solution at a final protein concentration of 2.5 mg ml<sup>-1</sup> and improved with seeding from initial crystals obtained without seeding and at a protein concentration of 9.1 mg ml<sup>-1</sup>. For soaking, a solution containing the reservoir solution plus 600 mM of methyl- $\alpha$ -L-fucopyranoside (Carbosynth) was prepared. Upon incubation in this solution, crystals were cryo protected by brief transfer to a drop containing soaking solution, glycerol, and reservoir solution without ligand at a ratio of 5:3:2 before flash-freezing in liquid nitrogen. In all, 2400 diffraction images with 0.1° per frame were collected at the X06SA beamline of the Swiss Light Source in Villigen, Switzerland, and processed with XDS<sup>69</sup>.

**Crystallographic model building.** The previously reported Saga P-dimer structure<sup>16</sup> was used for initial refinement against the collected data in *REFMAC*<sup>70</sup>. Difference density maps revealed unambiguous electron density for the methylated fucopyranoside (ePDB ligand description MFU) in both chains of the P-dimer. Water molecules were removed, the ligand was placed manually in the difference map, and the model went through several cycles of real-space refinement in COOT version 0.8.6<sup>71</sup> and restrained reciprocal refinement using PHENIX<sup>72</sup>. Once electron density had improved to the level where the iso-Asp bond became visible in one chain (chain A in the deposited crystallographic model) it was manually inserted in the peptide chain in COOT (ePDB ligand description IAS). Restraints for the isoD373-E372 and isoD373-D374 bonds were generated with the help of the CCP4 JLigand tool<sup>73</sup>. Final model quality was assessed using MolProbity<sup>74</sup>. A table summarizing data collection and refinement is found in Supplementary Note 5 (Supplementary Table 11).

**Reporting summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

### Code availability

Matlab scripts used for data analysis are available from the corresponding author upon request.

### Data availability

Coordinates of the refined crystallographic model and structure factors have been deposited to the protein data bank (pdb) with the accession code 6H9V. NMR assignments have been deposited with the BioMagResBank with the accession code 27445. HDX data tables as well as MS raw data have been deposited to the ProteomeXchange Consortium<sup>67</sup> via the PRIDE<sup>68</sup> partner repository with the dataset identifier PXD011914. HDX MS data are also provided as a Supplementary Data 5 and deuterium uptake plots are provided in Supplementary Data 1–4. Source data for Figs. 2c, 4a, 4b, 4f, 4g, and 5c, and for Supplementary Figures 1a, 1b, 1c, 1d, 1f, 2, 4, 5a, 6, 8, 11, 12a, 12b, 14b, 15, 16, 17, and 21 are provided as a Source Data file. A reporting summary for this Article is available as a Supplementary Information file. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

Received: 12 September 2018 Accepted: 27 February 2019

Published online: 21 March 2019

### References

1. Lindesmith, L. C. et al. Mechanisms of *GII.4* norovirus persistence in human populations. *PLoS Med.* **5**, e31 (2008).

2. Bok, K. et al. Evolutionary dynamics of GII.4 noroviruses over a 34-year period. *J. Virol.* **83**, 11890–11901 (2009).
3. Siebenga, J. J. et al. Epochal evolution of GII.4 norovirus capsid proteins from 1995 to 2006. *J. Virol.* **81**, 9932–9941 (2007).
4. de Graaf, M., van Beek, J. & Koopmans, M. P. Human norovirus transmission and evolution in a changing world. *Nat. Rev. Microbiol.* **14**, 421–433 (2016).
5. Zhang, X. F. et al. Inhibition of histo-blood group antigen binding as a novel strategy to block norovirus infections. *PLoS ONE* **8**, e69379 (2013).
6. Feng, X. & Jiang, X. Library screen for inhibitors targeting norovirus binding to histo-blood group antigen receptors. *Antimicrob. Agents Chemother.* **51**, 324–331 (2007).
7. Rademacher, C. et al. Targeting norovirus infection-multivalent entry inhibitor design based on NMR experiments. *Chemistry* **17**, 7442–7453 (2011).
8. Guiard, J., Fiege, B., Kitov, P. I., Peters, T. & Bundle, D. R. “Double-click” protocol for synthesis of heterobifunctional multivalent ligands: toward a focused library of specific norovirus inhibitors. *Chemistry* **17**, 7438–7441 (2011).
9. Ettayebi, K. et al. Replication of human noroviruses in stem cell–derived human enteroids. *Science* **353**, 1387–1393 (2016).
10. Jones, M. K. et al. Human norovirus culture in B cells. *Nat. Protoc.* **10**, 1939–1947 (2015).
11. Jones, M. K. et al. Enteric bacteria promote human and mouse norovirus infection of B cells. *Science* **346**, 755–759 (2014).
12. Taube, S., Mallagaray, A. & Peters, T. Norovirus, glycans and attachment. *Curr. Opin. Virol.* **31**, 33–42 (2018).
13. Shanker, S. et al. Structural features of glycan recognition among viral pathogens. *Curr. Opin. Struct. Biol.* **44**, 211–218 (2017).
14. Heggelund, J. E., Varrot, A., Imberty, A. & Krenzel, U. Histo-blood group antigens as mediators of infections. *Curr. Opin. Struct. Biol.* **44**, 190–200 (2017).
15. Arda, A. & Jimenez-Barbero, J. The recognition of glycans by protein receptors. Insights from NMR spectroscopy. *Chem. Commun. (Camb.)* **54**, 4761–4769 (2018).
16. Cao, S. et al. Structural basis for the recognition of blood group trisaccharides by norovirus. *J. Virol.* **81**, 5949–5957 (2007).
17. Singh, B. K., Leuthold, M. M. & Hansman, G. S. Human noroviruses’ fondness for histo-blood group antigens. *J. Virol.* **89**, 2024–2040 (2015).
18. Tan, M., Hegde, R. S. & Jiang, X. The P domain of norovirus capsid protein forms dimer and binds to histo-blood group antigen receptors. *J. Virol.* **78**, 6233–6242 (2004).
19. Mallagaray, A., Lockhauserbäumer, J., Hansman, G. S., Uetrecht, C. & Peters, T. Attachment of norovirus to histo blood group antigens: a cooperative multistep process. *Angew. Chem. Int. Ed.* **54**, 12014–12019 (2015).
20. Koromyslova, A. D., Leuthold, M. M., Bowler, M. W. & Hansman, G. S. The sweet quartet: binding of fucose to the norovirus capsid. *Virology* **483**, 203–208 (2015).
21. Fiege, B. et al. Epitope mapping of histo blood group antigens bound to norovirus VLPs using STD NMR experiments reveals fine details of molecular recognition. *Glycoconj. J.* **34**, 679–689 (2017).
22. Mallagaray, A., Rademacher, C., Parra, F., Hansman, G. & Peters, T. STD NMR titrations reveal complex multistep-binding of L-fucose to norovirus particles. *Glycobiology* **27**, 80–86 (2017).
23. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).
24. Frueh, D. P. Practical aspects of NMR signal assignment in larger and challenging proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **78**, 47–75 (2014).
25. Engen, J. R. Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS. *Anal. Chem.* **81**, 7870–7875 (2009).
26. Trabjerg, E., Nazari, Z. E. & Rand, K. D. Conformational analysis of complex protein states by hydrogen/deuterium exchange mass spectrometry (HDX-MS): challenges and emerging solutions. *TrAC Trends Anal. Chem.* **106**, 125–138 (2018).
27. Lindesmith, L. C. et al. Conformational occlusion of blockade antibody epitopes, a novel mechanism of GII.4 human norovirus immune evasion. *mSphere* **3**, <https://doi.org/10.1128/mSphere.00518-17> (2018).
28. Lindesmith, L. C. et al. Particle conformation regulates antibody access to a conserved GII.4 norovirus blockade epitope. *J. Virol.* **88**, 8826–8842 (2014).
29. Donaldson, E. F., Lindesmith, L. C., Lobue, A. D. & Baric, R. S. Viral shape-shifting: norovirus evasion of the human immune system. *Nat. Rev. Microbiol.* **8**, 231–241 (2010).
30. Lee, D., Hilty, C., Wider, G. & Wüthrich, K. Effective rotational correlation times of proteins from NMR relaxation interference. *J. Magn. Reson.* **178**, 72–76 (2006).
31. Dechavanne, V. et al. A high-throughput protein refolding screen in 96-well format combined with design of experiments to optimize the refolding conditions. *Protein Expr. Purif.* **75**, 192–203 (2011).
32. Kay, L. E. & Gardner, K. H. The use of  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  Multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 357–406 (1998).
33. Ulrich, E. L. et al. BioMagResBank. *Nucleic Acids Res.* **36**, D402–D408 (2008).
34. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
35. Müller, M. M. Post-translational modifications of protein backbones: unique functions, mechanisms, and challenges. *Biochemistry* **57**, 177–185 (2018).
36. Hagan, R. M. et al. NMR spectroscopic and theoretical analysis of a spontaneously formed Lys-Asp isopeptide bond. *Angew. Chem. Int. Ed. Engl.* **49**, 8421–8425 (2010).
37. Tugarinov, V., Muhandiram, R., Ayed, A. & Kay, L. E. Four-Dimensional NMR Spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase G. *J. Am. Chem. Soc.* **124**, 10025–10035 (2002).
38. Taube, S. et al. A mouse model for human norovirus. *mBio* **4**, <https://doi.org/10.1128/mBio.00450-13> (2013).
39. Hansman, G. S. et al. Crystal structures of GII.10 and GII.12 norovirus protruding domains in complex with histo-blood group antigens reveal details for a potential site of vulnerability. *J. Virol.* **85**, 6687–6701 (2011).
40. Koromyslova, A., Tripathi, S., Morozov, V., Schrotten, H. & Hansman, G. S. Human norovirus inhibition by a human milk oligosaccharide. *Virology* **508**, 81–89 (2017).
41. Baryshnikova, O. K., Williams, T. C. & Sykes, B. D. Internal pH indicators for biomolecular NMR. *J. Biomol. NMR* **41**, 5–7 (2008).
42. Hansman, G. S. et al. Structural basis for norovirus inhibition and fucose mimicry by citrate. *J. Virol.* **86**, 284–292 (2012).
43. Wegener, H. et al. Human norovirus GII.4(MI001) P dimer binds fucosylated and sialylated carbohydrates. *Glycobiology* **27**, 1027–1037 (2017).
44. Yan, Y. et al. Isomerization and oxidation in the complementarity-determining regions of a monoclonal antibody: a study of the modification-structure-function correlations by hydrogen-deuterium exchange mass spectrometry. *Anal. Chem.* **88**, 2041–2050 (2016).
45. Robinson, A. B., McKerrow, J. H. & Cary, P. Controlled deamidation of peptides and proteins: an experimental hazard and a possible biological timer. *Proc. Natl Acad. Sci. USA* **66**, 753–757 (1970).
46. Robinson, N. E. & Robinson, A. B. Molecular clocks. *Proc. Natl Acad. Sci. USA* **98**, 944–949 (2001).
47. Griffith, S. C. et al. Crystal structure of a protein repair methyltransferase from *Pyrococcus furiosus* with its L-isoaspartyl peptide substrate. *J. Mol. Biol.* **313**, 1103–1116 (2001).
48. McFadden, P. N. & Clarke, S. Conversion of isoaspartyl peptides to normal peptides: implications for the cellular repair of damaged proteins. *Proc. Natl Acad. Sci. USA* **84**, 2595 (1987).
49. Geiger, T. & Clarke, S. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *J. Biol. Chem.* **262**, 785–794 (1987).
50. Hao, P., Adav, S. S., Gallart-Palau, X. & Sze, S. K. Recent advances in mass spectrometric analysis of protein deamidation. *Mass Spectrom. Rev.* **36**, 677–692 (2017).
51. Grassi, L. et al. Complete NMR assignment of succinimide and its detection and quantification in peptides and intact proteins. *Anal. Chem.* **89**, 11962–11970 (2017).
52. Eakin, C. M., Miller, A., Kerr, J., Kung, J., Wallace, A. Assessing analytical methods to monitor isoAsp formation in monoclonal antibodies. *Front. Pharmacol.* **5**, <https://doi.org/10.3389/fphar.2014.00087> (2014).
53. Noguchi, S., Miyawaki, K. & Satow, Y. Succinimide and isoaspartate residues in the crystal structures of hen egg-white lysozyme complexed with tri-N-acetylchitotriose. *J. Mol. Biol.* **278**, 231–238 (1998).
54. Chazin, W. J. et al. Identification of an isoaspartyl linkage formed upon deamidation of bovine calbindin D9k and structural characterization by 2D proton NMR. *Biochemistry* **28**, 8646–8653 (1989).
55. Lee, J. C. et al. Protein L-isoaspartyl methyltransferase regulates p53 activity. *Nat. Commun.* **3**, <https://doi.org/10.1038/ncomms1933> (2012).
56. Deverman, B. E. et al. Bcl-xL deamidation is a critical switch in the regulation of the response to DNA damage. *Cell* **111**, 51–62 (2002).
57. Kumar, S. et al. Unexpected functional implication of a stable succinimide in the structural stability of *Methanocaldococcus jannaschii* glutaminase. *Nat. Commun.* **7**, <https://doi.org/10.1038/ncomms12798> (2016).
58. Corti, A. & Curnis, F. Isoaspartate-dependent molecular switches for integrin-ligand recognition. *J. Cell Sci.* **124**, 515–522 (2011).
59. Tan, M. & Jiang, X. Vaccine against norovirus. *Hum. Vaccin. Immunother.* **10**, 1449–1456 (2014).
60. Robinson, N. E. & Robinson, A. B. Deamidation of human proteins. *Proc. Natl Acad. Sci. USA* **98**, 12409–12413 (2001).

61. Robinson, N. E. & Robinson, A. B. Prediction of protein deamidation rates from primary and three-dimensional structure. *Proc. Natl Acad. Sci. USA* **98**, 4367–4372 (2001).
62. Plotnikov, N. V., Singh, S. K., Rouse, J. C. & Kumar, S. Quantifying the risks of asparagine deamidation and aspartate isomerization in biopharmaceuticals by computing reaction free-energy surfaces. *J. Phys. Chem. B* **121**, 719–730 (2017).
63. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
64. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
65. Arora, J. et al. Hydrogen exchange mass spectrometry reveals protein interfaces and distant dynamic coupling effects during the reversible self-association of an IgG1 monoclonal antibody. *Mabs* **7**, 525–539 (2015).
66. Kavan, D. & Man, P. MSTools—Web based application for visualization and presentation of HXMS data. *Int. J. Mass Spectrom.* **302**, 53–58 (2011).
67. Deutsch, E. W. et al. The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. *Nucleic Acids Res.* **45**, D1100–D1106 (2017).
68. Vizcaino, J. A. et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456 (2016).
69. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 133–144 (2010).
70. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
71. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
72. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
73. Lebedev, A. A. et al. JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 431–440 (2012).
74. Davis, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).

## Acknowledgements

T.P., B.S.B., T.S., and C.U. thank the Deutsche Forschungsgemeinschaft (DFG) for grants Pe494/12–1, BL1294/3–1, STE 1463/7–1, and UE 183/1–1, respectively (all three FOR2327, ViroCarb). We would like to thank Dr. Thorsten Biet for support with the NMR experiments, Dr. Hanne Peters for enzymatic synthesis of blood group B trisaccharide, and Prof. Javier Pérez Castells for providing us with blood group H disaccharide. We are grateful for support within the iNEXT program (PID 1483 and PID 2254) giving us access to the NMR high-field facilities at the Bijvoet Center in Utrecht. In particular, we would like to thank Dr. Hans Wienk from the Bijvoet Center for assisting us in any respect. We thank Dr. Mario Schubert (University of Salzburg, Austria) for helpful discussions. J.D. and C.U. would like to thank Prof. H. Schlüter (UKE, University of Hamburg, Germany) for access to high-resolution mass spectrometers and Dr. Daniel Kavan, Dr. Petr Man, and Dr. Alan Kádek for providing the DeutEx software. J.D. acknowledges funding from COST BM1403 and FOR2327 ViroCarb. C.U. acknowledges funding from the Leibniz Association through SAW-2014-HPI-4 grant.

The Heinrich-Pette-Institute, Leibniz Institute for Experimental Virology is supported by the Freie und Hansestadt Hamburg and the Bundesministerium für Gesundheit (BMG). We thank Dr. Grant Hansman (University of Heidelberg, Germany) for providing us with the plasmids of the P-domains of *GII.4* Saga, *GII.10* Vietnam 026, and *GII.17* Kawasaki 308. Prof. Dr. Stefan Taube (University of Lübeck, Germany) is thanked for providing the plasmid of the P-domain of *GII.4* MI001. J.M.O. thanks to Fundación Universitaria San Pablo CEU for a mobility grant.

## Author contributions

A.M., R.C., J.D., B.S.B., C.U., and T.P. designed experiments. A.M., R.C. and L.L.G. performed the NMR experiments, and A.M., R.C., L.L.G., and T.P. analyzed and interpreted the NMR data. A.M. and R.C. expressed and purified the P-domains. A.M. established the unfolding-refolding protocol and wrote Matlab scripts for data analysis. R.C. established the IEX protocol and analyzed deamidation rates. J.M.O. expressed and purified Kawasaki 308 P-domains. J.D., E.T., and K.D.R. performed the MS experiments and J.D., E.T., K.D.R., and C.U. analyzed and interpreted the MS data. P.H.O.M. and B.S.B. performed the X-ray experiments, and P.H.O.M., T.S., and B.S.B. analyzed and interpreted the corresponding data. B.S.B., C.U., and T.P. wrote the paper. All authors contributed to data interpretation and commented on the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-09251-5>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Journal peer review information:** *Nature Communications* thanks Bidadi Prasad and the other anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019