# Establishing structural characterization of noroviruses by hydrogen-deuterium exchange mass spectrometry (HDX-MS)

## Dissertation

with the aim of achieving a doctoral degree at the Faculty of

Mathematics, Informatics and Natural Sciences

Department of Chemistry
Universität Hamburg

submitted by

## Jasmin Dülfer

born in Potsdam

Hamburg, 2020

# TABLE OF CONTENTS

## PUBLICATION LIST

Parts of this thesis are published or in preparation for publication. A detailed overview of my contributions to these publications can be found in the supplement.

(1) Pogan, R., **Dülfer, J.** & Uetrecht, C. (2018). Norovirus assembly and stability. Current opinion in virology, 31, 59-65.

(2) Mallagaray, A., Creutznacher, R., **Dülfer, J**., Mayer, P. H. O., Grimm, L. L., Orduna, J. M., Trabjerg, E., Stehle, T., Rand, K. D., Blaum, B. S., Uetrecht, C. & Peters, T. (2019). A post-translational modification of human Norovirus capsid protein attenuates glycan binding. Nature communications, 10(1), 1-14.

(3) **Dülfer, J.**, Kadek, A., Kopicki, J. D., Krichel, B. & Uetrecht, C. (2019). Structural mass spectrometry goes viral. Advances in Virus Research, 105, 189-238.

(4) **Dülfer, J.**, Yan, H., Brodmerkel, M., Creutznacher, R., Mallagaray, A., Peters, T., Caleman, C., Marklund, E. & Uetrecht, C.: Glycan-induced structural dynamics in human norovirus P domains depend on virus strain and deamidation status (in preparation)

(5) Krainer, G., Batet, M., **Dülfer, J.**, Anadamurugan, A., Textor, M., Frotscher, E., Hartmann, A., Uetrecht, C., Keller, S. & Schlierf, M. (in preparation)

(6) Creutznacher, R., Feldmann, C., **Dülfer, J.**, Maass, T., Knickmann, J., Uetrecht, C., Peters, T. & Taube, S.: Murine norovirus capsid plasticity probed by NMR - Bile acid binding stabilizes and rearranges P domain dimers and triggers immune escape. (in preparation)

## ABBREVIATIONS

| | |
|---|---|
| **ACN** | acetonitrile |
| **APS** | ammonium persulfate |
| **BB** | backbone |
| **BE** | back exchange |
| **CCS** | collisional cross-section |
| **CD** | circular dichroism |
| **CID** | collision-induced dissociation |
| **CMC** | critical micelle concentration |
| **CSP** | chemical shift perturbation |
| **Da** | Dalton |
| **DC** | direct current |
| **DDM** | $n$-dodecyl-β-D-maltoside |
| **DLS** | dynamic light scattering |
| **DOPC** | dioleoylphosphatidylcholine |
| **DPC** | n-dodecyl phosphocholine |
| **EM** | electron microscopy |
| **ESI** | electrospray ionization |
| **ETD** | electron-transfer dissociation |
| **FA** | formic acid |
| **FCV** | feline calicivirus |
| **FD** | fully deuterated |
| **fJAM-A** | feline junctional adhesion molecule A |
| **FWHM** | full width at half maximum |
| **GCDCA** | glycochenodeoxycholic acid |
| **GI** | genogroup I |
| **GndHCl** | guanidine hydrochloride |
| **H1** | helix 1 |
| **HBGA** | histo-blood group antigen |
| **HCD** | higher-energy collisional dissociation |
| **HDX** | hydrogen/deuterium exchange |
| **H-ESI** | heated-electrospray ionization |
| **HPLC** | high performance liquid chromatography |
| **HSQC** | heteronuclear single quantum coherence |
| **iDiD** | fully deamidated (isoAsp/isoAsp) P dimer |
| **iDN** | half-deamidated (isoAsp/Asn) P dimer |
| **IEX** | ion exchange |
| **IM** | ion mobility |
| **IPTG** | isopropyl β-d-1-thiogalactopyranoside |
| **IRMPD** | infrared multiple photon dissociation |

| | |
|---|---|
| $K_d$ | dissociation constant |
| LC | liquid chromatography |
| LDAO | lauryldimethylamine *N*-oxide |
| LUV | large unilamellar vesicle |
| MALDI | matrix-assisted laser desorption-ionization |
| MCP | multi-channel plate |
| MD | molecular dynamics |
| MNV | murine norovirus |
| MS | mass spectrometry |
| MS/MS | tandem mass spectrometry |
| MWCO | molecular weight cut-off |
| *m/z* | mass to charge ratio |
| NMR | nuclear magnetic resonance |
| NN | fully native (Asn/Asn) P dimer |
| OD | optical density |
| ORF | open reading frame |
| P domain | protruding domain |
| PDB | protein data bank |
| PFU | plaque-forming unit |
| PLRP-S | polymeric reversed-phase column |
| POPC | palmitoyloleoylphosphatidylcholine |
| PTM | post-translational modification |
| RdRp | RNA dependent RNA polymerase |
| RF | radio frequency |
| S domain | shell domain |
| SC | side chain |
| SD | standard deviation |
| SDS-PAGE | sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| SID | surface-induced dissociation |
| smFRET | single molecule fluorescence resonance energy transfer |
| *T* | triangulation number |
| TROSY | transverse relaxation-optimized spectroscopy |
| TEMED | *N,N,N',N'*-tetramethyl ethylenediamine |
| TOF | time-of-flight |
| UPLC | ultra-high-pressure liquid chromatography |
| UVPD | ultraviolet photodissociation |
| VLP | virus-like particle |
| VP1/2 | major/minor capsid protein |
| wt | wild type |

# ABSTRACT

Over the last twenty years, mass spectrometry (MS) has more and more entered the field of structural virology aiming to investigate the structure and dynamics of viral proteins as close to their native environment as possible. The use of non-perturbing labels in hydrogen-deuterium exchange (HDX) MS allows for the analysis of interactions between viral proteins and host cell factors as well as their dynamic responses to the environment.

Infection with norovirus is the leading cause of acute gastroenteritis worldwide. Attachment of human norovirus to histo blood group antigens (HBGAs) is essential for infection, but how this binding event promotes the infection of host cells is unknown. HBGA binding to the norovirus capsid is mediated by the P domain of the viral capsid protein VP1 which bind glycans in a strain-dependet manner, with fucose as minimal binding motif. To track strain-specific structural changes on norovirus P dimers upon glycan binding that could hypothetically prime for cellular uptake, an HDX-MS workflow was established.

A first set of experiments aimed to investigate the structural impact of a highly selective spontaneous transformation of asparagine 373 into an iso-aspartate residue, located in a loop adjoining the HBGA binding site. This post-translational modification (PTM) proceeds with an estimated half-life of a few days at physiological temperatures. N373 deamidation increases protein dynamics in the HBGA binding pocket and strongly attenuates HBGA recognition in fully deamidated P dimers. Sequence conservation and the surface-exposed position of this PTM suggest an important role in infection and immune recognition for many norovirus strains, especially those of the pandemic genogroup GII.4.

In a second set of experiments, HDX-MS was used to investigate glycan-induced structural dynamics in P dimers of different norovirus strains, which exhibit high structural similarity but different prevalence in humans. While the almost identical strains GII.4 Saga and GII.4 MI001 share similar glycan-induced dynamics, the dynamics differ in the emerging GII.17 Kawasaki 308 and rare GII.10 Vietnam 026 strain. Structural effects of N373 deamidation upon glycan binding were also investigated in partially deamidated GII.4 P dimers, which are likely present during infection. These mixed species exhibit increased solvent exposure in the P2 domain of the P dimer upon glycan binding in contrast to pure wild type. Furthermore, deamidated P dimers display increased flexibility and

dissociation into monomers. The results indicate that glycan binding induces strain-dependent structural dynamics, which are further altered by N373 deamidation. Therefore, a role of deamidation in modulating cell attachment and entry in GII.4 strains can be supposed.

The standard HDX-MS workflow was also adapted for the investigation of protein-membrane interactions. Using the membrane protein Mistic as a model protein, dynamics in different lipid and detergent environments were investigated. HDX-MS results show that three regions of the protein have significant deuterium uptake differences in lipid vesicles and detergent micelles, with two of these regions displaying bimodal peak distributions. The deuteration differences imply that Mistic conformation and dynamics are influenced by lipid properties, which could have some implications on the biological function.

The HDX-MS workflow established here enables structural dynamics studies of viral proteins in solution as well as in more challenging environments, paving the way for further HDX-MS measurements with virus-like particles and infectious virions and their interactions with membrane-displayed glycans.

## ZUSAMMENFASSUNG

In den letzten zwanzig Jahren hat die Massenspektrometrie (MS) mehr und mehr Einzug in den Bereich der Strukturvirologie gehalten mit dem Ziel, die Struktur und Dynamik viraler Proteine so nah wie möglich an ihrer nativen Umgebung zu untersuchen. Die Verwendung von nicht-störenden Markierungen im Wasserstoff-Deuterium-Austausch (HDX) MS ermöglicht die Analyse von Wechselwirkungen zwischen viralen Proteinen und Wirtszellfaktoren sowie deren dynamische Reaktionen auf die Umgebung.

Die Infektion mit dem Norovirus ist weltweit die Hauptursache für akute Gastroenteritis. Die Anheftung des Norovirus an Histo-Blutgruppenantigene (HBGAs) ist für die Infektion im Menschen essentiell, wie jedoch dieses Bindungsereignis die Infektion von Wirtszellen fördert, ist unbekannt. Die HBGA-Bindung an das Norovirus-Kapsid wird durch die P-Domäne des viralen Kapsidproteins VP1 vermittelt, die Glykane stammabhängig bindet. Fucose dient hierbei als minimales Bindungsmotiv. Um stammspezifische Strukturveränderungen an Norovirus-P-Dimeren bei der Glykanbindung zu verfolgen, die hypothetisch zur zellulären Aufnahme führen könnten, wurde ein HDX-MS-Workflow etabliert.

Eine erste Reihe von Experimenten zielte darauf ab, die strukturellen Auswirkungen einer hochselektiven spontanen Transformation von Asparagin 373 in einen Iso-Aspartat-Rest zu untersuchen, welcher sich in einer an die HBGA-Bindungsstelle angrenzenden Loop-Struktur befindet. Diese posttranslationale Modifikation (PTM) verläuft bei physiologischen Temperaturen mit einer geschätzten Halbwertszeit von wenigen Tagen. Die Deamidierung von N373 erhöht die Proteindynamik in der HBGA-Bindungstasche und schwächt die HBGA-Erkennung in vollständig deamidierten P-Dimeren stark ab. Die Sequenzkonservierung und die oberflächenexponierte Position dieser PTM deuten auf eine wichtige Rolle bei der Infektion und Immunerkennung für viele Norovirusstämme hin, insbesondere für die der pandemischen Genogruppe GII.4.

In einer zweiten Versuchsreihe wurde HDX-MS zur Untersuchung der glykaninduzierten Strukturdynamik in P-Dimeren verschiedener Norovirusstämme verwendet, die eine hohe strukturelle Ähnlichkeit, aber eine unterschiedliche Prävalenz beim Menschen aufweisen. Während die fast identischen Stämme GII.4 Saga und GII.4 MI001 eine ähnliche glykaninduzierte

Dynamik aufweisen, unterscheidet sich die Dynamik bei dem neu auftretenden Stamm GII.17 Kawasaki 308 und dem seltenen Stamm GII.10 Vietnam 026. Die strukturellen Auswirkungen der N373-Deamidierung auf die Glykanbindung wurden auch an teilweise deamidierten GII.4 P-Dimeren untersucht, die wahrscheinlich während der Infektion vorhanden sind. Diese gemischten Spezies zeigen im Gegensatz zum reinen Wildtyp eine erhöhte Lösungsmittelexposition in der P2-Dömane des P-Dimers bei der Glykanbindung. Des Weiteren zeigen deamidierte P-Dimere eine erhöhte Flexibilität und Dissoziation in Monomere. Die Ergebnisse deuten darauf hin, dass die Glykanbindung eine stammabhängige Strukturdynamik induziert, die durch die N373-Deamidierung weiter modifiziert wird. Daher kann eine Rolle der Deamidierung bei der Modulation der Zellanheftung und des Zelleintritts in GII.4-Stämmen vermutet werden.

Der Standard-HDX-MS-Workflow wurde auch für die Untersuchung von Protein-Membran-Interaktionen angepasst. Mit dem Membranprotein Mistic als Modellprotein wurde die Dynamik in verschiedenen Lipid- und Detergensumgebungen untersucht. Die Ergebnisse der HDX-MS zeigen, dass drei Regionen des Proteins signifikante Unterschiede in der Deuteriumaufnahme in Lipidvesikeln und Detergens-Mizellen aufweisen, wobei zwei dieser Regionen durch bimodale Signalverteilungen gekennzeichnet sind. Die Deuterierungsunterschiede implizieren, dass die Konformation und Dynamik von Mistic durch die Lipideigenschaften beeinflusst werden, was einige Auswirkungen auf die biologische Funktion haben könnte.

Der hier etablierte HDX-MS-Workflow ermöglicht strukturdynamische Studien viraler Proteine sowohl in Lösung als auch in anspruchsvolleren Umgebungen und ebnet damit den Weg für weitere HDX-MS-Messungen mit virusähnlichen Partikeln und infektiösen Virionen und deren Wechselwirkungen mit membranständigen Glykanen.

# 1 INTRODUCTION

## 1.1 Structural mass spectrometry

Parts of the mass spectrometry introduction have been published in the following book chapter: *Dülfer, J., Kadek, A., Kopicki, J. D., Krichel, B., & Uetrecht, C. (2019). Structural mass spectrometry goes viral. Advances in Virus Research, 105, 189-238.* Reprinted from Dülfer et al. (2019), Copyright (2019), with permission from Elsevier.

### 1.1.1 HOW STRUCTURAL MASS SPECTROMETRY CAN SERVE VIROLOGICAL QUESTIONS

The ultimate goal in structural virology is to understand how the structure of a virus influences its interactions with a host organism and thus governs its life cycle. Therefore, virologists nowadays employ a broad range of structural biology methods to uncover the organization of viral particles and to understand their function at the molecular level. Macromolecular crystallography techniques as well as nuclear magnetic resonance (NMR) spectroscopy routinely provide three dimensional structures with resolution down to single atoms, while electron microscopy-based (EM) methods are fast approaching this boundary with cryogenic sample preparation and recent developments in detector technology (Merk et al., 2016). However, all these techniques have their own challenges and limitations. These can range from the often frustrating efforts to produce sufficiently large, diffracting crystals, through sample size limitations, up to often very significant demands on sample amount and homogeneity or limited ability to assign subunit positions or identify post-translational modifications (PTM). Most importantly, however, there is very rarely a single technique to answer all questions at hand. Therefore, structural biology is moving more and more in the direction of so-called integrative approaches, where high-resolution structural data from different methods are combined with lower resolution information. These data then ultimately feed into computational modeling in order to obtain the most detailed functional understanding possible of the biological system (Marklund and Benesch, 2019, Politis and Schmidt, 2018, Ward et al., 2013). One of the experimental techniques, which provide invaluable input for the integrative structural efforts, is mass spectrometry (MS). Originally belonging

purely to the domains of physics and analytical chemistry, MS has changed significantly since the revolutionary developments of electrospray ionization (ESI) and matrix-assisted laser desorption-ionization (MALDI) in the late 1980s (Karas et al., 1987, Tanaka et al., 1988, Yamashita and Fenn, 1984). These two so-called 'soft-ionization techniques' have enabled the analyses of biomolecules without their unwanted fragmentation and allowed MS to expand hugely into the field of biochemistry. There, it has ever since been used in proteomics as a method of choice for identifying proteins, quantifying them and for analyzing their PTMs (Luo and Muesing, 2014, Yates et al., 2009). Moreover, and more importantly from the point of view of higher order macromolecular structure, so-called structural MS has also established itself especially in the last two decades as a powerful method to study proteins. Pioneering works by the groups of Carol Robinson, Albert Heck and Joseph Loo have exploited the softness of the ESI process to transfer whole non-covalently bound protein assemblies into the gas phase inside a mass spectrometer without disrupting their stoichiometry and interactions in a process termed native MS (Leney and Heck, 2017, Loo, 1995, Rostom and Robinson, 1999, Van Berkel et al., 2000). Additionally, as was shown later, the conformations of such protein complex ions often survive largely unperturbed, which opened up possibilities for further gas phase structural studies and 'gas-phase structural biology' (Ruotolo et al., 2005, Seo et al., 2016, Uetrecht et al., 2008). Furthermore, many techniques were developed which probe protein conformations, interactions and dynamics directly in solution. Usually, such methods use some form of chemical labeling – be it through isotopic exchange (Katta and Chait, 1991), chemical crosslinking (Rossi et al., 1995) or radical chemistry (Maleknia et al., 2001) – and capitalize from the use of MS as a detection method.

Together, all the various structural MS approaches offer a broad portfolio of experimental techniques, which can probe a wide range of protein structural features, while at the same time retrieving information about protein identity and PTMs. Moreover, structural MS nicely complements the three traditional high-resolution structural biology techniques, albeit with lower overall resolution. However, this is often more than balanced by the ability of MS to work with low amounts of samples, in low concentrations, and with high throughput. Importantly, structural MS can often work within complex mixtures, with several pioneering methods probing protein structures directly within cells (Chavez et al., 2018, Chorev et al., 2018, Espino et al., 2015, Gan et al., 2017, Garcia-Moreno

et al., 2019, Nguyen et al., 2018). Additionally, structural MS is in theory unlimited regarding molecular weight, offering information on systems ranging from small isolated protein domains all the way up to huge multi-megadalton assemblies. Perhaps most importantly, structural MS can also provide insight into protein conformational dynamics and co-existing transient species in solution, which is frequently missing or can be very challenging to access by static, averaged high-resolution structures provided by crystallography and EM.

## 1.1.2  NATIVE MASS SPECTROMETRY (NATIVE MS)

The basis of native MS is very soft ESI from volatile buffer surrogates. The main challenge is retaining non-covalent interactions of proteins and their stoichiometry in complexes upon transfer into the gas phase. However, ESI is soft enough that intact viruses have been shown to survive the process and keep their infectivity (Hogan et al., 2006, Siuzdak et al., 1996). The ideal protein ion is free of all solvent and buffer adducts but populated by multiple charges. In ESI, the protein sample is sprayed from an externally charged capillary into the mass spectrometer (Figure 1). The physical background regarding desolvation, protein charging and structural integrity of gas phase ions was unraveled experimentally and in MD simulations described elsewhere (Konermann et al., 2013, McAllister et al., 2015). The ESI process for native MS presented herein requires buffer exchange into an MS-compatible solution of volatile salt. For that purpose, even though not buffering at neutral pH, ammonium acetate is particularly suitable, due to its adjustable pH (4.6-10) and ionic strength (0 to several M), mimicking the natural environment and providing favorable conditions for the protein until its desolvation (Konermann, 2017). Small amounts of non-volatile salts or small molecules up to a few mM can be added when required, for example bivalent cations or co-factors. Ideally, native MS reveals the charge state distributions of all mass species, including any intermediates, present in the sample in solution. The mass can be deduced from the $m/z$ values of neighboring peaks in a charge state distribution, where charge states of the same species differ in $z$ by 1. For similar sized analytes, the signal intensity is proportional to the analytes concentration; however, large oligomeric complexes deviate from this direct correlation at higher concentrations, which hampers a reliable quantification (Root et al., 2017).
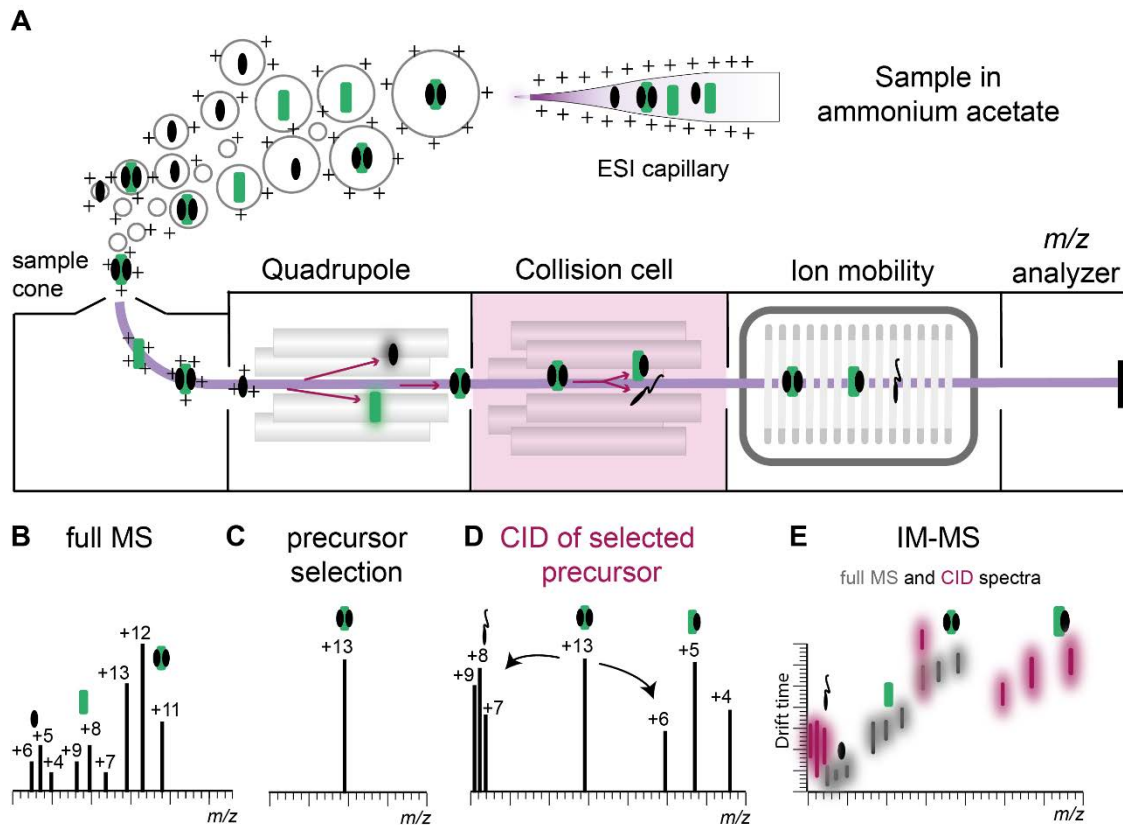
Figure 1: Exemplary workflow and readout in native MS. (A) A protein complex with two subunits (black and green) is sprayed from ammonium acetate solution, via an electric potential and enters the mass spectrometer as molecular ions. The modules of the mass spectrometer allow for manipulation and analysis *in vacuo* as depicted in schematic spectra. (B) Full ESI-MS spectra reveal a Gaussian distribution of the charge states of each mass of the sample proteins. An additional charge distribution indicates a trimeric complex. (C) In order to confirm subunit stoichiometry of the candidate complex, its molecular ions are filtered for in the quadrupole region of the mass spectrometer. (D) Selected ions are subsequently collided with a neutral gas in the collision cell. The complex dissociates into two subunits of characteristic masses, which confirms 2:1 stoichiometry. The dissociated subunit unfolds and takes a large number of charges with it. It is therefore detected in the lower $m/z$ range of the spectrum, whereas the remainder of the complex is detected at higher $m/z$. (E) Additional information about conformation is obtained by ion mobility MS (IM-MS), where drift time allows an approximation of the CCS in the gas phase. Here, IM spectra of full MS (B) and CID (D) are compared, where the smaller subunit has a substantially longer drift time when ejected from the complex, due to its unfolding process during dissociation. IM-MS studies become particularly powerful when a protein complex already has different subunit conformations in solution. Reprinted from Dülfer et al. (2019), Copyright (2019), with permission from Elsevier.

Nevertheless, semi-quantification and hence characterizing binding affinities is possible as has been shown in antibody dimerization (Rose et al., 2011). Determining the mass of a species of interest is in itself not always sufficient to deduce its stoichiometry so that further degrees of separation or MS/MS fragmentation approaches are necessary.

One approach often coupled with MS to provide a semi-orthogonal separation is ion mobility (IM). IM-MS is able to separate ions according to their "retention" time in a region with elevated pressure of an inert gas (usually nitrogen or helium). The retention or arrival time is proportional to $m/z$ and overall shape, which allows distinguishing coexisting protein conformers with same mass but different structures, as well as protein complex structure and ligand interactions (Eschweiler et al., 2017, Goth and Pagel, 2017). Since the $m/z$ is known from MS, the arrival time can be converted into a rotationally averaged collisional cross section (CCS) (Uetrecht et al., 2010). The CCS allows drawing conclusions from the gas phase structure relating back to the protein structure in solution.

The charge acquired during the ESI process allows protein manipulation and detection in the gas phase, opening the door to the application of methods that are incompatible with having the protein in solution. For example, to detect the stoichiometry unambiguously, a complex ion can be filtered out or selected in a quadrupole and dissociated subsequently (Figure 1C and D). For this purpose, collision induced dissociation (CID) is mostly used, in which additional energy is imparted to the analyte ions via high-velocity collisions with heavy inert gas atoms. This gradually increases the internal energy of the ion until individual subunits of a non-covalent complex unfold and are eventually ejected from the complex one-by-one (Pagel et al., 2010).

Mass information obtained by the detection of the ejected subunits and the remaining subcomplexes unequivocally proves stoichiometry of the original complex. The order of subunit ejection also correlates with peripheral location in the complex and thus provides even more insight into complex architectures. Notably, this dissociation behavior is usually distinct from what is observed in solution, where it is a spontaneous (often reversible) equilibrium reaction rather than an irreversible consequence of physical collisions with gas molecules (Boeri Erba and Petosa, 2015). The described method is especially informative if a specific complex species is singled out according to its $m/z$ in the gas phase before CID occurs, so that dissociation products can be traced back to one single complex species. In principle, any combination of protein fragmentation with

(native) MS analysis of intact proteins is called top down MS. When harsh CID conditions are used, backbone fragmentation of the intact protein occurs on top of subunit dissociation of the protein complex. Additionally, these backbone fragmentations are often induced by other fragmentation techniques such as ultraviolet photodissociation (UVPD) and infrared multiphoton dissociation (IRMPD) (Brodbelt, 2014), impacts with a surface (SID) (Stiving et al., 2019) or electron transfer reactions (ETD) (Lermyte et al., 2018) in the gas phase. Thus, information about the primary structure of a protein, its PTMs as well as its presence in a complex can be obtained.

### 1.1.3 HYDROGEN-DEUTERIUM EXCHANGE MASS SPECTROMETRY (HDX-MS)

Hydrogen-deuterium exchange (HDX)-MS is a sensitive technique to localize regions of conformational dynamics in proteins and identify amino acids engaged in ligand binding under close to native conditions in solution. As opposed to other MS experiments, HDX can be performed in practically any buffer condition and with additives or co-factors that are required for proper protein folding and function – the sole difference is buffer preparation from deuterated water. The method exploits the naturally occurring exchange between solvent-accessible protein backbone hydrogens and deuterium in solution over time, measured as mass increase using MS (Konermann et al., 2011, Trabjerg et al., 2018).

#### H/D EXCHANGE

The rate of exchange is influenced by several factors, including solution conditions such as pH and temperature as well as by the peptides' primary structure (Bai et al., 1993). The chemical HDX rate constant ($k_{ch}$) is the sum of acid-, base- and water-catalyzed exchange reactions of amide hydrogens with deuterated water. The rate of the base catalysis far exceeds that of acid and water catalysis, which results in a high pH dependence of $\log(k_{ch})$(Figure 2A and B).

The pH dependence of the exchange reaction enables measurement of incorporated deuterium by MS because the exchange reaction can be quenched by reducing pH to 2.5-3.0, where the rates of acid- and base-catalyzed exchange are equal and $k_{ch}$ is at its minimum. Thus, under quench conditions D/H back exchange is sufficiently low to allow time for detection of incorporated deuterium by mass spectrometry (Englander, 2006).
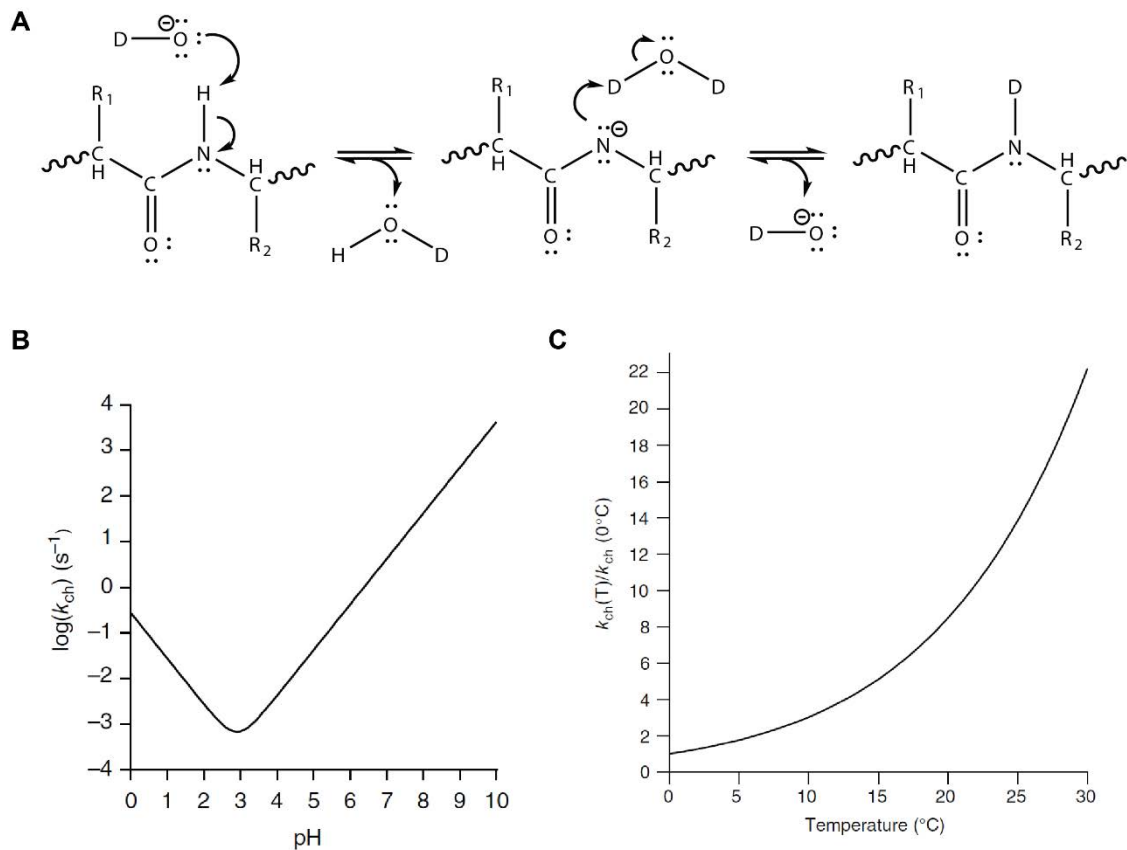
Figure 2: Effects of pH and temperature on the chemical HDX rate $k_{ch}$. (A) Base-catalyzed amide HDX reaction. (B) Example of a pH profile of $k_{ch}$ of an amide hydrogen in a random coil-like poly-DL-alanine peptide at 20°C. (C) The chemical exchange rate as a function of temperature for base-catalyzed exchange. Panels B and C are reprinted from Smith et al. (1997) and Jensen and Rand (2016) with permission from John Wiley and Sons.

As $k_{ch}$ is also exponentially dependent on temperature (Figure 2C), labeling is usually done at 25 °C, while temperature is reduced to 0 °C under quench conditions (Jensen and Rand, 2016). The main impacts on HDX rates, however, come from interactions in native protein structures. The involvement of hydrogens in hydrogen bonding within higher order elements of protein structure (e.g. in α-helices and β-sheets) or in interactions with ligands slows down the exchange rate for respective backbone amide hydrogens. Protection from solvent (e.g. in a hydrophobic core of the protein or at a protein-protein interface) also typically leads to a decrease in exchange in the affected protein region. On the other hand, conformational dynamics and flexibility of protein regions influence the exchange rate too, as the exposure to deuterated solvent accelerates HDX.

# 1 Introduction

All proteins naturally fluctuate through transient, locally unfolded states in solution. In the Linderstrøm-Lang model (Englander et al., 1997), deuterium exchange occurs at these moments when a region of protein structure is temporarily loosened up and the hydrogen bonds are momentarily interrupted (illustrated in Equation 1).

Equation 1

$$N\text{—}H_{closed} \xrightleftharpoons[k_{cl}]{k_{op}} N\text{—}H_{open} \xrightarrow[D_2O]{k_{ch}} N\text{—}D_{open} \xrightleftharpoons[k_{op}]{k_{cl}} N\text{—}D_{closed}$$
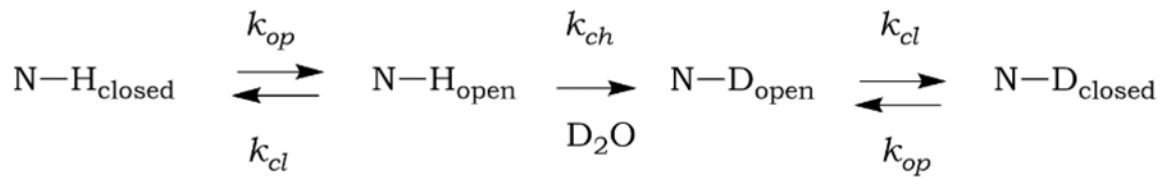
Here, $k_{op}$ and $k_{cl}$ are the rate constants for the opening and closing reaction and $k_{ch}$ the intrinsic chemical H/D exchange rate. The isotopic exchange rate ($k_{HDX}$) can then be defined as

Equation 2
$$k_{HDX} = \frac{k_{op} * k_{ch}}{k_{op} + k_{cl} + k_{ch}}$$

Under native conditions, proteins typically refold fast after unfolding ($k_{cl} \gg k_{op}$) and Equation 2 can be approximated by

Equation 3
$$k_{HDX} = \frac{k_{op} * k_{ch}}{k_{cl} + k_{ch}}$$

If the refolding of such temporarily more exposed regions is slower than the chemical exchange rate, the respective population will be more deuterated than the folded population, showing as a discreet jump in deuteration (EX1 kinetics or correlated exchange). As a result, the isotopic exchange rate is equal to the rate of opening (Equation 4).

**EX1 kinetics** (correlated exchange):  $k_{cl} \ll k_{ch}$

Equation 4
$$k_{HDX} = k_{op}$$

EX1 kinetics result in bimodal peak distributions in deuterated peptide spectra (Figure 3 B). They are often observed for proteins showing large concerted structural transitions in parts of their structure, such as is for example often the

case for viral fusion proteins, and can be a powerful tool to study the chronology of structural intermediates in a complex refolding process (Benhaim and Lee, 2020, Benhaim et al., 2020).

For most proteins, however, the rate of this temporary unfolding and refolding is much faster than the deuterium uptake rate, so that the deuteration occurs gradually and only a single peak distribution can be detected (EX2 kinetics or uncorrelated exchange)(Figure 3 A). In this case, the exchange rate is the product of the equilibrium constant of the opening reaction ($K_{op}$) and the chemical exchange rate (Equation 5).

**EX2 kinetics** (uncorrelated exchange): $k_{cl} >> k_{ch}$

Equation 5 $$k_{HDX} = \frac{k_{op}}{k_{cl}} * k_{ch} = K_{op} * k_{ch}$$
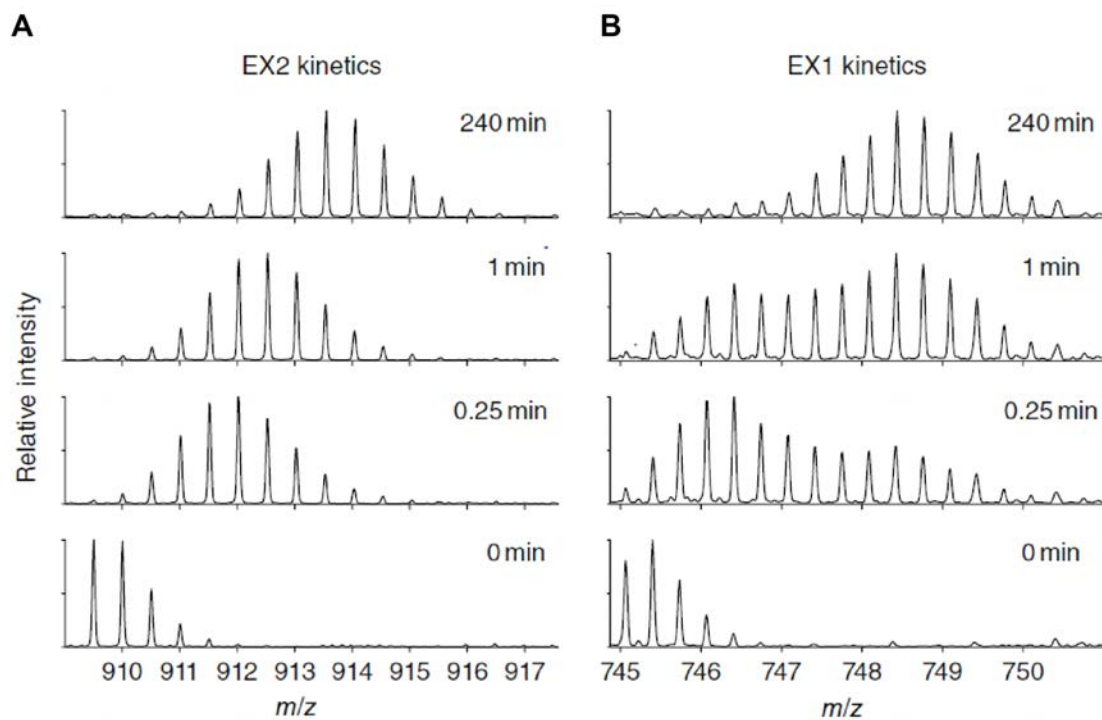


Figure 3: MS characteristics and data analysis of EX1 and EX2 kinetics. A 19-residue peptide displaying EX2 kinetics (A) and a 18-residue peptide displaying EX1 kinetics (B). EX2 kinetics are characterized by binomial isotope patterns with gradual $m/z$ centroid value shifts over deuteration time. On the other hand, EX1 kinetics cause bimodal isotope patterns, that can be either separated on the $m/z$ scale when a large number of residues participate in correlated unfolding, or are merged together when the amount of unfolded residues is low. In this case broad

bell-shaped peak distributions are visible and binomial fitting is required to extract the deuterium uptake of the individual populations. Reprinted from Jensen and Rand (2016) with permission from John Wiley and Sons.

EX1 kinetics can be difficult to characterize if the individual populations are less distinct (Weis et al., 2006b). In such cases, there is the possibility that the protein undergoes EX1 and EX2 kinetics at the same time (so-called "EXX" or "mixed exchange" kinetics (Xiao et al., 2005)). The distance between the individual peak envelopes is an indicator of how many amide hydrogens are involved in cooperative unfolding. The relative intensity of the peak envelopes is directly related to the fraction of the individual populations in solution (Engen and Smith, 2000). It has to be noted that bimodal peak distributions can also arise from different protein populations in solution. Multiple populations of the same protein that differ structurally will incorporate different amounts of deuterium and can be observed as bimodal distributions in the mass spectra (Marcsisin and Engen, 2010). Constant relative intensities of the individual peak distributions over time can be indicative of two distinct protein populations, while an intensity shift towards the higher deuterated species at later time points indicates EX1 kinetics (Guttman et al., 2013).

A detailed overview of EX1/2 kinetics and the factors that govern HDX of proteins in solution can be found in a number of reviews and book chapters (Jensen and Rand, 2016, Konermann et al., 2011, Weis et al., 2006b). Considering the range of monitored changes within proteins, HDX-MS is a versatile tool to gain insights into the dynamics and structural responses of protein backbones to perturbed solution conditions and ligand or protein binding (Chalmers et al., 2011, Engen, 2009). Furthermore, conformational dynamics at different pH and temperature can also be probed, even though this requires slight adjustments in the experimental procedure and data analysis (Bai et al., 1993, Engen, 2009).

## CONVENTIONAL BOTTOM-UP HDX-MS WORKFLOW

Depending on the research question, several types of experiments can be performed to provide information on different aspects of the protein structure and dynamics. In most cases, deuterium uptake of a protein incubated under a certain condition (e.g. pH, temperature or ligand presence) is compared with deuterium uptake of the protein under standard conditions at several time points (so-called differential approach with continuous labeling).

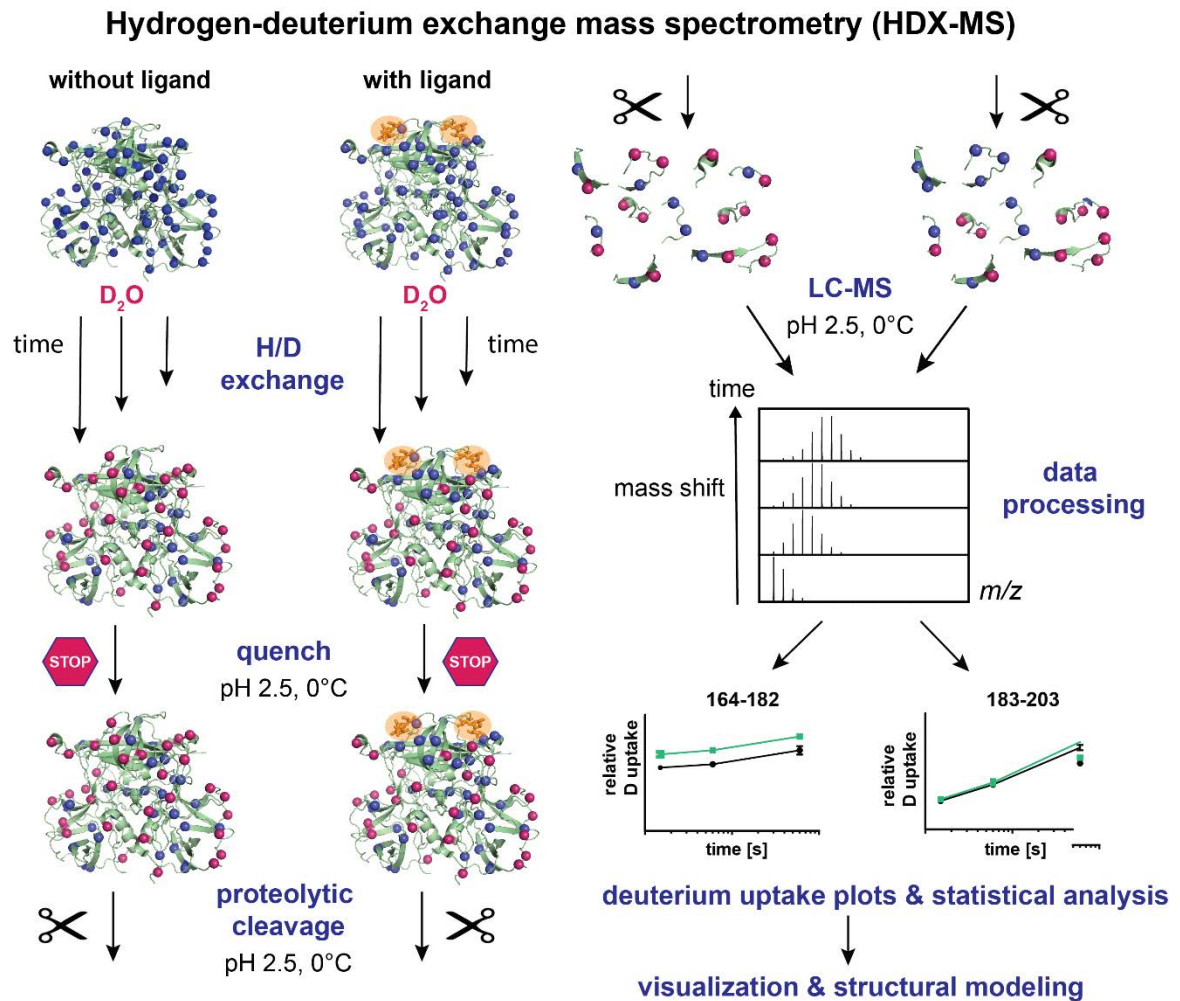**Hydrogen-deuterium exchange mass spectrometry (HDX-MS)**



Figure 4: HDX-MS workflow for a bottom-up continuous labeling experiment. After labeling for several time intervals, followed by a quench step, which reduces HDX to a minimum, the protein alone or with an interaction partner is proteolytically digested. The resulting peptides are separated and analyzed by LC-MS. From the observed mass shifts, deuterium uptake plots can be generated and peptides showing differences under the analyzed conditions can be identified. Adapted from Dülfer et al. (2019), Copyright (2019), with permission from Elsevier.

Using this approach, protein regions that undergo changes in dynamics and therefore change their hydrogen-bonding pattern and solvent accessibility, can be identified. Deuterium uptake differences can provide information about various aspects of the proteins' structural dynamics, such as conformational changes induced by ligand binding and location of binding epitopes, folding or unfolding depending on solution conditions, protein stability, protein-protein interactions and overall flexibility of different parts of the protein (Marcsisin and Engen, 2010, Morgan and Engen, 2009). Additionally, HDX can detect effects

more distant from the ligand binding site, for example allosteric interactions (Konermann et al., 2014). Another approach is so-called pulsed labeling, where the protein is first forced to undergo structural changes (e.g. by being exposed to abruptly changed pH, temperature or by addition of a perturbant) and is then labeled by a short 'pulse' of deuterium exposure. This approach is mainly used to identify and monitor transient folding intermediates (Konermann et al., 2011, Morgan and Engen, 2009).

The overall workflow of the bottom-up HDX experiment is depicted in Figure 4. The studied protein is labeled in a deuterated buffer under native conditions. After various time intervals, the labeling reaction is quenched by rapidly reducing the pH to 2.5 and the temperature to 0 °C, where the HDX rate and therefore also the rate of D/H back exchange, is minimal. Therefore, deuterium incorporation can be localized when the protein is split up into smaller peptides prior to MS analysis. These peptides can be generated via digestion with an aspartic protease, either in bulk solution or more conveniently using protease immobilized onto a flow-through column (Wang et al., 2002). As the low pH is required to keep deuterium back exchange loss minimal, porcine pepsin is routinely used as the protease of choice for HDX-MS applications. However, a number of alternatives are used, which can yield different peptides and more precise deuteration localization (Cravello et al., 2003, Kadek et al., 2014, Rey et al., 2009, Yang et al., 2015).
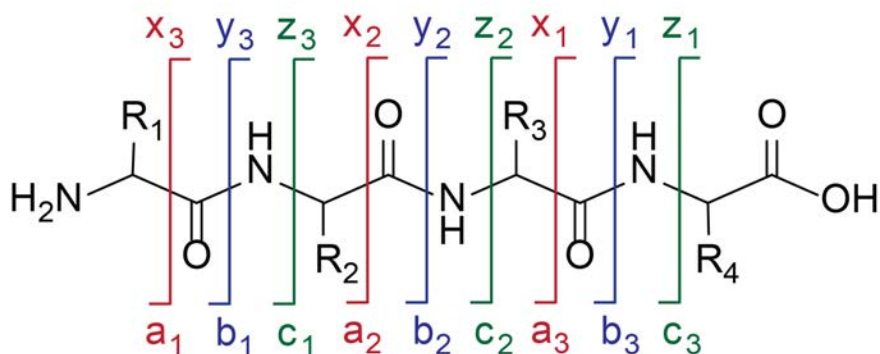


Figure 5: Peptide fragment nomenclature. Low energy fragmentation methods (such as CID) produce predominantly b and y ions by cleavage of the thermally labile peptide bonds. Losses of water or ammonia are also common in these fragments. Electron-based fragmentation methods (such as ETD) produce c and z ions resulting from cleavage of the N- $C_a$ bonds (Brodbelt, 2016).

Following digestion, peptides are separated in a liquid chromatography (LC) system and directly electrosprayed into a mass spectrometer. There, peptides are first identified by their $m/z$ and tandem MS (MS/MS) fragmentation patterns from non-deuterated samples. Based on the fragmentation method, different kinds of peptide fragment pairs are generated (Figure 5). Subsequently, deuterium content is measured for all peptides and compared between experimental conditions to localize any differences.

Besides the so-called bottom-up approach presented above, where the protein is cleaved by a protease, top-down approaches exist as well. There, the intact protein is sprayed in the mass spectrometer and fragmented in the gas phase (Leurs et al., 2015).

## METHODOLOGICAL ADVANCES TO TACKLE CHALLENGING SYSTEMS

The advantages of HDX-MS are a low sample consumption ($\approx$ 20- 100 pmol per deuteration time point) and accessibility of large protein complexes, which makes it possible to measure conformational dynamics and ligand binding of large virus-like particles (VLPs) or even intact viruses (Bereszczak et al., 2013, Lim et al., 2017, van de Waterbeemd et al., 2017). The protein digestion step can sometimes be a limitation in HDX experiments, as poor proteolytic digestion may result in low sequence coverage and therefore loss of information. Obtaining sufficient sequence coverage is often hampered by the occurrence of PTMs, such as glycosylation or disulfide bridges, adding to the spectral complexity or hindering efficient protease cleavage. However, this can in most cases be overcome by protocol optimization, e.g. addition of reducing or denaturing agents during the quench step or choice of a different protease (Trabjerg et al., 2018). In particular, the broad specificity of pepsin can be detrimental on this account as a large number of variably glycosylated peptides may be generated. A way to overcome this can be the implementation of a deglycosylation step before HDX (Guttman et al., 2011, Houde et al., 2009), or even after quenching (Jensen et al., 2016).

The use of an ultra-high-pressure liquid chromatography (UPLC) rather than a high-pressure liquid chromatography (HPLC) system increases chromatographic resolution and analysis speed and facilitates the investigation of more complex datasets (Engen and Wales, 2015). A conventional HDX-MS experiment usually provides resolution on the peptide level, even though more sophisticated fragmentation techniques, such as ETD or UVPD, can be used to bring resolution down to ultimately the single amino acid level (Mistarz et al., 2018, Rand, 2013).

## MEMBRANE PROTEINS

Membrane proteins play an important role in a variety of biological processes, but have so far been challenging to study by high-resolution biophysical techniques. This is mainly due to their hydrophobic character and strong interactions with the lipid bilayer, which usually results in rapid unfolding and aggregation once the protein is removed from the lipid environment. To perform membrane protein structure and dynamics studies, detergent micelles and phospholipid vesicles have been used as membrane mimetics. The high lipid content presents a challenge for LC-MS based workflows because lipids deteriorate LC performance and produce strong signals in MS. A number of groups have used HDX-MS to study the conformation and dynamics of proteins in detergent and lipid environments (Duc et al., 2015, Eisinger et al., 2017, Martens et al., 2018, Merkle et al., 2018, Moller et al., 2020, Reading et al., 2017).

To increase digestion and sequence coverage of membrane proteins, alternative proteases such as rhizopuspepsin and nepenthesin I and II are often used instead of or in addition to porcine pepsin (Moller et al., 2019). Another adjustment in the HDX-MS workflow is the implementation of a lipid depletion step with zirconium oxide ($ZrO_2$) beads (Hebling et al., 2010), which can also be implemented in automated workflows (Anderson et al., 2018). Given the high variety of membrane proteins and membrane mimetic systems, digestion conditions as well as LC workflows and MS parameters need to be tested and optimized for every protein of interest. Parameters for optimization and exemplary protocols for membrane protein HDX-MS in lipid and detergent environments can be found in various reviews and method papers (Martens and Politis, 2020, Martens et al., 2019, O'Brien et al., 2020).

## 1.2 Norovirus and the role of glycans in infection

Parts of the norovirus introduction have been published in the following review: *Pogan, R., Dülfer, J., & Uetrecht, C. (2018). Norovirus assembly and stability. Current opinion in virology, 31, 59-65.* Reprinted from Pogan et al. (2018a) under CC BY 4.0 license.

### 1.2.1 HUMAN NOROVIRUS

Approximately one-fifth of all acute gastroenteritis outbreaks are caused by human noroviruses (Ahmed et al., 2014). Acute gastroenteritis involves fever, vomiting, cramping and diarrhea (Kapikian et al., 1972). This illness usually persists for several days only but can take longer in children, the elderly and immunocompromised patients (Lindsay et al., 2015). Although the prototypical Norwalk virus was discovered over four decades ago, so far no vaccines and antiviral treatments are available (Kapikian et al., 1972).

Outbreak causing noroviruses emerge frequently and are seasonal especially in temperate regions (Ahmed et al., 2013). The dynamics of this diversity are poorly understood. Contributing factors like population immunity, virus evolution and transmission are mainly followed by epidemiological observations as a robust infection model is still missing. Since the virus is mostly foodborne and has to persist on surfaces and in the environment, particle stability at different environmental conditions is of interest. Data on human noroviruses are limited and are complemented with findings on murine, feline and bovine noroviruses.

### 1.2.2 GENETIC DIVERSITY

The genus Norovirus belongs to the *Caliciviridae* family, which is divided into seven genogroups (GI-GVII) based on the major capsid protein (VP1) sequence. Noroviruses infecting humans can be found in genogroup I, II and IV, while murine noroviruses (MNV) are GV (Vinje, 2015)(Figure 6).

Noroviruses of distinct prevalence emerge frequently and surveillance networks help to capture trends of respective genogroups and genotypes (Koopmans et al., 2003). While GI, which also includes the prototype Norwalk virus, is less frequently found nowadays, GII.4 isolates are known as the most common cause of clinical gastroenteritis cases (Eden et al., 2013). Surveillance studies also imply

that GII.17 variants increase in prevalence (de Graaf et al., 2015). Interestingly, genogroup and transmission routes correlate. While GII.4 strain transmission was often found to be person-to-person dependent, non GII.4 viruses like several GI isolates, GII.6 and GII.12 were mostly foodborne and other GI isolates waterborne (reviewed in de Graaf et al. (2016)).
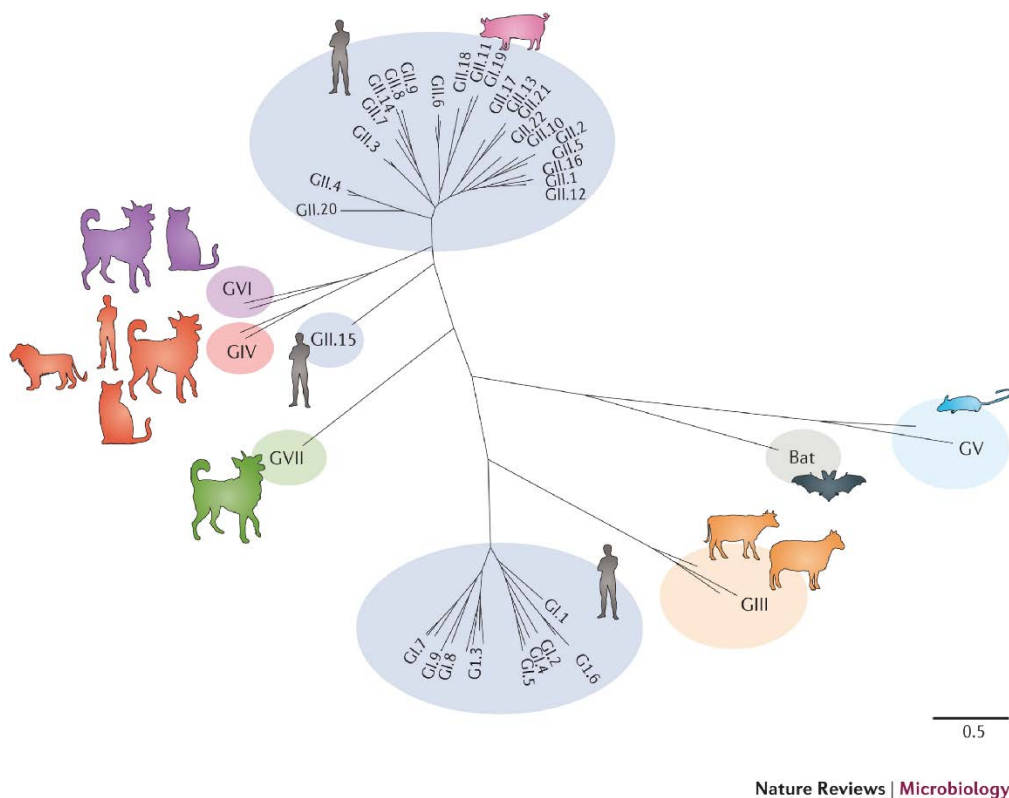


Figure 6: Norovirus classification by VP1 sequence. Genogroups are further subdivided into genotypes, which differ by at least 43 % (Zheng et al., 2006). The scale bar reflects the number of nucleotide substitutions per site. Reprinted by permission from Springer Nature, Nature Reviews Microbiology, de Graaf et al. (2016), Copyright (2016). The figure was originally adapted from Vinje (2015).

## 1.2.3  HUMAN NOROVIRUS VIRIONS AND SURROGATES

Many norovirus mediated gastroenteritis outbreaks originate in the consumption of fecal-contaminated food like mussels, oysters, berries and vegetables (Tuan Zainazor et al., 2010). Hence, these enteric pathogenic agents have to persist on various surfaces and in harsh environments. Studies on intact human noroviruses are hampered as no robust cell culture system is available. A cultivation in enterocytes has been established (Ettayebi et al., 2016), but still remains challenging. However, for infectivity assays on virions reliable cultivation

systems are inevitable and thus studies focus on norovirus surrogates. Proposed norovirus surrogates are amongst others feline calicivirus (FCV), which belongs to the genus Vesivirus. Due to its transmission by the respiratory route and its general pH instability, FCV is considered less suitable (Cannon et al., 2006, Fallahi and Mattison, 2011). MNV is often used as a surrogate because it shares many genetic and biochemical features with human norovirus, such as capsid size and shape and genome organization. Furthermore, it can infect cells in culture  as well as replicate in the gastrointestinal tract of its host (Wobus et al., 2004, Wobus et al., 2006). However, human and murine noroviruses have different cell specificities and MNV infection is usually asymptomatic (Wobus et al., 2006). MNV is more stable than FCV regarding thermal resistance and acid tolerance (Bozkurt et al., 2013, Cannon et al., 2006).

### 1.2.4 NOROVIRUS CAPSID STRUCTURE

To overcome the limitations of a lacking cultivation system, human norovirus studies focus on VLPs. The human norovirus capsid is icosahedral and comprises 180 copies of the capsid protein VP1 in $T = 3$ capsid symmetry (virus capsid symmetries and the triangulation number $T$ are described in more detail in Prasad and Schmid (2012)). Expression of VP1 in insect cells leads to self-assembly of VLPs of approximately 36-42 nm in diameter that are morphologically and antigenically comparable to native virions (Xi et al., 1992). X-ray crystallography of GI.1 Norwalk VLPs revealed the division of the capsid protein into a protruding (P) domain and a shell (S) domain (Figure 7). The P domain is further divided into P1 and P2 subdomains. Subdomain P2 is highly variable and, as it is exposed to the surface, involved in determination of antigenicity and cell attachment via binding of glycans (Baclayon et al., 2011, Choi et al., 2008, Prasad et al., 1999, Prasad et al., 1994). The capsid protein adopts three quasi-equivalent structures, namely A, B and C. The S domain of A/B and C/C dimers is arranged bent and flat, respectively. Switching between A/B and C/C dimers of the S domain is necessary to facilitate curvature whereas P domains build the capsid protrusions. The hinge domain that connects the P domain to the S domain provides a certain degree of flexibility to the P domain in a strain-dependent manner (reviewed in Smith and Smith (2019)). Icosahedral contacts between VP1 dimers are modulated by the S domain, while intra-dimer contacts are modulated by the P domain (Prasad et al., 1999).
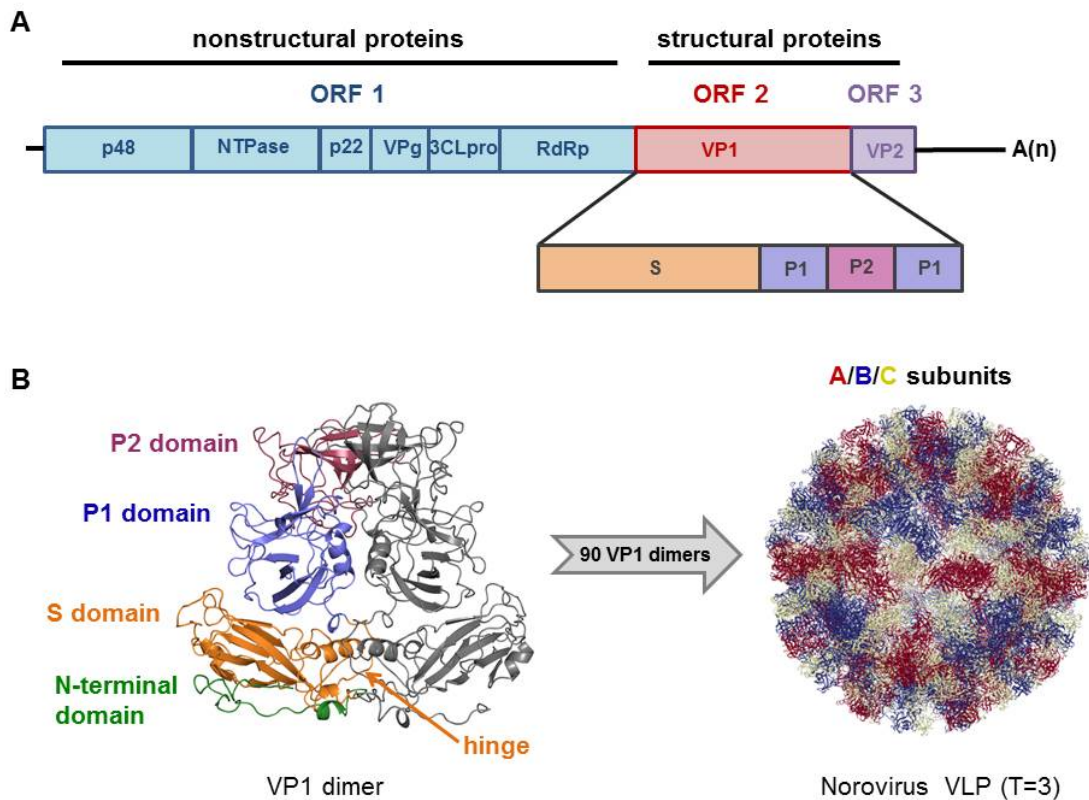
Figure 7: Genomic organization and structure of human noroviruses. (A) Noroviruses are RNA viruses with a positive sense single-stranded genome with three open reading frames (ORFs). A polyprotein, which includes the nonstructural proteins such as the RNA dependent polymerase (RdRp), is encoded by ORF1. Two structural proteins, the major and minor capsid protein (VP1 and VP2), are encoded by ORF2 and ORF3, respectively (Jiang et al., 1993, Xi et al., 1990) The VP1 protein is divided into shell (S) and protruding (P) domains. (B) 90 dimers of the major capsid protein VP1 assemble into icosahedral $T$ = 3 VLPs. The S domain of the VP1 monomers builds a shell that surrounds the viral RNA in form of a scaffold. The more flexible P domain is subdivided into P1 and P2 and connected to the S domain via a hinge. Subdomain P2 is highly variable and, as it is exposed to the surface, involved in determination of antigenicity and cell attachment (Baclayon et al., 2011, Choi et al., 2008, Mallagaray et al., 2015, Prasad et al., 1999, Prasad et al., 1994). The domains are highlighted in the VP1 dimer structure (left) and the three quasi-equivalent subunits (A/B/C) forming the capsids are shown in the VLP structure (right, pdb 1IHM) (Prasad et al., 1999). Reprinted from Pogan et al. (2018a) under CC BY 4.0 license.

Particles can be formed by assembly of the S domain only, showing that the S domain interactions are sufficient for capsid formation (Bertolotti-Ciarlet et al., 2003, Bertolotti-Ciarlet et al., 2002). Deletion of the P domain results in smaller, smooth particles ($T$ = 3, 29 nm) with reduced stiffness compared to wild type particles indicating that the P domain stabilizes the viral particles (Baclayon et al., 2011). Recombinant VLPs are assembled in absence of genome and therefore lack putative RNA-capsid interactions, which could alter capsid stability. These interactions have been shown to play important roles in assembly and stability of other RNA viruses (Dykeman et al., 2013, Snijder et al., 2013). Another factor likely contributing to capsid stability and assembly behavior is the presence of the minor capsid protein VP2 (Glass et al., 2000). Although the function of VP2 could not be deciphered completely, comparison of a GII norovirus VP1 only and VP1/VP2 particles indicate a slight stabilizing influence under alkaline conditions (Yao et al., 2014). Additionally, VP1/VP2 particles show decreased protease degradation (Bertolotti-Ciarlet et al., 2003). The amount of incorporated VP2 molecules is still unknown.

## 1.2.5 P DIMER STRUCTURE CONSERVATION IN DIFFERENT NOROVIRUS STRAINS

The P domain adopts a characteristic and well-conserved topology. The P2 domain is formed by a large insertion of ≈ 126 residues into the P1 domain and folds into a compact barrel-like structure of six β-strands. The P1 domain consists of several β-strands and one α-helix. In the capsid, the P domain is involved in dimeric contacts that are also formed in the P dimer. (Choi et al., 2008, Prasad et al., 1999). P domain sequence alignment of the four human norovirus strains used in this work (GII.10 Vietnam, GII.17 Kawasaki, GII.4 MI001 and GII.4 Saga) resulted in 92 % (GII.4 MI001), 59 % (GII.17 Kawasaki) and 56 % (GII.10 Vietnam) sequence identity to the GII.4 Saga P domain as reference (Figure 8A). On the structural level, the P dimer structure is well conserved, with biggest differences in loop regions of the P2 domain (Figure 8B). Even though most structural variations can be found in this region, glycan binding pockets on the top of the P2 domain are also highly conserved among all strains. P domains of the murine norovirus strains GV MNV07 and CR10 are 34 % and 32 % identical to the human norovirus GII.4 Saga P domain, respectively (Figure 8C). Despite the large difference on the amino acid level, murine norovirus P domains show tertiary

structures similar to those of human norovirus, with two antiparallel β-sheets and one α-helix in P1 and a β-barrel of six antiparallel β-strands in P2 (Figure 8D). The greatest structural variations can be found in the loop regions on the outer surface of the P domain (Taube et al., 2010).
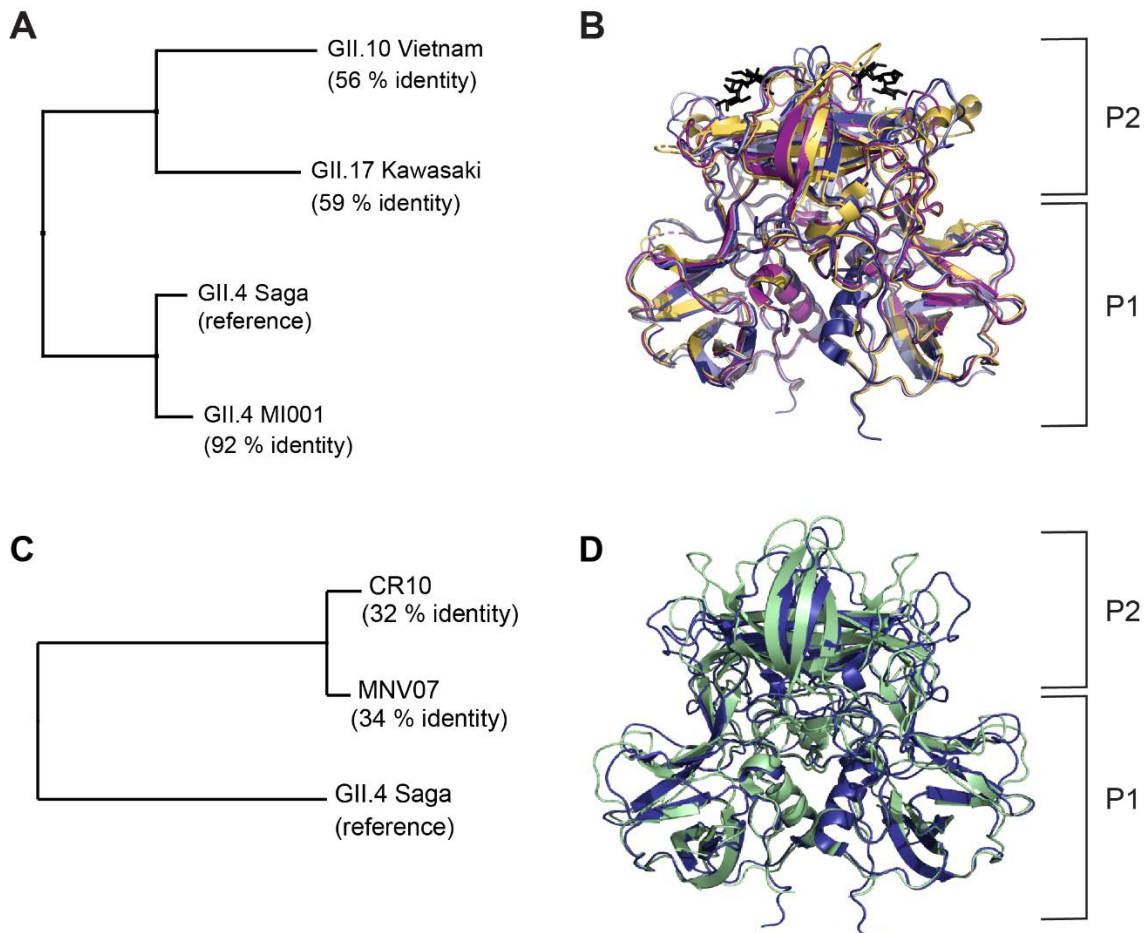


Figure 8: Comparison of human norovirus P domain sequences (A) and crystal structures with HBGA B trisaccharide (black) bound to the P2 domain (B). P dimer sequences of three virus strains were aligned to the GII.4 Saga strain as reference. For the GII.4 MI001 P dimer, a homology model was created for comparison using the GII.4 Farmington Hills P dimer structure with 94% sequence identity to MI001 as reference. Crystal structures of GII.4 Saga (pdb 4X06, dark blue), GII.10 Vietnam (pdb 3ONY, yellow), GII.17 Kawasaki (pdb 5F4O, purple) and the homology modelled structure of GII.4 MI001 (light blue) were superimposed in PyMOL. (C) Comparison of human norovirus GII.4 Saga P domain sequence with two murine P domain sequences used in this work. (D) Crystal structures of GII.4 Saga (pdb 4X06, dark blue) and GV CR10 (pdb 6H6M, light green) P dimer structures were aligned in PyMOL. P1 and P2 indicate the respective domains of the P dimer.

## 1.2.6    CELL ATTACHMENT AND ENTRY

Norovirus attachment to glycans is an essential factor in host cell infection. Human noroviruses interact with Lewis and ABH Histo blood group antigens (HBGAs, Figure 9A) on glycoproteins or glycolipids presented on host cell surfaces. Whiles these interactions are an important factor for infection, the exact mode how binding to glycans promotes host cell entry is not known. Furthermore, presentation of HBGAs on cell surfaces alone is not sufficient for infection (Taube et al., 2018). The discovery of an entry receptor for murine norovirus (Haga et al., 2016, Orchard et al., 2016) as well as the modulating role of bile acids in infectivity (Creutznacher et al., 2019, Kilic et al., 2019, Nelson et al., 2018), points towards a more complex interplay of human norovirus with host cell factors (Figure 9B).

Even though carbohydrate-protein affinities are usually low, with dissociation constants ($K_d$) in the mM range, noroviruses have high strain-dependent specificities to glycans. Human norovirus predominantly interacts with HBGAs, but whether glycan binding plays a role in MNV infection is still unclear (Creutznacher et al., submitted, Taube et al., 2012, Wegener et al., 2017). Structures of Lewis antigens and HBGAs of different types are highly similar and all share an L-fucose residue, which has been shown to be the minimal structural element responsible for binding to the majority of human noroviruses (Taube et al., 2018).

Evidence of glycan interactions with human noroviruses mainly come from biochemical and biophysical techniques on the level of VLPs or P dimers. A number of NMR (Mallagaray et al., 2015, Mallagaray et al., 2016, Wegener et al., 2017) and crystallography studies (Cao et al., 2007, Hansman et al., 2011, Koromyslova et al., 2015, Schroten et al., 2016, Shanker et al., 2011, Singh et al., 2015b, Weichert et al., 2016) confirm the binding of α-L-fucose and HBGAs to VLPs and P dimers of different strains in conserved binding pockets on the top of the P2 domain (Figure 9C). In GII.4 Saga P dimers L-fucose interacts via six hydrogen bonds to either the backbone (BB) or side chain (SC) of five amino acids: Asp374 (2xSC), Arg 375 (2xSC), Thr344 (BB), Gly443 (BB) and Tyr444 (hydrophobic interaction). In GII.10 Vietnam two additional fucose binding sites exist in the P2 β-sheet cleft (Koromyslova et al., 2015)(Figure 9C and D).
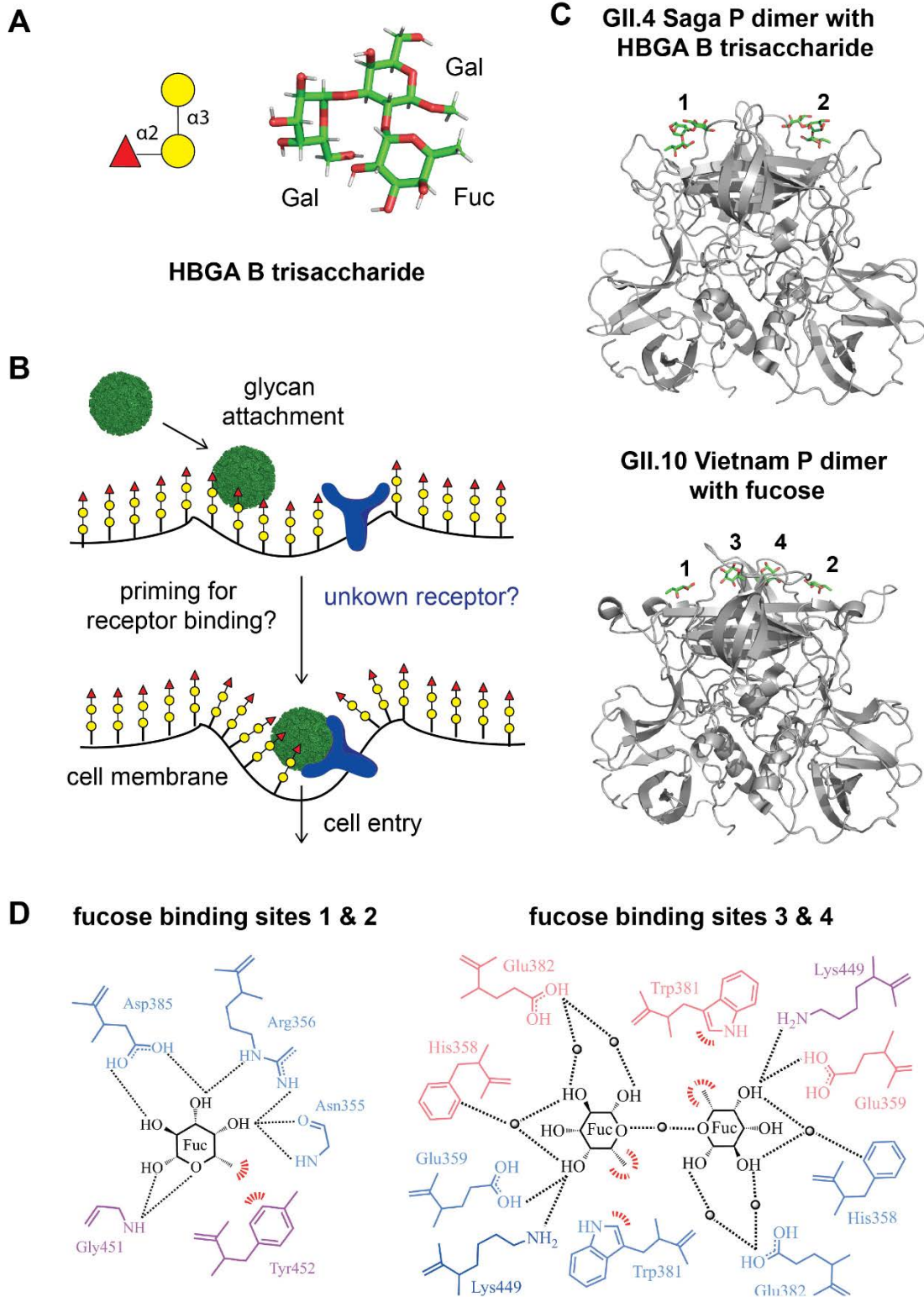
Figure 9: Human norovirus-glycan interactions. (A) Structural and symbolic representation of human HBGA B trisaccharide. The structure has been modeled as minimum energy conformation using the Glycam web server (Woods). (B) Model of glycan-mediated norovirus attachment and entry. The norovirus attaches to glycans displayed on host cell surfaces which hypothetically results

in structural changes that prime the virus for binding of a (so far unknown) receptor. Receptor binding and membrane invaginations then facilitate cell entry. (C) P dimer crystal structures of different strains show the two canonical glycan binding sites (1, 2) on the top of the P2 domain. Two additional binding sites (3, 4) have been reported for fucose in the GII.10 Vietnam strain (pdb 4X06, 4Z4R). (D) Molecular interactions of fucose via hydrogen bonds (black dashes) and hydrophobic interactions (red dashes) with P domain amino acids in GII.10 Vietnam. Panel D is reprinted from Koromyslova et al. (2015), Copyright (2015), with permission from Elsevier.

GII.4 norovirus VLPs induce membrane invaginations on giant unilamellar vesicles after binding to ABH glycosphingolipids (Rydell et al., 2013). A follow-up study showed that the mode of glycan presentation strongly influences their recognition (Nasir et al., 2017). This highlights the importance of more complex systems for studying human norovirus-glycan interactions on a native-like level. As an efficient cell culture system and animal model for human norovirus is still lacking, murine norovirus has been broadly used as a model to unravel general mechanisms of norovirus infection. These studies enabled the discovery of murine CD300If and CD300Id as functional protein receptors for murine norovirus (Haga et al., 2016, Orchard et al., 2016), which are required and sufficient for infection. Structural studies revealed that the CD300If receptor binds to a site in close proximity to the glycan binding site in human noroviruses (Kilic et al., 2018, Nelson et al., 2018)(Figure 10A), but while the receptor binding site is conserved in murine noroviruses, this region is very variable in human noroviruses which points against a genotype-wide receptor binding at this site (Kilic et al., 2018). It was also suggested that each virion binds multiple CD300If receptor molecules, which points towards an avidity-driven process as suggested for attachment of human noroviruses to cell surface clustered HBGAs (Nelson et al., 2018). In FCV, receptor binding to the VP1 protein induces large structural changes in the capsid, that leads to the formation of a portal-like assembly of VP2 molecules and opening of a pore in the capsid shell (Conley et al., 2019)(Figure 10D).

Recently, it was shown that bile acids can act as co-factors by enhancing cell-binding and increasing infectivity in MNV (Nelson et al., 2018) and human norovirus VLPs and P domains (Kilic et al., 2019). Binding of bile acids to a high-affinity binding site near the HBGA binding pocket switched critical loop positions to turn a human HBGA non-binder strain into a HBGA binder (Kilic et al., 2019)(Figure 10B).
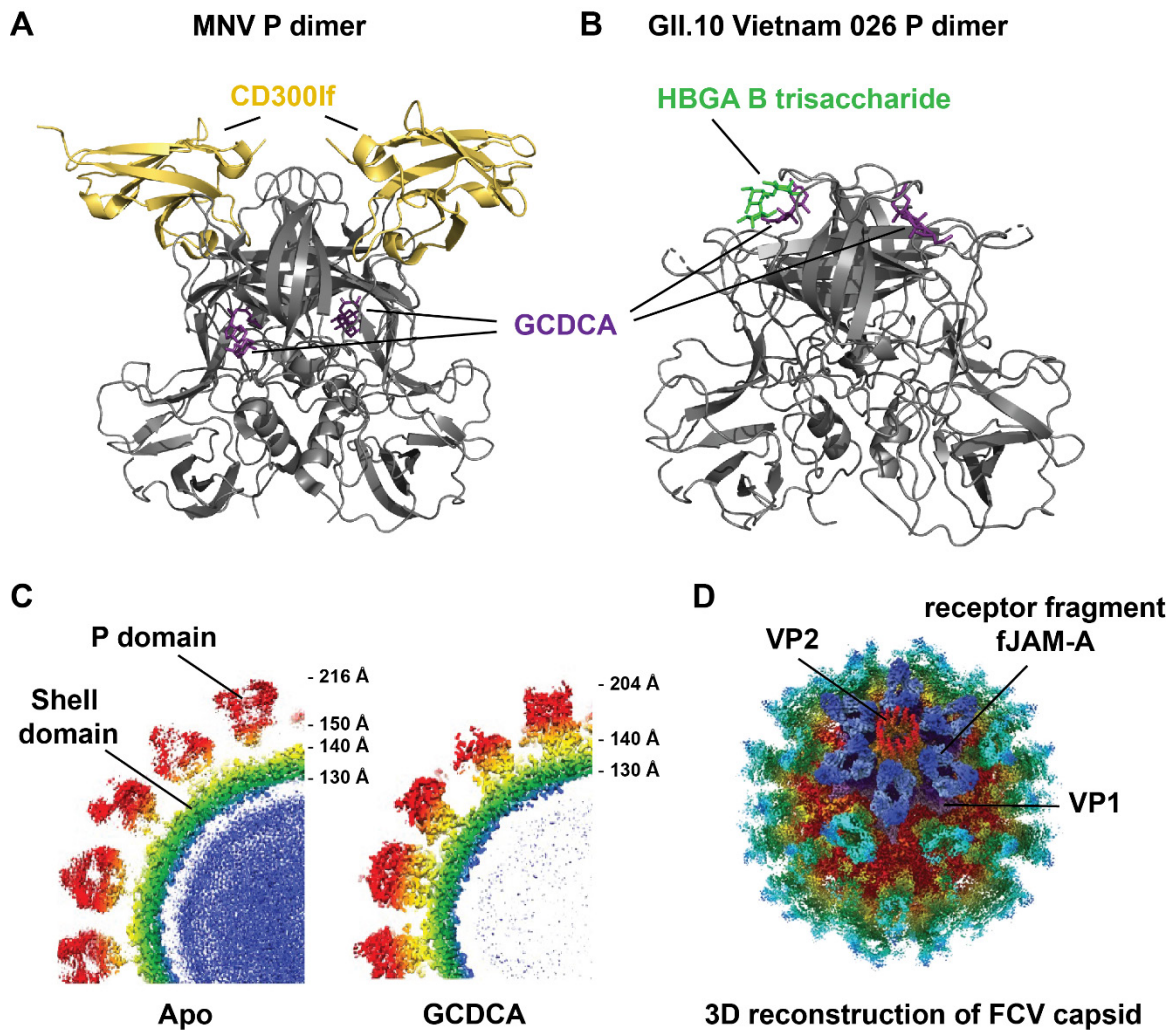
Figure 10: (A) Binding of the CD300If receptor and glycochenodeoxycholic acid (GCDCA) to the MNV P dimer (pdb 6E47). (B) GCDCA binds near the HBGA binding site in human GII.10 Vietnam P dimer (superposition of pdb 6GW1 and 3Q38 in PyMOL). (C) Cryo-EM image reconstruction of MNV-1 in absence and presence of GCDCA. In the apo structure the P domain "floats" above the shell domain, whereas it collapses onto the shell when GCDCA is bound. (D) Cryo-EM structure of the FCV capsid bound to feline junctional adhesion molecule A (fJAM-A), a fragment of its cellular receptor. One receptor molecule (blue and green) binds to each VP1 (pink and purple). Twelve copies of the VP2 protein (orange and red) form a portal-like assembly. Panel C is adapted from Sherman et al. (2019) with permission from the American Society of Microbiology. Panel D is adapted from Springer Nature, Nature, Conley et al. (2019), Copyright 2019.

In contrast, binding of bile acids to another low-affinity binding site identified by NMR did not significantly influence HBGA binding of VLPs and P dimers of different human strains (Creutznacher et al., 2019). In MNV, binding of bile acids causes a conformational change that contracts the P domain onto the shell

surface which optimizes receptor avidity (Sherman et al., 2020, Sherman et al., 2019)(Figure 10C).

Taken together, structural biology approaches have contributed to the understanding of human norovirus-glycan interactions and infection studies with murine noroviruses urged speculations about the possible existence of a human norovirus receptor. However, further studies are required to elucidate the dynamic structural changes during the attachment and entry process that so far cannot be captured by crystallography. In this regard, HDX-MS is a valuable tool to investigate norovirus structural dynamics following interaction with different entry co-factors and can contribute to the understanding of the attachment and entry process from a structural perspective. Undoubtedly, further advancements in the implementation of a robust cell culture system for human noroviruses are required to link the observed structural features to infectivity *in vivo*.

## 1.3  Mistic – an unconventional membrane protein

Mistic is a small 13 kDa integral membrane protein from *Bacillus subtilis*. It comprises 110 amino acid residues that are arranged into a bundle of four α-helices with a central kink in helix 2. Its structure has so far been determined by solution phase NMR spectroscopy in the presence of lauryldimethylamine *N*-oxide (LDAO) detergent (Roosild et al., 2005)(Figure 11A). It is thought to fold autonomously into the membrane, which appears to be in conflict with its polar and uncommonly hydrophilic surface (Figure 11B).

Despite comprising the α-helical structure of a canonical integral membrane protein (Roosild et al., 2005), its intrinsic features render a transmembrane orientation unlikely and are thought to be linked to its biological function. Mistic has been shown to serve as a chaperone for the high-level native-fold production and integration of other membrane proteins into the bacterial lipid bilayer (Kefala et al., 2007, Roosild et al., 2005, Roosild et al., 2006), as reflected by the acronym "**M**embrane-**i**ntegrating **s**equence for **t**ranslation of **i**ntegral membrane protein **c**onstructs". Together with the potassium channel YugO located downstream of Mistic in the gene sequence of *B. subtillis*, Mistic plays a role in biofilm formation (Lundberg et al., 2013). The exact functional connection of these two proteins still remains unclear.
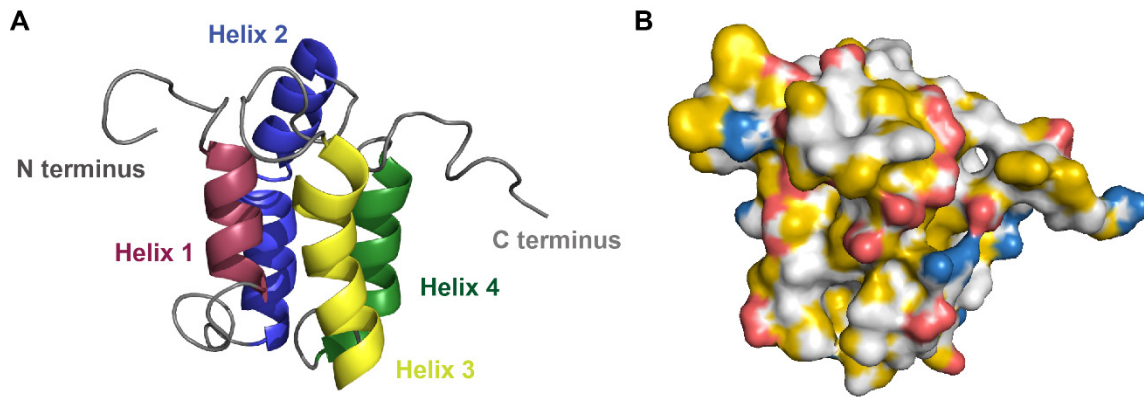
Figure 11: Mistic structure. Structure of the lowest energy conformer solved by solution phase NMR spectroscopy in LDAO detergent micelles (pdb 1YGM). (B) Surface representation of Mistic in the same orientation as in A, with colored hydrophobic (yellow), negatively charged (red) and positively charged (blue) amino acid side chains (Hagemans et al., 2015). The remaining atoms and the backbone are colored in grey.

So far Mistic biophysical properties and (un-)folding behavior have been studied by a variety of biochemical and biophysical techniques, e.g. circular dichroism (CD) spectroscopy, isothermal titration calorimetry, dynamic light scattering (DLS), NMR and different fluorescence spectroscopy approaches amongst others (Broecker et al., 2014, Debnath et al., 2011, Dvir et al., 2009, Frotscher et al., 2018, Hartmann et al., 2015, Jacso et al., 2013, Roosild et al., 2005, Textor, 2016). Mistic dynamics and topology have mainly been investigated in molecular dynamics (MD) simulations (Psachoulia et al., 2006, Sansom et al., 2008). All these studies have in common that they mainly examined Mistic in micelles of various detergents, such as LDAO, n-dodecyl phosphocholine (DPC) and n-dodecyl-β-D-maltoside (DDM), but not in more native-like lipid environments.

In a recent PhD thesis (Textor, 2016) a quantitative reconstitution protocol for Mistic in large unilamellar vesicles (LUVs) was established and the protein orientation in the lipid bilayer was investigated by various biophysical techniques, suggesting a rather in-plane topology in the membrane. Recent studies using single molecule fluorescence resonance energy transfer (smFRET) suggest that lipid composition in the bilayer has an effect on Mistic conformation (Krainer et al., in preparation).

## 2 AIM AND OBJECTIVE

Viral infections are initiated through the recognition of specific cell surface receptors. These receptors often contain glycan structures, which are attached to either proteins or lipids. Different types of highly diverse glycan structures have been found to serve as viral attachment factors and/or entry receptors (Ströh and Stehle, 2014). For many viruses specific knowledge on interactions between viral proteins and cellular protein receptors is available and has enabled detailed models of conformational changes during virus-receptor interaction (Chen et al., 2005, Persson et al., 2007, Stehle and Casasnovas, 2009).

In contrast, the contributions of protein-glycan interactions in viral attachment and entry have so far hardly been clarified, even though attachment to specific glycans is a crucial step that determines susceptibility and cell tropism. One reason for this is that glycans on cell surfaces form complex heterogeneous mixtures and different types of these glycan structures, including sialic acids, glycosaminoglycans and histo-blood group antigens, or even combinations of them, can act as attachment factors. Furthermore, interactions between glycans and viral proteins have a relatively low affinity, which makes it difficult to measure these interactions with standard technologies (Kang et al., 2016). For noroviruses, X-ray crystallography and NMR have been used to characterize binding sites for HBGAs and other glycans on VLPs and P dimers of different strains. However, it is important to understand glycan-induced changes in protein dynamics to uncover their role in the cell entry process.

To overcome this limitation the objective of this project was to implement an HDX-MS workflow to identify glycan binding sites on norovirus P dimers and track strain-specific structural changes upon glycan binding. These changes could hypothetically prime for cellular uptake, make additional binding sites accessible or modify capsid stability. Time resolved HDX-MS data on structural changes was combined with crystallography and NMR performed by collaborators to gather a more detailed understanding of norovirus-glycan interactions. In addition, native MS was used to confirm the dimeric state of P domains of different norovirus strains and retrieve information about the monomer-dimer equilibrium in human and murine P domains.

After having implemented an HDX-MS workflow for soluble proteins, the next step was to adapt this workflow for the measurement of protein-membrane interactions. Apart from the huge variety of membrane proteins that can be

targeted with such an adapted workflow, the investigation of norovirus binding to membrane-attached glycans will be of special interest in the future. As norovirus VLPs have been shown to form membrane invaginations (Parveen et al., 2018, Rydell et al., 2013) after attachment to glycolipids, it would be of interest to map the occurring structural changes in the virus capsid under these conditions by HDX-MS. More detailed information about this approach can be found in the Outlook section.

The protein Mistic has been studied extensively by the group of Prof. Michael Schlierf and collaborators over the last years, giving insights about its folding and unfolding trajectories in different detergents (Frotscher et al., 2018, Hartmann et al., 2015, Krainer et al., 2018). Recently, the focus moved to studies of Mistic conformation and dynamics in lipid vesicles which makes it a good model protein for the systematic screening and adaptation of HDX-MS conditions for membrane-associated proteins. At the same time, HDX-MS can contribute valuable information about Mistic protein dynamics in different lipid environments, which complements existing data.

Therefore, a collaboration project was set up for screening of optimal HDX-MS conditions as well as measurement of Mistic dynamics in lipid vesicles of different compositions and in detergent micelles. This included tests of quench buffer and protease digestion conditions and the implementation of a lipid depletion step in the HDX-MS workflow.

# 3 RESULTS AND DISCUSSION

Parts of the following results and discussion chapters have been published or are in preparation for publication in:

*Mallagaray, A., Creutznacher, R., Dülfer, J., Mayer, P. H. O., Grimm, L. L., Orduna, J. M., Trabjerg, E., Stehle, T., Rand, K. D., Blaum, B. S., Uetrecht, C., Peters, T. (2019). A post-translational modification of human Norovirus capsid protein attenuates glycan binding. Nature communications, 10(1), 1-14. (Reprinted under CC BY 4.0 license.)*

*Dülfer, J., Yan, H., Brodmerkel, M., Creutznacher, R., Mallagaray, A., Peters, T., Marklund, E., Uetrecht, C.: Glycan-induced structural dynamics in human norovirus P domains depend on virus strain and deamidation status. (in preparation)*

*Creutznacher, R., Feldmann, C., Dülfer, J., Maass, T., Knickmann, J., Uetrecht, C., Peters, T. & Taube, S.: Murine norovirus capsid plasticity probed by NMR - Bile acid binding stabilizes and rearranges P domain dimers and triggers immune escape. (in preparation)*

## 3.1  Deamidation influences glycan binding in norovirus GII.4 P dimers

P dimers of the GII.4 Saga strain were chosen as the first target for the newly implemented HDX-MS workflow. Infections with GII.4 strains are highly prevalent in humans and therefore information about capsid protein-glycan interactions is of great interest to understand the infection process in humans. In parallel, NMR and crystallography experiments with the same samples were conducted, so that complimentary data is available. The present sub-chapter presents findings from HDX-MS, NMR and crystallography, but focuses on MS data that was acquired by me. All results discussed in this subchapter are published in Mallagaray et al. (2019).

### 3.1.1  IDENTIFICATION OF N373 DEAMIDATION

Before HDX-MS experiments can be conducted, peptides generated by peptic cleavage need to be identified, a process also called "peptide mapping". To do this, the unlabeled protein is injected into a LC system for on-column peptide cleavage, the resulting peptides are separated on a reverse-phase column and fragmented

in the mass spectrometer. Individual fragmentation patterns can then be mapped to peptide sequences at specific retention times by computational approaches and protein sequence coverage can be visualized in sequence coverage maps. Sequence coverage maps with all peptides used in each HDX-MS experiment (effective coverage) can be found in the supplement (Figure S 27 - Figure S 33). This approach also allows identification of PTMs, such as for instance *N*-terminal acetylation, oxidation, deamidation or disulfide bonds. If a sufficient number of peptide fragment ions is generated, the exact amino acid that carries the PTM can be mapped with high precision.

For the GII.4 Saga P dimer, a deamidation of N373 into D373 could be identified by MS/MS, after NMR spectra had also shown the conversion of N373 into iso-aspartate over time. N373 is located directly next to the glycan-interacting D374 residue; therefore, it was suspected that deamidation could influence glycan recognition. The reaction mechanism for the conversion of N373 into isoD373 (iD373) is depicted in Figure 12.
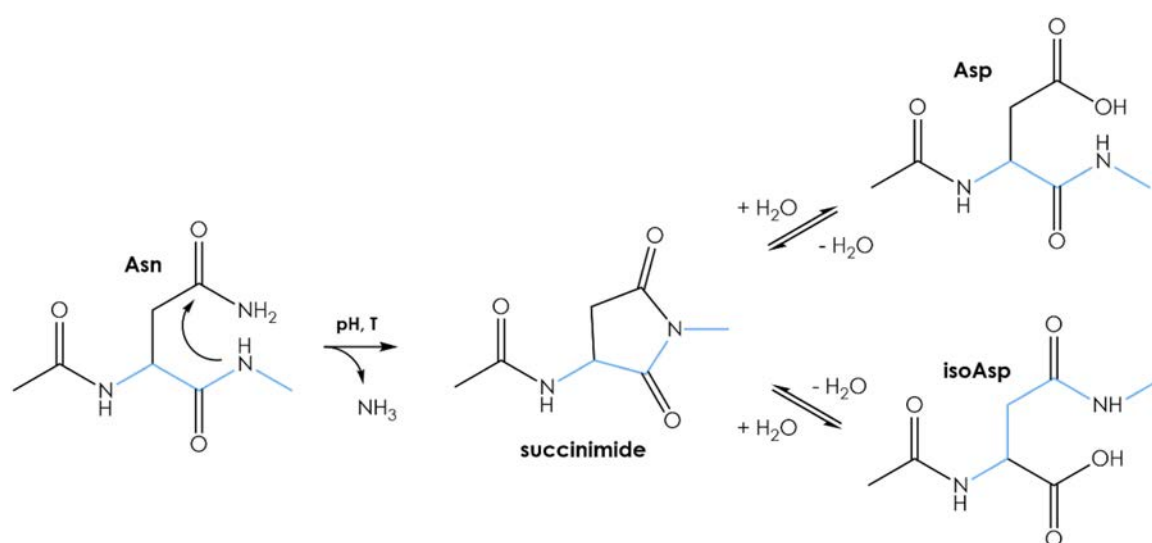


Figure 12: Formation of Asp and isoAsp following Asn deamidation. Nucleophilic attack of the C-terminal backbone NH leads to deamidation of Asn and formation of a succinimide as intermediate. Subsequent hydrolyzation of the succinimide ring creates a mixture of Asp and isoAsp, with isoAsp being the main species with a ratio of about 3:1 (Geiger and Clarke, 1987) and L-isoAsp being the main isomer (Riggs et al., 2017). Basic pH as well as increased temperature accelerates deamidation (Tyler-Cross and Schirch, 1991). Blue bonds highlight the peptide backbone.

Even though deamidation of Asn can result in formation of Asp or isoAsp, NMR spectra do not indicate the presence of Asp, so that formation of isoAsp seems to be significantly preferred. As deamidation results in an additional negative charge on the P domain, it was possible to separate fully native (NN), mixed (iDN) and fully deamidated (iDiD) P dimers by ion exchange (IEX) chromatography and quantify their fractions. Freshly expressed GII.4 Saga P dimers contained the isoAsp species in various amounts, which increased further over time (Figure S3). The rate of conversion into isoAsp seems to be strongly dependent on temperature. At room temperature and sample storage temperature (5°C) the half-life of deamidation is in the range of several days to weeks. In contrast, at body temperature (37°C), it is much faster with an estimated half-life of 1.6 days (Figure S 4).

The IEX separated native (NN) and deamidated (iDiD) GII.4 Saga P dimers were also subjected to peptic cleavage followed by LC separation and peptide identification using MS/MS. Peptides containing the potential deamidation site showed 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 (Figure 13). Analysis of peptides covering the entire protein sequence revealed no other deamidation sites apart from N373, even though the protein contains 17 asparagine residues (Table 1).

Table 1: Identification of deamidation sites GII.4 Saga P dimers. Most likely deamidation position (bold) and probability for deamidation at this position were calculated in MaxQuant based on peptide fragment spectra. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.

| Peptide sequence (charge) | Theoretical *m/z* (wt) | Observed *m/z* | Probability for D373 | deamidated fraction (%) |
|---|---|---|---|---|
| STDTE**N$^{373}$**DFETHQ (z=2) | 712.284 | 712.776 | 1 | 100 |
| STDTE**N$^{373}$**DFETHQNTKFPVG (z=3) | 756.673 | 757.001 | 0.708 | 96.6 |
| STDTE**N$^{373}$**DFETHQNTKFPVG (z=2) | 1134.506 | 1134.998 | 0.998 | 98.1 |
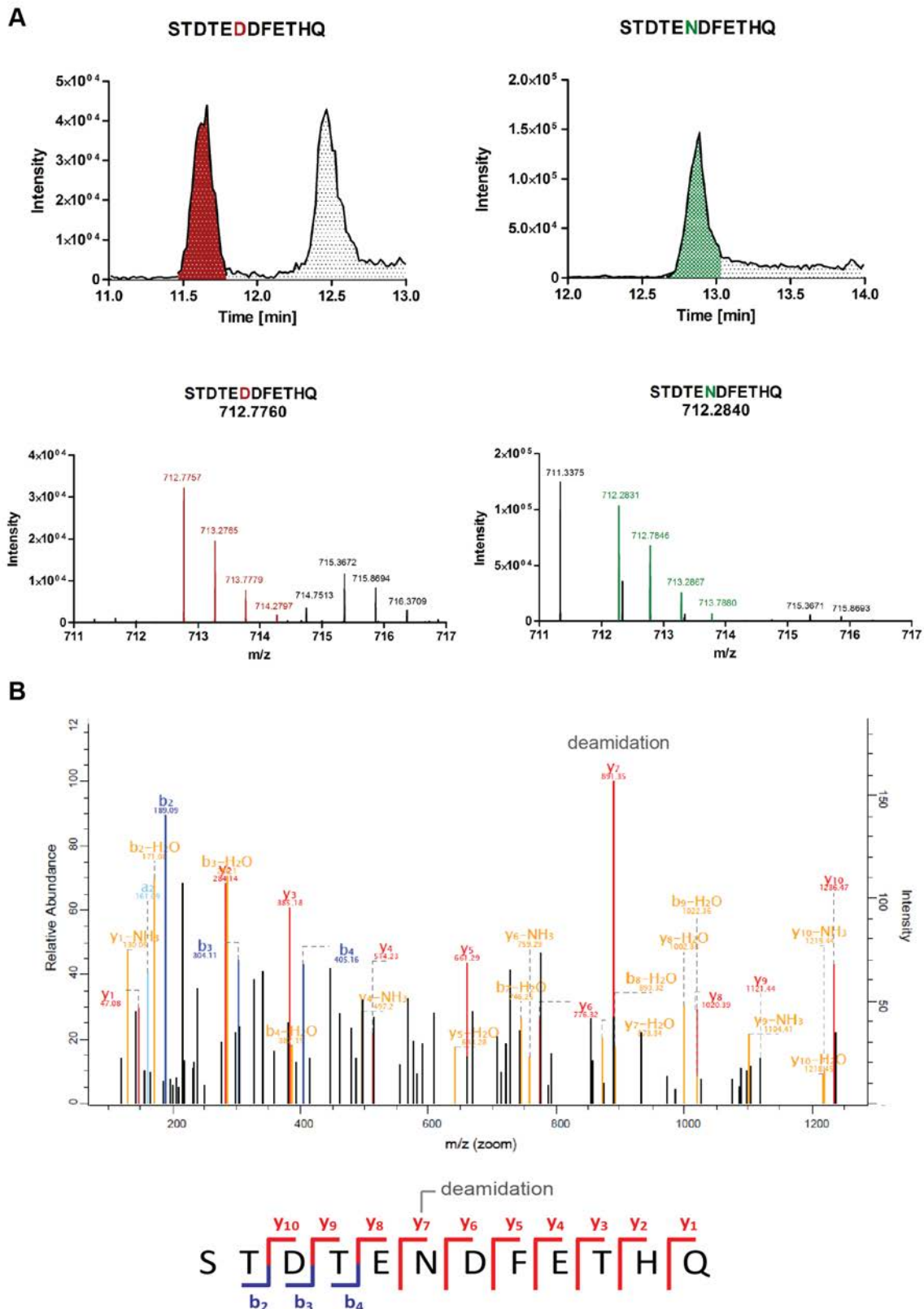
Figure 13: (A) Chromatogram and spectrum of peptide STDTENDFETHQ ($z$=2) in both the deamidated (red) and wild type (green) protein sample. The theoretical monoisotopic mass is indicated below the peptide sequence in the lower panel. (B) Fragmentation spectrum of peptide STDTENDFETHQ in the deamidated

sample. The fragment y7 (and all upstream y-fragments) shows a characteristic mass shift, which clearly indicates the presence of the deamidation at N373. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.

P dimers of three other norovirus strains, GII.4 MI001 (Taube et al., 2013), GII.10 Vietnam 026 (Hansman et al., 2011), and GII.17 Kawasaki 308 (Koromyslova et al., 2017) were also measured to test whether deamidation of N373 occurs in other strains as well. For GII.4 MI001 P dimers no crystal structure is available, but the P domain exhibits 92 % sequence identity with GII.4 Saga. Furthermore, the GII.4 MI001 loop region (STDTS**N**D374) is almost identical with the corresponding sequence in GII.4 Saga (STDTE**N**D374). Therefore it was suggested, that N373 deamidation could occur as well. The other two strains have no asparagine at the N373 equivalent position, but the overall loop conformations are rather similar to the one in GII.4 Saga. The sequences in that loop region are STWET**Q**D385 (GII.10 Vietnam 026) and RISDN**D**D378 (GII.17 Kawasaki 308). For GII.10 Vietnam 026 a glutamine precedes D385. Glutamine deamidation has also been described but is much slower (Geiger and Clarke, 1987). To test for deamidation, samples were incubated at elevated temperatures and IEX chromatograms were obtained before and after incubation. As hypothesized, deamidation was observed for GII.4 MI001 but not for the GII.17 Kawasaki 308 or GII.10 Vietnam 026 strains (Figure S 5). In MS/MS, deamidation of N373 was detected for GII.4 MI001 along with minor deamidation of another asparagine residue (either N446 or N448) after incubation at elevated temperatures. The other two strains showed no indication of deamidation in the entire protein confirming the results from IEX chromatography (Figure S 6 and Table S 1).

## 3.1.2 EFFECT OF DEAMIDATION ON GLYCAN BINDING AND PROTEIN DYNAMICS

Due to the direct neighborhood of N373 to D374, which takes part in direct interactions with glycan ligands, a possible influence of deamidation on glycan recognition was tested by HDX-MS, NMR and X-ray crystallography with methyl α-L-fucopyranoside and blood group B trisaccharide as representative HBGA ligand. HDX-MS and NMR were applied to provide information about the structural response to ligand binding and changes in protein dynamics, while in crystallography binding geometry and contributing amino acids were of interest.

In all setups apart from crystallography, IEX separated fully native (NN) P dimers were compared to fully deamidated (iDiD) P dimers.

## DEAMIDATION INCREASES DYNAMICS IN THE P2 DOMAIN

Notably, the deamidated P dimer shows a higher H/D exchange in peptides covering large parts of the top P2 subdomain irrespective of the presence of glycans. This indicates higher dynamics of respective regions of the P2 subdomain in the deamidated protein (Figure 14A). However, this difference was not detected in the peptide containing the deamidation site.

A comparison of NMR spectra of native and deamidated P dimers (Figure S 7) also shows a shift of some peaks in the deamidated protein relative to the native protein. The largest chemical shift perturbations (CSPs) are found in the vicinity of the site of deamidation 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.

## DEAMIDATION ATTENUATES GLYCAN BINDING

In the presence of 10 mM B trisaccharide the wild type protein is significantly protected from H/D exchange in peptide 434-449 along the canonical binding site (G443, Y444), indicating occupation of the binding pocket. In contrast, binding is highly reduced in the deamidated protein showing essentially no protection (Figure 14B). In NMR, CSPs are observed for iDiD P dimers demonstrating B trisaccharide binding, but a comparison to CSPs for NN P dimers reveals significant differences (Figure S 8). 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. In addition, protection of some residues of a β-sheet in the center of the binding cleft (Y286-L303) can be detected in HDX-MS. This region contains N298, which displays the largest CSP of all assigned backbone NH resonances (Figure S 8A and C), supporting a long-range structural effect of glycan binding into this region of the protein. Apart from overlapping peptides covering the protein regions presented in Figure 14 no other peptides of either the wild type or the deamidated P dimer show significant deuteration differences. This indicates 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.

**A**



**B**
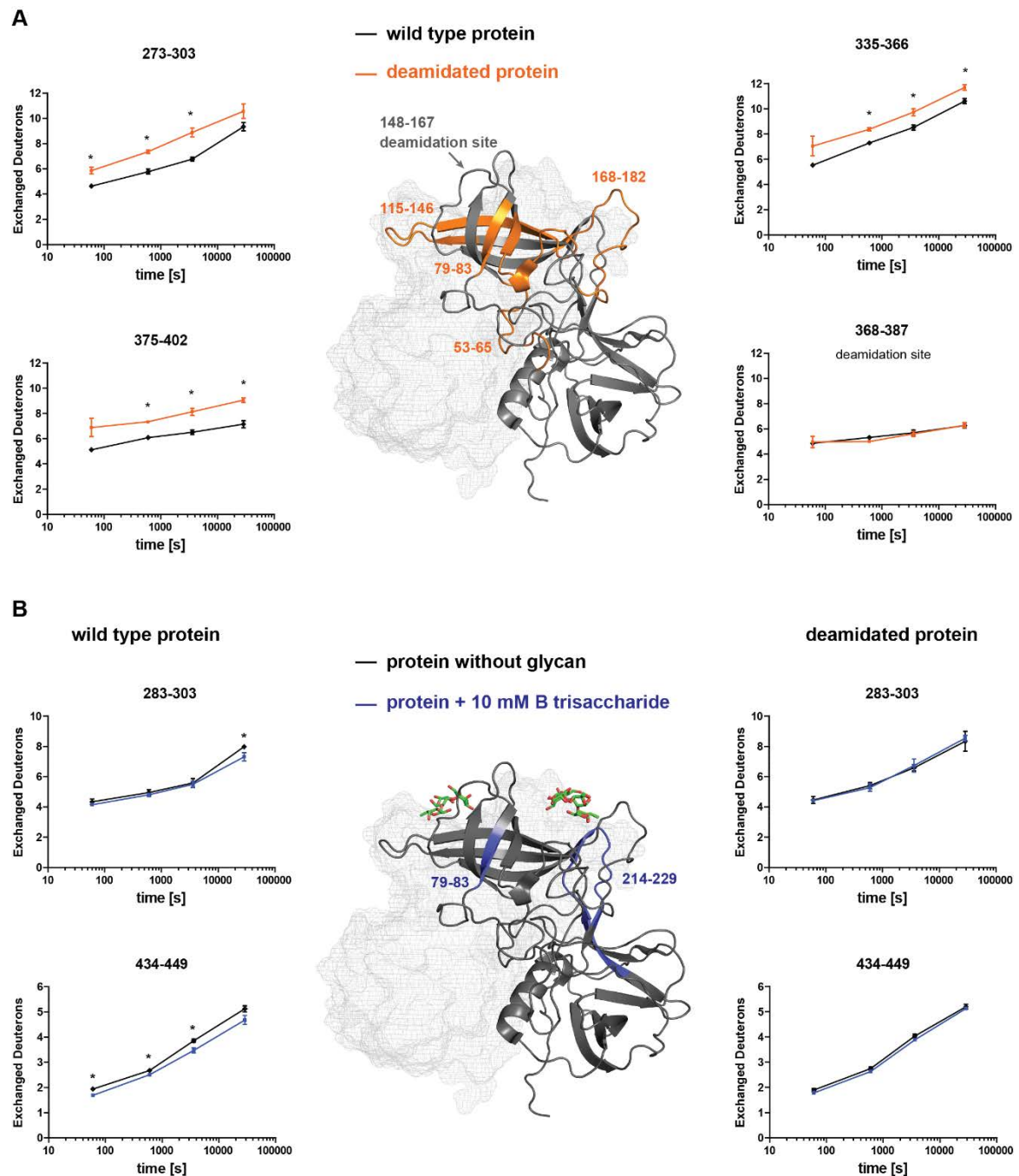


Figure 14: HDX-MS experiments reveal changes in protein dynamics upon deamidation. (A) Deuterium uptake differences of the wild type and deamidated GII.4 Saga P dimer mapped to the crystal structure (pdb 4X06). (B) Deuterium uptake differences of the wild type and deamidated P dimer in presence and absence 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. All measurements were performed in triplicate and error bars

indicate the standard deviation. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.

Binding of 100 mM methyl α-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 HDX differences can be detected for the negative control, 100 mM D-galactose. Also in NMR, regions with CSPs are similar for B trisaccharide and methyl α-L-fucopyranoside binding (Figure S 8 and Figure S 9). As hypothesized from the qualitative HDX-MS and NMR data, chemical shift titration experiments with NN and iDiD P dimers at varying concentrations of methyl α-L-fucopyranoside revealed strong binding attenuation in the iDiD P dimer represented by a 10-fold higher dissociation constant $K_d$. Re-evaluation of previous NMR data (Mallagaray et al., 2015) based on the new findings yielded a $K_d$ of 5.5 mM for B trisaccharide binding to NN Saga P dimers. A recent study confirms these results and supports an approximately 10-fold decrease in binding affinity also for B trisaccharide (Creutznacher et al., submitted)(Table 2).

Table 2: Summary of glycan binding affinities of GII.4 P dimers measured by NMR.References: 1: Creutznacher et al. (submitted), 2: Mallagaray et al. (2019), 3: Mallagaray et al. (2015), [RE] re-evaluation

| Norovirus strain | Composition | ligand | $K_d$ in mM [reference] |
|---|---|---|---|
| GII.4 Saga | NN | B trisaccharide | 6.7 ± 0.4 [1] <br> 5.5 [3, 2[RE]] |
| GII.4 Saga | iDiD | B trisaccharide | approx. 60 [2] |
| GII.4 Saga | NN | Methyl α-L-fucopyranoside | 22.3 ± 0.04 [2] <br> 15.6 ± 1.9 [3, 1[RE]] |
| GII.4 Saga | iDiD | Methyl α-L-fucopyranoside | 219.8 ± 4.5 [2] |
| GII.4 MI001 | NN | B trisaccharide | 4.3 ± 0.4 [1] |

The deamidated GII.4 Saga P dimer was also subjected to crystal structure analysis in the presence of methyl α-L-fucopyranoside. Upon incubation of deamidated P dimer apo crystals with 600 mM methyl α-L-fucopyranoside a co-crystal structure at 1.52 Å resolution was obtained, in which the isoD373 isopeptide bond is clearly visible in one chain of the P dimer (pdb 6H9V). Binding of methyl α-L-fucopyranoside to the two canonical HBGA binding sites in the GII.4 Saga P dimer was also observed. However, no binding to additional sites (sites three and four) as observed for GII.10 P dimers (Koromyslova et al., 2015) was detectable, even at this drastic ligand concentration. Whereas the overall P dimer structure is almost identical to previously published GII.4 Saga P dimer structures (Cao et al., 2007) binding of methyl α-L-fucopyranoside to the modified binding pocket shows some deviation from the native P dimer. In the native site, a bivalent hydrogen bond is formed between the side chain carboxyl group of Asp374 and the hydroxyl groups at positions C2 and C3 of methyl α-L-fucopyranoside. In the deamidated site, the Asp374 side chain carboxyl group makes only one hydrogen bond with the C2 hydroxyl group of methyl α-L-fucopyranoside (Figure S 10). The altered D374 side-chain orientation and the loss of one hydrogen bond can be explained by altered conformational dynamics and could contribute to the observed loss in binding affinity.

### 3.1.3  THE ROLE OF N373 DEAMIDATION IN GLYCAN BINDING

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 isoAsp affects HBGA recognition. In fact, loss of binding affinity due to isoAsp formation has been described for monoclonal antibodies (Yan et al., 2016). Chemical shift titrations show a dramatic loss of affinity of about one order of magnitude upon formation of the isopeptide linkage for the case of methyl α-L-fucopyranoside binding to GII.4 Saga P dimers. 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.

Crystal structure analysis provides an explanation for the observed loss of binding affinity. 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. The altered D374 side chain orientation and the loss of one hydrogen bond can be explained

by altered conformational dynamics in the loop. In addition, HDX MS reveals alterations of the whole P2 subdomain dynamics upon formation of the isopeptide linkage in both GII.4 strains. Except for the loop carrying the isopeptide bond, the overall structure of the P dimer is very similar to the structures published before (Singh et al., 2015b). However, numerous long range CSPs and deuterium uptake differences measured by HDX MS are observed upon deamidation. This implies that effects of deamidation are not confined to the vicinity of the deamidation site suggesting differences in overall protein dynamics.

## PARAMETERS DETERMINING THE HIGH REGIOSELECTIVITY OF N373 DEAMIDATION

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 (Geiger and Clarke, 1987, Robinson et al., 1970). It has been recognized early on that for proteins other factors are dominant (Robinson and Robinson, 2001a, Robinson and Robinson, 2001c). In particular, the conformation of the sequence containing the deamidation site plays a crucial role (Plotnikov et al., 2017). One important finding is that the deamidation motif requires a defined local conformation of the Asn deamidation site i and the i+1 residue. This specific conformation 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. Analysis of the crystal structures of GII.4 Saga P dimers shows that N373 in both monomers matches the required reactive conformation. This conformation is also found in known crystal structures of other GII.4 P dimers carrying the N373-D374 motif (Figure S 11).

Superposition of the loops containing the N373-D374 motif in published crystal structures shows that the conformation of this loop is highly conserved and, therefore, may be critical for the formation of a special local environment promoting deamidation. It can be assumed 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 GII.4 MI001 the critical loop has the sequence STDTS**N**D[374], which is almost identical to the one found in GII.4 Saga P dimers (STDTE**N**D[374]). Although there is no crystal structure available for GII.4 MI001, this suggests

that the conformation of the critical loop is identical to the conformation found for GII.4 Saga.

## BIOLOGICAL SIGNIFICANCE OF IRREVERSIBLE ASN DEAMIDATION

Deamidation and isomerization of asparagine residues as well as isomerization of aspartate residues are spontaneous processes generating PTMs associated e.g. with protein aging. Therefore, these modifications serve as molecular clocks controlling the life time of proteins (Robinson et al., 1970, Robinson and Robinson, 2001b). On the other hand, a cellular enzymatic repair system exists, reconverting iso-aspartate into aspartate and restoring the conventional protein backbone (Geiger and Clarke, 1987, Griffith et al., 2001, McFadden and Clarke, 1987). Identification and characterization of these PTMs, which are linked to a single unit change in molecular mass, is a challenging analytical problem (Chazin et al., 1989, Eakin et al., 2014, Grassi et al., 2017, Hao et al., 2017, Noguchi et al., 1998). 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 (Deverman et al., 2002, Lee et al., 2012) and confer thermal stability in enzymes present in hyperthermophilic bacteria (Kumar et al., 2016). In general, there are a number of PTMs modifying the protein backbone and thus modulating protein functions (Muller, 2017). 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 (Mallagaray et al., 2019). In fact, almost all GII.4 strains, for which crystal structures of P dimers have been obtained, carry this motif. Given a half-life of N373 of the order of only one to two 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 (Donaldson et al., 2010, Lindesmith et al., 2014, Lindesmith et al., 2018). Therefore, it would be of interest to further characterize the role of this deamidation *in vivo*, to draw conclusions for antiviral therapy and vaccine development.

## 3.2   Structural impacts of glycan binding in P dimers of other strains

Given the high structural similarity of the analyzed P dimers with at the same time strain-dependent glycan binding behavior, the next step was to study if glycan binding induces different dynamic changes in P dimers of different strains. Furthermore, the structural effects of fucose binding to a mixture of wild type and deamidated GII.4 MI001 and GII.4 Saga P dimers were examined, to get more insights about the dynamic changes induced by N373 deamidation. In this study, the differences in structural responses to glycan binding of norovirus P dimers belonging to the following strains were addressed: The Asian epidemic strain GII.17 Kawasaki 308 (Chan et al., 2017), the rarely detected strain GII.10 Vietnam 026, which can bind up to four fucose molecules (Koromyslova et al., 2015) and the GII.4 MI001 strain, which belongs to the highly pandemic GII.4 genotype and has been shown to infect mice as well (Taube et al., 2013).

### 3.2.1   P DIMER DEAMIDATION IN OTHER HUMAN STRAINS

To verify the deamidation status of the P dimer samples, they were subjected to peptic cleavage followed by LC-MS for peptide identification (Table 3). No deamidated peptides were identified for GII.10 Vietnam and GII.17 Kawasaki after 2 years and several months of storage at 4°C, respectively. For GII.4 MI001 stored several months at pH 7.3 and 4°C, a fraction of approximately 64 % carried a deamidation at N373. Based on this identification, the ratio of purely native (NN) to half deamidated (iDN) to fully deamidated (iDiD) P dimers would be 13:46:41 %. Furthermore, a small fraction (< 10 %) was deamidated at N239 and N448, respectively. In GII.4 MI001 P dimers stored at pH 4.9, no deamidation apart from a small fraction of deamidated N448 was identified. GII.4 Saga P dimers stored for several months at pH 7.3 were also partially deamidated at N373 with a deamidated fraction of approximately 88 % (ratio 1.5:21:77.5 % NN:iDN:iDiD) (Table 3 and Figure S 14 - Figure S 20)

Table 3: Identification of deamidation sites in GII.4 MI001 and GII.4 Saga P dimers stored in buffers of different pH. No deamidation sites were identified in GII.10 Vietnam and GII.17 Kawasaki P dimers. Deamidation positions with highest probability were calculated in MaxQuant based on peptide fragment spectra.

| Strain (storage pH) | Peptide sequence | Deamidated position | Deamidation position probability | deamidated fraction (%) |
|---|---|---|---|---|
| GII.4 MI001 (pH 4.9) | RSTMPGCSGYPNM**N**L | N448 | 0.99 | 8 |
| GII.4 MI001 (pH 7.3) | FRSTMPGCSGYPNM**N**L | N448 | 0.82 | 2 |
| | STDTS**N**DFETGQNTRF | N373 | 1 | approx. 64 |
| | **N**SRFPIPLEKL | N239 | 0.98 | 7 |
| GII.4 Saga (pH 7.3) | STDTE**N**DFETHQ | N373 | 0.96 | approx. 88 |

## VERIFICATION OF OLIGOMERIC STATE BY NATIVE MS

Prior to HDX-MS analysis P dimers were subjected to native MS for quality control. Furthermore, ion exchange separated wild type (NN) and fully deamidated (iDiD) GII.4 Saga P dimers were measured for comparison. GII.17 Kawasaki, GII.10 Vietnam, wild type GII.4 MI001 and wild type GII.4 Saga P domains showed dimers with the expected molecular masses, apart from a small fraction of unspecific tetramers formed during the ESI process (Figure 15). Interestingly, both deamidated GII.4 P domains were also present as monomers. Increased monomer fractions correlate with the amount of N373 deamidated species, with 16 % monomer in 64 % deamidated GII.4 MI001 and 32 % monomer in 100 % deamidated GII.4 Saga (Table 4). As deamidation rate in these strains is identical (Creutznacher et al., submitted) this suggests that monomers are a result of iDiD P dimer dissociation, while iDN species are still primarily dimeric.
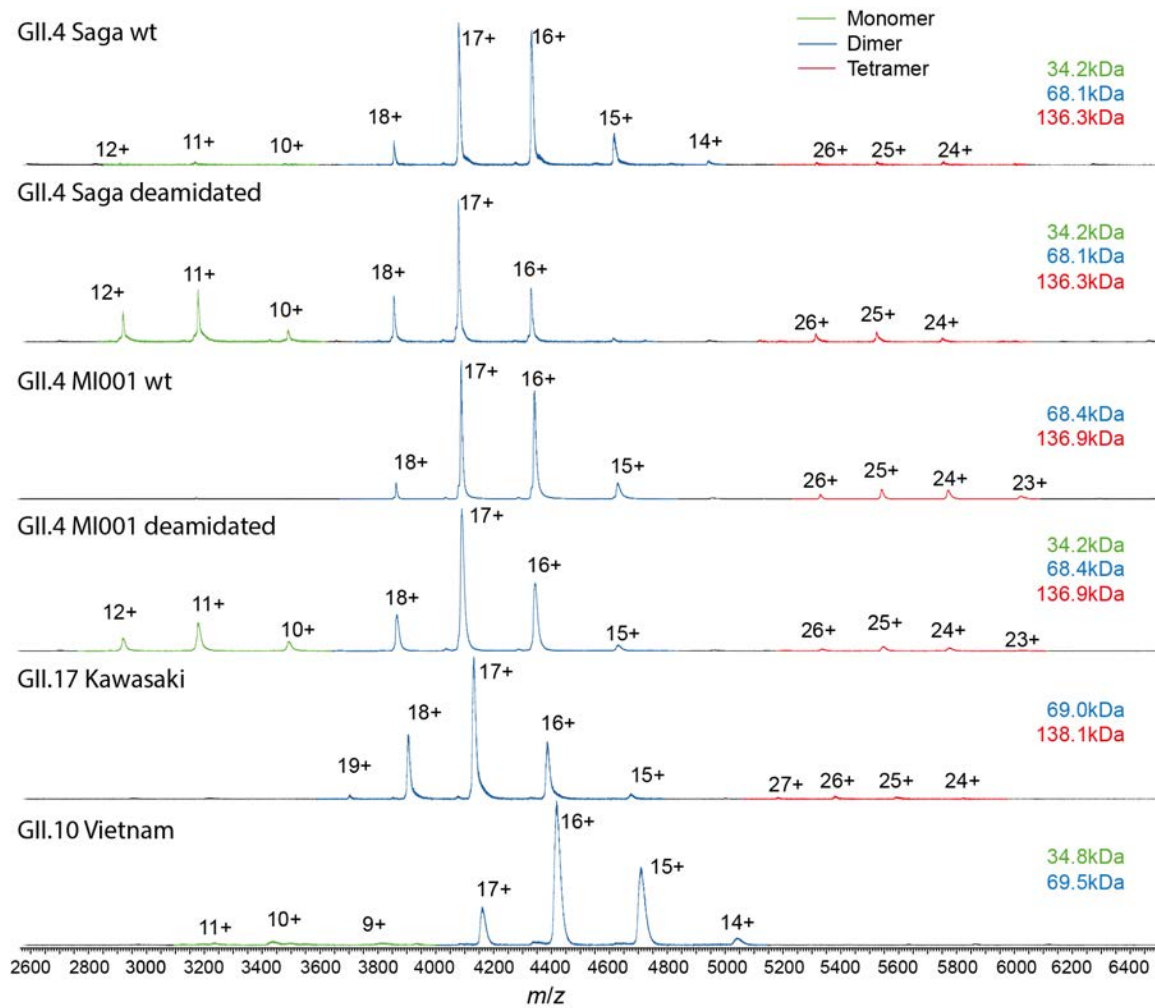
Figure 15: Native mass spectra of different human norovirus P domains. GII.17 Kawasaki, GII.10 Vietnam, wild type GII.4 MI001 and wild type GII.4 Saga P domains are present as dimers with the expected molecular masses, apart from unspecific tetramers formed during the ESI process. Both deamidated GII.4 P domains (100 % deamidated GII.4 Saga and 64 % deamidated GII.4 MI001) are also present as monomers. No native MS data is available for the partially (88 %) deamidated GII.4 Saga P dimer.

Table 4: Fractions of P monomers, dimers and tetramers from native MS

| Strain | Monomer fraction (%) | Dimer fraction (%) | Tetramer fraction (%) |
|---|---|---|---|
| GII.4 Saga deamidated (100 %) | 32 | 60 | 8 |
| GII.4 Saga wild type | 0 | 100 | 0 |
| GII.4 MI001 deamidated (64 %) | 16 | 80 | 4 |
| GII.4 MI001 wild type | 0 | 91 | 9 |
| GII.17 Kawasaki | 0 | 100 | 0 |
| GII.10 Vietnam | 4 | 96 | 0 |

### 3.2.2 BIMODALITY INDICATES A SECOND, MORE PROTECTED CONFORMATION

To analyze possible strain specific differences in the structural response to glycan binding wild type GII.4, GII.10 and GII.17 P dimers were incubated with 10 mM HBGA B trisaccharide and 100 mM fucose at pH 7 and differences between the unbound and ligand-bound state were measured using HDX-MS. During inspection of the glycan binding data, a bimodal character with one low intense, low deuterated and a second high intense, higher deuterated peak distribution was observed in the deuterated spectra of some peptides (Figure 16 A). These bimodal peak distributions can have many causes, e.g. two distinct protein conformations, conformational rearrangements that lead to EX1 exchange kinetics, insufficient ligand saturation or peptide carry over from the analytical or protease column (Fang et al., 2011, Hodge et al., 2019).

To rule out effects induced by carry over, an additional dataset with randomized sample order and additional washing of the pepsin column between sample injections was measured for GII.4 MI001 NN P dimers incubated with fucose, which still showed bimodality. Moreover, bimodality is also observed in absence of any ligand, strongly indicating that undersaturated binding sites are not the origin of bimodality. Furthermore, ligand concentrations were chosen to provide

high and comparable saturation of binding sites. HBGA B trisaccharide affinities are identical for GII.4 Saga and GII.4 MI001 P dimers (Creutznacher et al., submitted). Assuming that P dimers of other strains bind glycans with affinities similar to GII.4 P dimers, binding pocket occupancy can be estimated based on $K_d$ values measured by NMR (GII.4 Saga: 5.5 mM for HBGA B trisaccharide and 22 mM for fucose) (Mallagaray et al., 2019). Binding of HBGA B trisaccharide in our setup would hereby correspond to 95 % binding site occupancy during equilibration with ligand (98 % for fucose) and 65 % during deuterium labeling (82 % for fucose). For GII.4 P dimers, recent NMR measurements suggest two independent binding sites (Creutznacher et al., submitted, Mallagaray et al., 2019), whereas occupancy of up to four binding sites has been reported for 100 mM fucose in crystal structures of GII.10 Vietnam P dimers (Koromyslova et al., 2015).

## COMPARISON OF BIMODALITY IN DIFFERENT STRAINS

Bimodality can be seen in GII.4 MI001, GII.10 Vietnam and GII.17 Kawasaki, in the unbound proteins as well as in ligand bound proteins, but has not been seen in the GII.4 Saga P dimer (Mallagaray et al., 2019). For all three strains, bimodality is almost exclusively present in the P2 domain and in the lower part of the P1 domain (Figure 16 B). Peptides that are affected by glycan binding are often bimodal as well; therefore, it was necessary to manually analyze these regions again by binomial fitting of the individual peak distributions (an example analysis can be seen in Figure S 13). Residues 334-354, for example, are bimodal and also involved in glycan binding in all three strains. However, not all peptides that are protected upon glycan binding are bimodal. Peptides covering the canonical binding site are unimodal in GII.17 Kawasaki and GII.10 Vietnam and only show slight bimodality in GII.4 MI001 in the presence of fucose, indicating that protein conformation is conserved in these regions to facilitate glycan interaction.

Relative intensities of the individual peak distributions are constant over time (Figure S 13) and highly similar for peptides within the same protein, which lead to the assumption that the P dimer adopts two distinct conformations, a compact and a more flexible one (Guttman et al., 2013). The relative intensity ratios of the peak distributions vary between experiments, but can still be compared within a certain experiment (Figure 16 C).

**A**



**B**   **bimodal regions in P dimers of different strains**



**C**   **relative intensity of first peak distribution in bimodal spectra**
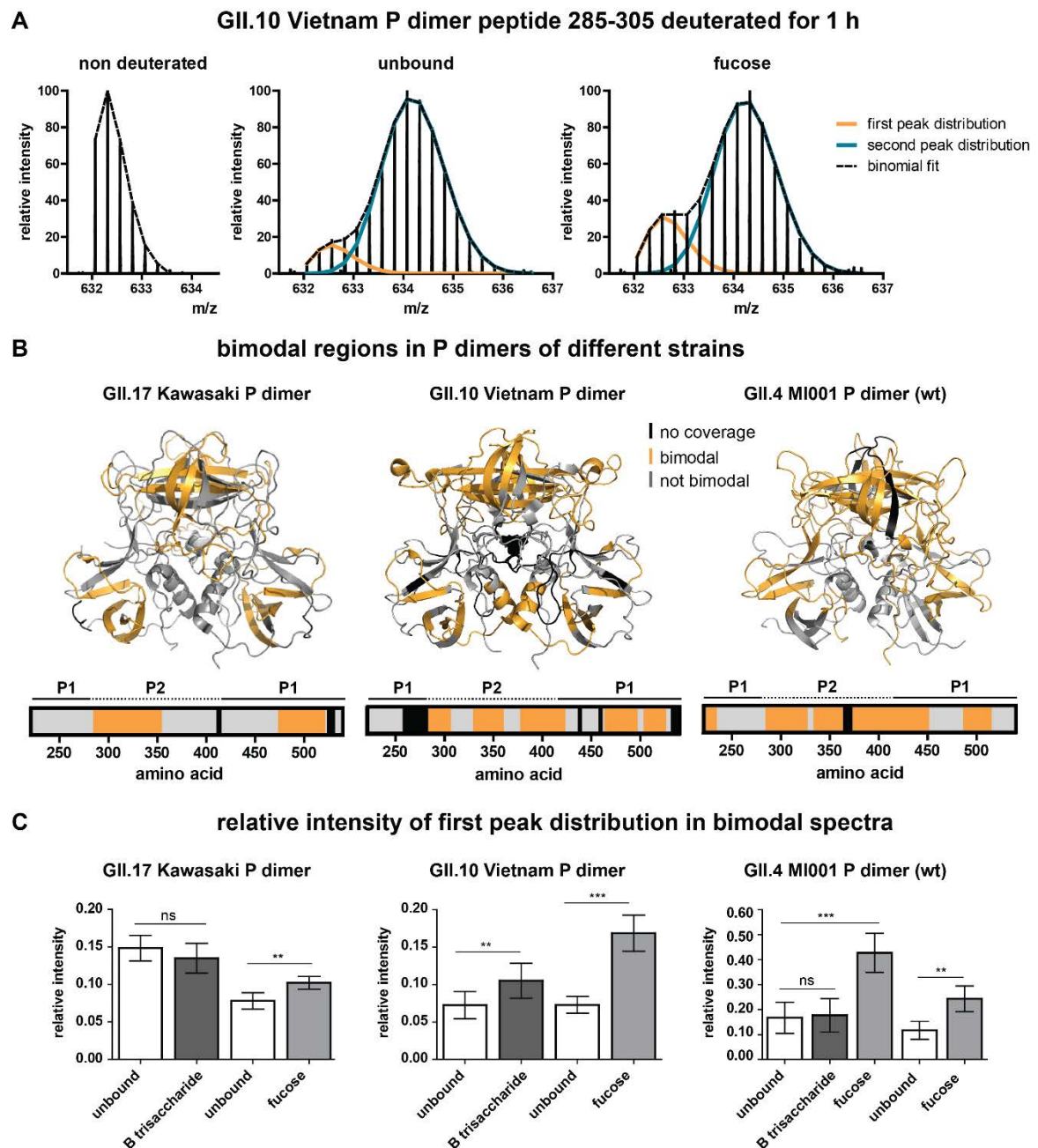


Figure 16: Bimodal peak distributions in some peptides point towards the existence of a second, more protected conformation. (A) Bimodality in deuterated spectra of an exemplary peptide of the GII.10 Vietnam P dimer. Bimodality occurs in both the unbound and the glycan bound state. Single peak distributions can be separated by binomial fitting. (B) Regions in P dimers of different strains that show bimodality in deuterated spectra (orange). Bimodality mainly occurs in the P2 domain and the lower part of the P1 domain in all strains. The amino acid numbering is based on the VP1 sequence. (C) Relative intensity of the first (lower deuterated) peak distribution for different strains and experiments calculated by binomial fitting. For GII.4 MI001 both HDX-MS experiments are shown (triplicate measurement with B trisaccharide and fucose vs. single measurement with fucose and additional wash steps between the injections). Relative intensity of the

first peak distribution stayed constant over time, so the averages over all bimodal time points for all states were calculated for several peptides of different protein regions and combined into bar graphs. The error bar represents the standard deviation of the average relative intensity calculated from $N \geq 5$ peptides. Significant differences between the unbound and the glycan bound state were assessed using a two-sided Student's t-test for each pair in an individual experiment. P values are indicated by asterisks: $p < 0.001$ (***), $p < 0.01$ (**) and not significant (ns).

Depending on the strain and the individual experiment, the relative intensity of the first peak distribution in the unbound P dimer varies between 7 and 17 %. For GII.17 Kawasaki and GII.4 MI001 P dimers, incubation with HBGA B trisaccharide has no significant effect on the relative intensity of the first distribution, while there is a slight increase for GII.10 Vietnam P dimers. Presence of fucose, in contrast, significantly increases the relative intensity of the first distribution in all strains. While there is only a slight increase for GII.17 Kawasaki P dimers, the relative intensity of the first distribution increases by a factor of 2 for GII.10 Vietnam and GII.4 MI001 P dimers.

## BIMODAL PEAK DISTRIBUTIONS COULD ORIGINATE FROM OLIGOMER FORMATION

The bimodality in deuterated peak distributions of certain protein areas points towards two distinct protein conformations that experience a different level of HDX over the whole exchange period. The low deuteration of the first peak distribution suggests the presence of a compact conformation that is shielded from HDX. This could be in principle true for protein aggregates, however, the location of this compact conformation only on the top and bottom of the protein appears very distinct. In case of protein aggregation, bimodality all over the protein surface would be expected.

So, if the low deuterated population is no artifact, what could it be instead? The P domain can form larger oligomers of different stoichiometry, up to whole 24-mer P particles (Figure S 12)(Bereszczak et al., 2012, Tan et al., 2008). P particles form contacts through interactions near the bases of each P dimer, which could explain the reduced deuteration in this area (Tan et al., 2008). Closer inspection of the cryo EM structure also suggests more contacts between the P2 domains compared to free P dimer. P particles can bind HBGAs and are even suspected to interact with them in the same way as VLPs (Tan et al., 2008, Tan and Jiang, 2005). However, in presence of glycans no significant deuteration difference could

be detected in the low deuterated peak distributions. Nevertheless, this does not explicitly mean that there is no binding in these areas. Peak distributions in peptides covering the canonical glycan binding site are unimodal and show protection in all strains, meaning this interaction can be found in the entire protein population.

The lack of deuteration changes in the low deuterated populations could mean that dynamics in these P oligomers are different from P dimers. However, interpretation of the absence of a significant deuteration difference on a certain time scale as no structural changes should be treated with caution for several reasons (Masson et al., 2019). First, the intensity and thus the signal-to-noise ratio in the lower deuterated peak distribution is low, which makes it difficult to detect statistically significant changes in deuteration. Secondly, HDX is reduced in the lower deuterated population compared to P dimers and would therefore need longer exchange times to reflect potential differences. There is a clear increase of the potential P oligomer population in presence of 100 mM fucose, which could mean that interaction with glycans supports P oligomerization that is otherwise less efficient (Tan and Jiang, 2005). It has to be noted that P oligomers were not observed in native MS of a 4.5 µM P dimer solution. This could be explained by the fact that the less than 20% of monomers, assumingly assembled into 24-mer P particles in absence of glycans, would amount to around 1% of total signal intensity split up into many charge states in native MS, which likely drop below detection limit. In contrast, fractions of structural variants of less than 5 % can be detected in a properly conducted HDX-MS experiment (Hageman and Weis, 2019b). Importantly, the results imply that GII.4 Saga has a different ability to form P oligomers than the closely related GII.4 MI001 strain.

### 3.2.3 STRUCTURAL DYNAMICS IN WILD TYPE P DIMERS

Most of the deuteration changes in presence of HBGA B trisaccharide or fucose are detected in peptides that show bimodality. This commonly causes falsely low deuteration differences in centroid analysis, so individual binomial fitting of the two peak distributions was performed for some representative peptides in these regions to validate the observed deuteration differences in the main (second) peak distribution. As expected, binding of HBGA B trisaccharide and fucose induces changes in P dimers of all three strains, primarily in the P2 domain, indicating occupation of the glycan binding pocket (Figure 17).
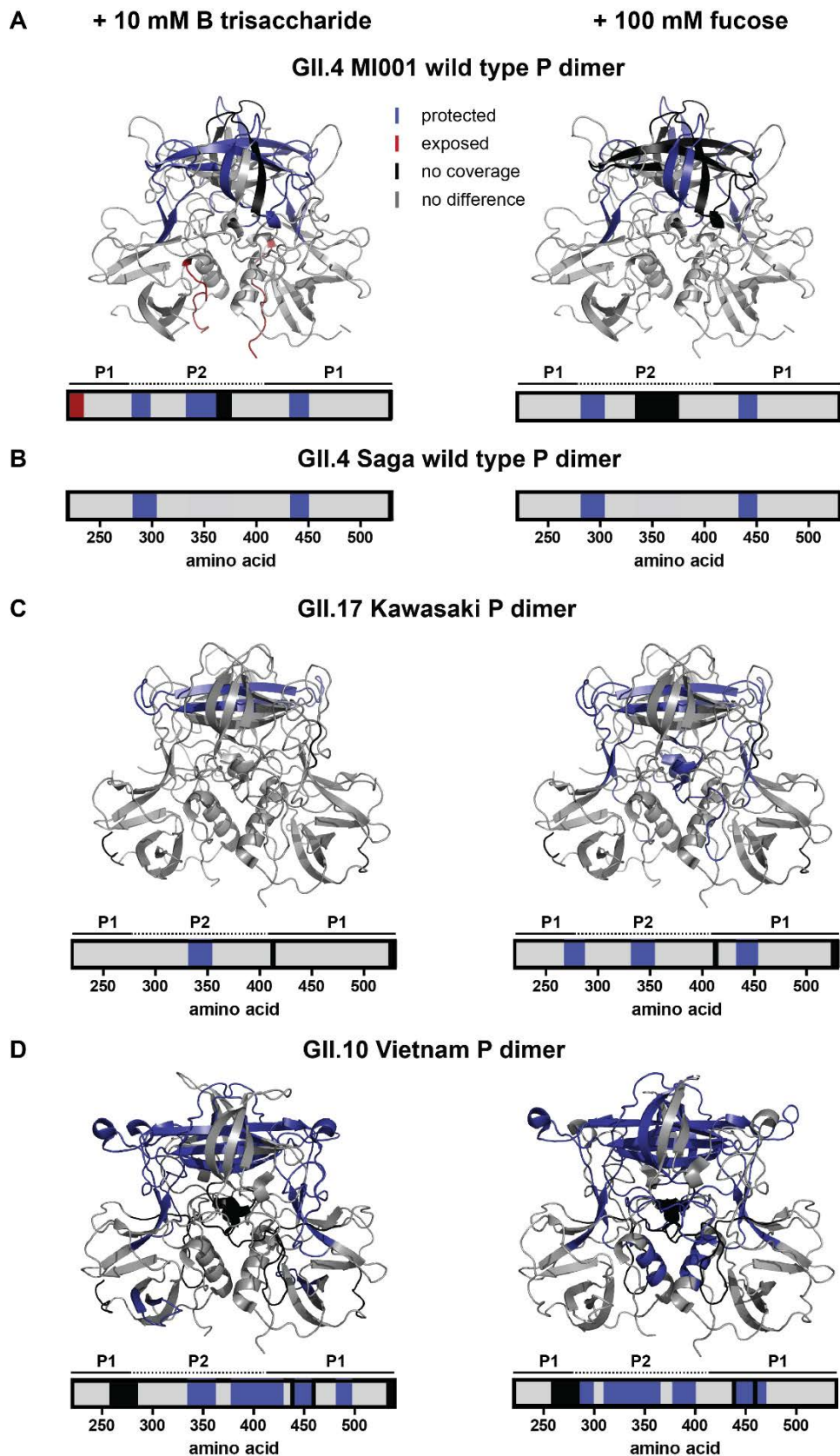
Figure 17: Structural dynamics induced by glycan binding in human GII.4 MI001 (A), GII.17 Kawasaki 308 (C) and GII.10 Vietnam (D) norovirus P dimers measured by HDX-MS. (B) Protected regions in wild type GII.4 Saga P dimers are shown for comparison (Mallagaray et al., 2019). Depicted are protein regions with significant

differences in deuterium uptake between unbound P dimers and P dimers with either 10 mM HBGA B trisaccharide or 100 mM fucose ($p < 0.05$, Student's t-test and $\Delta D > 2x$ pooled average SD). The deuteration difference in the second peak distribution has been manually validated by binomial fitting in case of bimodal spectra. Bar graphs and colored structures indicate regions of P dimers, which get more protected (dark blue) or exposed (red) upon interaction with glycans. Areas colored in grey show no significant difference in the chosen HDX time regime and black areas have no peptide coverage. P1/P2 refers to the two domains of the P dimer (shown in Figure 8).

Protected regions in wild type GII.4 MI001 P dimers are highly similar to GII.4 Saga P dimers (Mallagaray et al., 2019) for both glycans (canonical binding site G443, Y444 and residues 283-303, Figure 17 A and B). In addition, protection of a β-sheet region in the top cleft of the P2 domain (residues 333-353) can be detected in GII.4 MI001 in presence of HBGA B trisaccharide. Chemical shift perturbations in this region could also be seen in NMR experiments with GII.4 Saga P dimers in presence of glycans (Creutznacher et al., submitted, Mallagaray et al., 2019). Overall, protected regions in GII.4 MI001 match exactly with regions showing CSPs in GII.4 Saga NMR data, suggesting that both strains respond similarly to glycan binding.

Protection of residues 333-353 in the P2 domain can be seen in all three strains. For GII.17 Kawasaki P dimers, significant protection in presence of HBGA B trisaccharide is only present in this specific region. When incubated with fucose, additional protection of the canonical glycan binding site (G443, Y444) and residues 269-286, located in the protein center below the P2 domain, can be detected (Figure 17 C). In contrast to GII.4 MI001 and GII.17 Kawasaki, GII.10 Vietnam P dimers show protection in a lot of regions not only in the P2 domain including the canonical binding site (G451, Y452) and the β-sheet region in the binding cleft, but also some parts of the lower part of the P1 domain (Figure 17 D). The observed differences can only be seen in the second, highly deuterated peak distribution. The lowly deuterated peak distribution shows no significant differences between the unbound and the glycan-bound state in any of the strains indicating that only the highly deuterated species can bind glycans or labeling time was too short to detect deuteration differences in already strongly protected regions.

## STRAIN DEPENDENT DIFFERENCES IN WILD TYPE P DIMER GLYCAN BINDING

Apart from the presence of two distinct protein species in the sample, some differences in structural response to glycan binding in different strains could be detected as well. For GII.4 MI001, protected regions are almost identical to the earlier investigated GII.4 Saga P dimer (canonical binding site G443, Y444 and residues 283-303) (Mallagaray et al., 2019), apart from additional protection in the upper P2 binding cleft (residues 333-353). Involvement of this region has been seen in NMR data of GII.4 Saga P dimers as well (Mallagaray et al., 2019). Furthermore, a recent NMR study suggests identical glycan binding behavior of both GII.4 strains (Creutznacher et al., submitted). The same study also shows that MNV P dimers do not bind HBGAs, underscoring that infectivity of GII.4 MI001 in mice cannot be explained by different glycan-induced dynamics between GII.4 Saga and MI001.

GII.17 Kawasaki P dimer crystal structures with fucose and HBGA A trisaccharide show backbone interactions in T348 and G443 and side chain interactions in R349, D378 and Y444 (Koromyslova et al., 2017, Singh et al., 2015a). When incubated with HBGA B trisaccharide and fucose protection from HDX is observed for residues 333-353 corresponding to interactions with T348 and R349. In presence of 100 mM fucose, the canonical binding site (G443, Y444) is protected, as well as residues 269-286, which cannot be explained by the known interactions from the crystal structures. This region is located below the glycan binding cleft in the protein center, so protection from HDX could rather be the result of a long-distance structural change than of direct interaction with fucose. It would be interesting to see how long-distance structural changes would further propagate into the S domain in VLPs and if they would influence the dynamic P domain lift off from the S domain that has been seen for different norovirus strains (Jung et al., 2019, Smith and Smith, 2019, Snowden et al., 2020, Song et al., 2020, Tubiana et al., 2017).

For the GII.10 Vietnam strain, binding of two HBGA B trisaccharides and up to four fucose molecules has been seen in crystal structures (Hansman et al., 2011, Koromyslova et al., 2015). Compared to GII.4 MI001 and GII.17 Kawasaki, protection in more protein areas for both HBGA B trisaccharide and fucose can be seen, which mainly corresponds to the known glycan interactions summarized in Table 5. Due to close proximity of interacting amino acids in fucose binding sites 1/2 and 3/4 these binding sites cannot be distinguished in HDX data at peptide resolution, but occupation of all four binding sites is likely at the given

concentration (Koromyslova et al., 2015). Interestingly, a protection of several residue stretches cannot be explained by known glycan interactions. Residues 285-298 are protected in presence of fucose, similar to GII.4 strains. Residues 311-336 belong to an unstructured region below the P2 binding cleft and could link the protection observed in the cleft to the one in residues 285-298. In presence of HBGA B trisaccharide, protection of the aforementioned residues is not present under the chosen conditions. A possible explanation could be that these structural changes are triggered by occupation of binding sites 3 and 4 in the P2 cleft, which so far has not been seen for HBGA B trisaccharide at similar concentrations (Koromyslova et al., 2015). HBGA B trisaccharide binding is mainly mediated by the fucose residue, with an additional interaction of galactose with G451 and some water mediated interactions (Hansman et al., 2011). In HDX-MS a protection of residues 483-496 on the bottom of the P dimer is detected in addition, which could be a long-range effect not triggered by fucose alone.

Taken together, P dimers of all investigated strains show protection of the upper P2 binding cleft (residues 333-353) underscoring the importance of this region for glycan binding. Protection of the canonical glycan binding site (G443, Y444 for GII.4 and GII.17; G451, Y452 for GII.10) was also detected in all strains and for all glycans apart from HBGA B trisaccharide binding with GII.17 Kawasaki P dimers. HBGA B trisaccharide could have a lower binding affinity in GII.17 Kawasaki compared to the other strains that leads to smaller deuteration changes that are below the detection limit in the current experimental setup.

It has to be noted that the GII.17 Kawasaki datasets have a higher back exchange (D/H) than the other datasets so that small glycan induced deuteration changes are more likely to be lost during the measurement. GII.4 and GII.10 P dimers show protection of residues 285-298, which is also absent in GII.17 P dimers. Interestingly, P dimers of the more prevalent strains GII.4 and GII.17 (Chan et al., 2017, de Graaf et al., 2015) show less changes in HDX upon glycan binding compared to GII.10 Vietnam, which is rarely detected in patients (Hansman et al., 2011).

Table 5: Comparison of protected residues in HDX with known glycan interactions in crystal structures. Binding sites 1/2 are conserved for many strains and glycans, binding sites 3/4 have so far only been detected in GII.10 Vietnam. For GII.17 Kawasaki only crystal structures with fucose and A trisaccharide are available, so protected residues for B trisaccharide are compared to binding sites seen for A trisaccharide. P dimer chain annotations are given for fucose binding sites 1 and 3. No crystal structure is available for GII.4 MI001 P dimers, so binding sites are marked as unknown (NA). Protected residues for wild type GII.4 Saga P dimers (Mallagaray et al., 2019) are shown for comparison.

| Protected residues in HDX | Fucose binding site 1/2 | Fucose binding site 3/4 |
|---|---|---|
| GII.10 Vietnam + 100 mM fucose (Koromyslova et al., 2015) | | |
| 285-298 | - | - |
| 311-336 | - | - |
| 337-364 | N355 (chain A) R356 (chain A) | E359 (chain A) |
| 379-399 | D385 (chain A) | W381 (chain A) |
| 442-458 | G451 (chain B) Y452 (chain B) | L449 (chain A) |
| GII.10 Vietnam + 10 mM HBGA B trisaccharide (Hansman et al., 2011) | | |
| 336-361 | N355 (chain A) R356 (chain A) | - |
| 379-428 | D385 (chain A) | - |
| 440-458 | G451 (chain B) Y452 (chain B) | - |
| 483-496 | - | - |
| GII.17 Kawasaki + 100 mM fucose (Koromyslova et al., 2017, Singh et al., 2015a) | | |
| 269-286 | - | - |
| 333-353 | T348 (chain A) R349 (chain A) | - |
| 434-452 | G443 (chain B) Y444 (chain B) | - |
| - | D378 (chain A) | - |

Table 5: continued

| Protected residues in HDX | Fucose binding site 1/2 | Fucose binding site 3/4 |
|---|---|---|
| GII.17 Kawasaki + 10 mM HBGA B trisaccharide (Koromyslova et al., 2017) | | |
| 333-353 | T348 (chain A)<br>R349 (chain A) | - |
| - | G443 (chain B)<br>Y444 (chain B) | - |
| GII.4 MI001 + 100 mM fucose | | |
| 283-303 | NA | NA |
| 434-449 | NA | NA |
| GII.4 MI001 + 10 mM HBGA B trisaccharide | | |
| 283-298 | NA | NA |
| 333-353 | NA | NA |
| 434-450 | NA | NA |
| GII.4 Saga + 100 mM fucose and GII.4 Saga + 10 mM HBGA B trisaccharide (Mallagaray et al., 2019, Singh et al., 2015b) | | |
| 283-303 | - | - |
| no coverage | D374 (chain A) | |
| 434-449 | G443 (chain B)<br>Y444 (chain B) | - |

### 3.2.4 STRUCTURAL DYNAMICS IN PARTIALLY DEAMIDATED P DIMERS

To test whether deamidation of N373 attenuates glycan binding as observed for HBGA B trisaccharide in fully deamidated GII.4 Saga (Mallagaray et al., 2019) (Figure 18 C), HDX-MS experiments with the partially deamidated GII.4 MI001 and GII.4 Saga P dimer in presence of 100 mM fucose were performed.

#### FUCOSE-INDUCED DYNAMICS IN GII.4 SAGA AND MI001 P DIMERS

Strikingly, protection of the canonical fucose binding site (G443, Y444) could be detected in both strains (Figure 18 A and B). This shows that fucose binding is still possible in partially deamidated (iDN) or even fully deamidated (iDiD) P dimers at the given concentration, even though binding is attenuated compared to the N373 wild type. Occupation of the canonical binding sites has been seen in crystal structures of deamidated GII.4 Saga P dimers at elevated concentrations of 600 mM fucose, but binding interactions are slightly different from wild type (chapter 3.1).
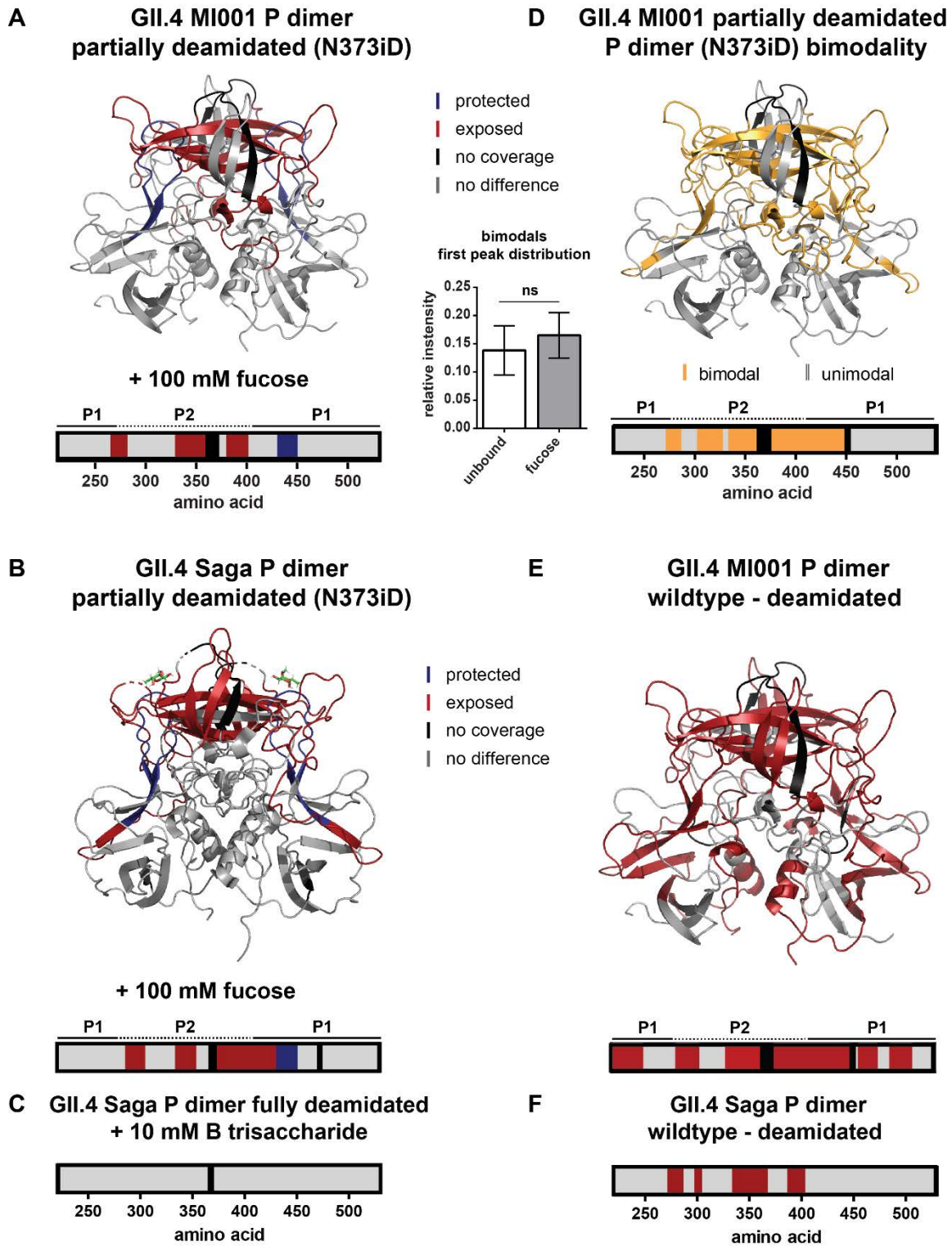
Figure 18: Structural dynamics in partially deamidated (N373iD) GII.4 P dimers. Fucose can still bind to the canonical glycan binding site in partially deamidated GII.4 MI001 (A) and GII.4 Saga (B) P dimers (pdb 6H9V)(significance criteria: Students t-test with p < 0.05 and ΔD > 2x pooled average SD). In contrast to the wild type N373 parts of the P2 domain get more exposed upon interaction with fucose. (C) HDX differences in fully deamidated GII.4 Saga P dimers in presence of 10 mM HBGA B trisaccharide are shown for comparison (chapter 3.1).

(D) Bimodality occurs in similar regions as for the wild type P dimer, implying that the more protected conformation is also present in the deamidated protein. In contrast to the native P dimer, the relative intensity of the first peak distribution does not significantly increase under fucose treatment. (E) The partially deamidated MI001 N373iD P dimer shows a higher deuterium uptake in large parts of the structure, which points towards higher flexibility, like in the fully deamidated GII.4 Saga P dimer (F) (chapter 3.1).

In contrast to wild type GII.4 MI001 P dimer, no other region is protected from HDX under fucose treatment, but increased deuteration is observed in the P2 domain of both GII.4 strains, suggesting a more exposed conformation. Interestingly, residues 335-362, which are protected in the wild type proteins of all strains, show increased deuteration in partially deamidated GII.4 P dimers.

As there is a mixture of wild type and deamidated P domains in the sample, the mass shift in the deuterated spectra reflects the average of all components, unbound and fucose-bound NN, iDN and iDiD P dimers, which cannot be discriminated. However, binding probability calculations can give a hint, which species contribute most to the observed increase in deuteration in presence of fucose. Considering the fractions of wild type and deamidated P domains and their different $K_d$s for fucose binding (chapter 3.1 Table 2), in GII.4 MI001 only 17 % of binding events occur in pure wild type NN dimers, 54 % in half-deamidated iDN dimers and 29 % in fully deamidated iDiD dimers. For GII.4 Saga even more binding events take place in fully deamidated P dimers (2:31:67 % for NN, iDN; iDiD), clearly showing that the detected increased deuteration is caused by fucose binding to at least half-deamidated P dimers. This suggests that fucose binding results in different structural responses in the wild type and partially deamidated protein.

## BIMODALITY IN PARTIALLY DEAMIDATED GII.4 MI001 P DIMERS

In GII.4 MI001 regions with bimodal peak distributions (Figure 18 D) and relative intensities of both distributions are similar to the ones in the wild type protein. However, interaction with fucose in the partially deamidated GII.4 MI001 does not lead to a significant increase in the relative intensity of the lower deuterated peak distribution, as seen in the wild type protein. Slight bimodality is also present in peptides covering the canonical fucose binding site in the fucose bound state but relative intensities are similar to the ones observed for the wild type protein.

## DEAMIDATION INCREASES DYNAMICS ALSO IN GII.4 MI001

A comparison of FD normalized deuteration levels for wild type and deamidated GII.4 MI001 P dimers without glycans revealed increased deuterium incorporation in large parts of the P2 domain, as well as the P1 domain (Figure 18 E). Highest deuteration differences ($\Delta D > 1$ Da) are detected for residues 335-432 located in the P2 β-sheet cleft. This is in line with the increased flexibility in the P2 domain of the deamidated GII.4 Saga P dimer (Figure 18 F) described in chapter 3.1, however, in GII.4 MI001 this effect is propagated into regions more distant from the glycan binding pocket and deamidation site.

The increased dynamics could weaken the dimer interfaces and therefore explain the dissociation into monomers in the deamidated protein (Figure 15). Additionally, monomers will most probably also experience higher HDX because of missing dimer interactions in the P1 domain and the P2 β-sheets (Chen et al., 2017) that will add to the observed increase in deuteration compared to the exclusively dimeric wild type protein.

## POSSIBLE ROLES OF N373 DEAMIDATION IN AN INFECTION CONTEXT

Deamidation is site-specific and happens over a timescale of 1-2 days at pH 7.3 and 37 °C correlating with the length of the infection cycle (Mallagaray et al., 2019). The high specificity of this deamidation under infection conditions lets suggest that this process could occur *in vivo* as well; however, the biological relevance for infection remains unclear.

To test if this effect can also be found in the closely related GII.4 MI001 strain, HDX-MS on a spontaneously deamidated P dimer sample was performed, resulting in mixed populations of NN, iDN and iDiD dimers, which is more likely to be found in a natural infection context. For comparison, fucose binding to a partially deamidated GII.4 Saga P dimer was also measured, which contained an even higher fraction of fully deamidated iDiD P dimers. Strikingly, protection of the canonical glycan binding site could still be detected in both isolates under fucose treatment, mainly corresponding to binding to iDN and iDiD P dimers. Additionally, the P2 cleft was more exposed in the partially deamidated sample under fucose treatment, in contrast to the protection observed in the wild type NN P dimer. This indicates that under natural deamidation conditions, glycan binding at the canonical binding site still happens, but induces different structural changes than in the purely wild type P dimer. The exposure of the P2 cleft suggests that after glycan binding this area gets more flexible, which could

be required to interact with other factors or the until now unknown receptor. As such an increase in deuteration is not present in wild type NN P dimers, this effect must be caused either by direct binding to deamidated P domains in iDN or iDiD dimers or by binding to wild type P domains in iDN dimers, whose overall dynamics are altered by the influence of the neighboring deamidated monomer. Neither protection of binding sites nor increased deuteration in P2 has been seen in previous HDX-MS measurements of fully deamidated GII.4 Saga P dimer with 10 mM HBGA B trisaccharide under nearly identical conditions (chapter 3.1). A possible explanation could be that HBGA B trisaccharide concentration was too low to induce the observed effects because of decreased binding affinity. Notably, NMR measurements of fully deamidated GII.4 Saga P dimers with HBGA B trisaccharide and fucose show large chemical shift perturbations around residues 370-380 (Figure S 8), a region where increased deuteration in presence of fucose is observed.

N373 deamidation could serve as a pH and temperature dependent mechanism to control infectivity of the virus. P dimers and VLPs have been shown to be stable under low pH conditions and temperature (Pogan et al., 2018a), where deamidation rate is low (Mallagaray et al., 2019). After entering the human host via contaminated food and reaching the intestine, the rise in pH and temperature facilitates conversion of wild type to deamidated P dimers that are able to attach to glycans and perform the structural change required for interaction with the receptor and infection of the target cell. This theory is supported by the observation of increased flexibility in the P2 cleft of iDiD GII.4 Saga P dimers compared to the wild type (chapter 3.1), which is also present in GII.4 MI001 P dimers. In summary this could mean that deamidation creates the required flexibility for host cell attachment and subsequent receptor binding. Attenuation of glycan binding in the deamidated P dimer could be counteracted by avidity due to high glycan presentation on cell surfaces *in vivo*. Native MS measurements of deamidated GII.4 Saga and MI001 P dimers also show that with increasing deamidation, dissociation into monomers occurs, whereas in NN P dimers no monomers are present. This could as well be linked to the increased flexibility of iDiD P dimers that weakens the dimer interface and shifts the monomer-dimer equilibrium. It would be interesting to investigate whether increased flexibility is limited to the P domain or whether it is propagated into the S domain in VLPs as well, which as a result could destabilize the particle and prepare for uncoating.

## CONSERVATION OF DEAMIDATION SITE IN DIFFERENT NOROVIRUS STRAINS

The question remains, which advantage the evolutionary conserved N373 deamidation site provides for the most prevalent GII.4 strains over other strains. One possibility is that higher flexibility induced by deamidation indeed enables better interactions with host receptors; another possibility is that it is part of an immune escape mechanism. N373 is located in the immunodominant antibody epitope A and even minor changes in the epitope sequence during viral evolution have resulted in the loss of monoclonal antibody response (Mallory et al., 2019). From all residues in this specific epitope, N373 seems to be the most conserved over time. Due to its close proximity to the glycan interacting D374, mutation of N373 could possibly impair glycan binding, which makes deamidation as a regulative mechanism for immune evasion without permanent mutation of the asparagine residue more favorable.

On the other hand, prevalence of GII.17 strains increased over the last years (de Graaf et al., 2015), and based on its RISDN**D**D[378] sequence the GII.17 Kawasaki P dimer is not able to deamidate at the equivalent position 377. Interestingly, a N373D mutated GII.4 Saga P dimer shows similar affinities to glycans as the N373 wild type (Creutznacher et al., submitted) clearly illustrating that iso-aspartate is required at this position to induce the observed changes. The absence of deamidation could also increase stability under a wide range of pH conditions, as dissociation into monomers is less likely to occur. Increased stability under alkaline conditions has been seen for GII.17 Kawasaki VLPs, however, other strains without potential deamidation sites were less stable at alkaline pH (Pogan et al., 2018b, Shoemaker et al., 2010). GII.10 Vietnam carries a glutamine at the equivalent position 384, which in theory can deamidate but deamidation is much slower and has not been observed after more than two years of storage at 5 °C and pH 7 (Mallagaray et al., 2019).

Further research is required to clarify the role of N373 deamidation in norovirus infection. Therefore, research focus should be shifted from wild type P dimers alone to the more likely occurring mixture of wild type and partially deamidated P dimers to elucidate the potentially important role of deamidation during infection. Furthermore, glycan binding studies with wild type and partially deamidated VLPs can give further information about the propagation of structural changes throughout the capsid.

## 3.3 Monomer-dimer equilibrium in human and murine P domains

Human and murine norovirus P domains adopt a very similar three-dimensional structure (see also Figure 8D), but show differences in their dimerization behavior. While human P dimers almost exclusively exist as dimers, murine P dimers are also present in a monomeric state after protein purification, but form dimers during crystallization (Nelson et al., 2018, Taube et al., 2010). To characterize these differences further and investigate the oligomeric state of murine P domains under native-like conditions, P domains of the two murine norovirus strains MNV07 and CR10 with 95 % sequence identity were measured at different protein concentrations by means of native MS. Mass spectrometry parameters were chosen to ensure native conditions and prevent protein unfolding and fragmentation.

### 3.3.1 MURINE P DOMAINS EXIST AS MONOMERS AND DIMERS

Both murine P domains are present as monomers and dimers in the native MS spectra and appear to be properly folded under the chosen conditions (Figure 19). P domain trimers and tetramers are also visible, but can be considered as unspecific clustering artifacts. Comparison of monomer and dimer peak areas at different protein concentrations shows that the CR10 P domain exists predominantly as monomer at low protein concentrations, whereas the MNV07 P domains shows a considerable amount of dimer at the respective protein concentrations.

This is also reflected by the different murine P dimer dissociation constants, which were calculated based on monomer/dimer peak area ratios obtained at different protein concentrations in native MS (Figure 20). The relative dimer peak area at different protein concentrations allows calculation of the P dimer dissociation constant, which is 3.5 times higher for the CR10 strain (32 µM) than for the MNV07 strain (9 µM). In contrast, human norovirus P domains are almost exclusively dimeric, even at very low protein concentrations (Figure 21), suggesting a much lower $K_d$ than in murine P domains.
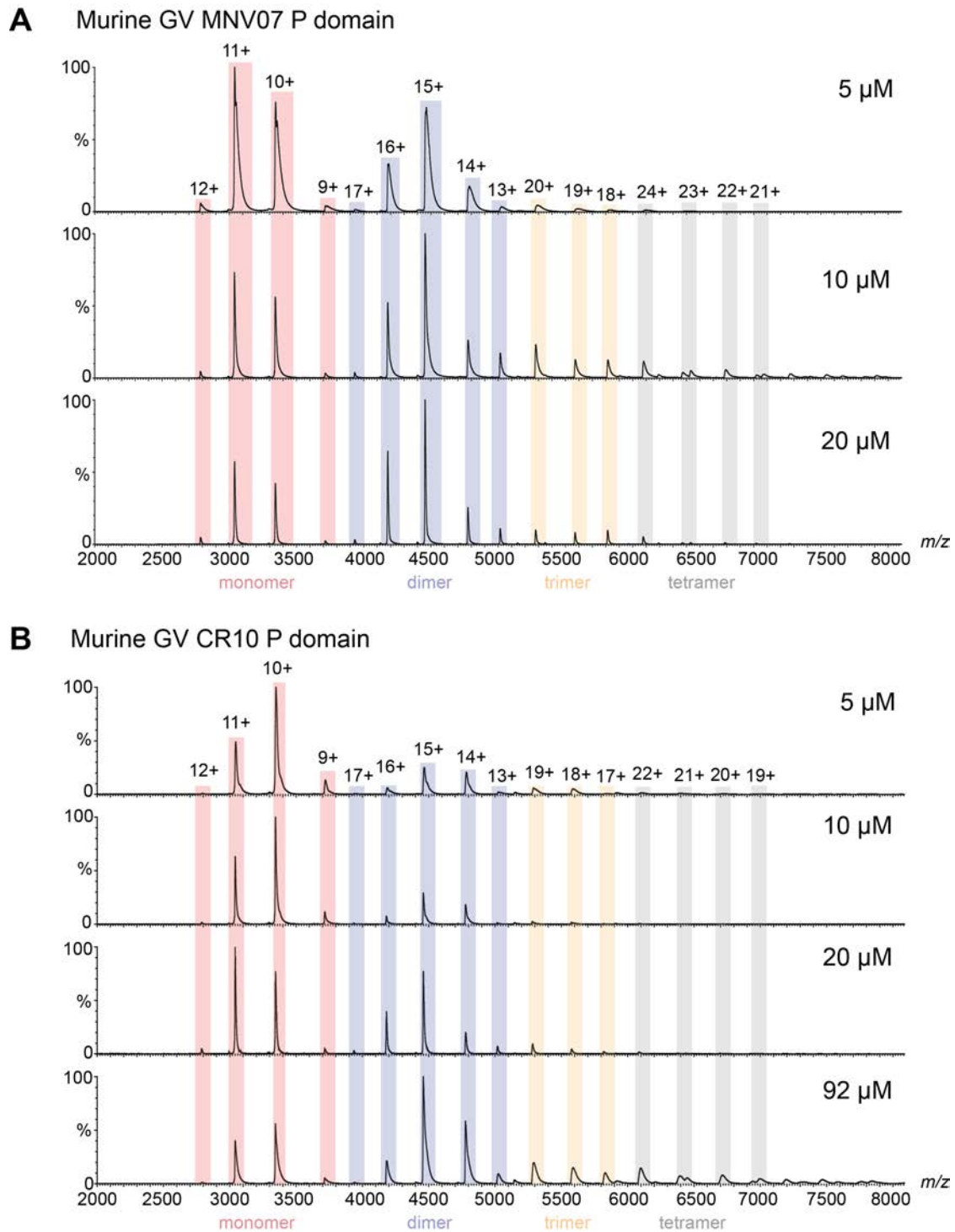
Figure 19: Native MS spectra of murine norovirus P domains at varying monomer concentrations. (A) The GV strain MNV07 forms a considerable amount of dimers even at lower concentrations, whereas the P domain of the CR10 strain (B) is mostly monomeric and only forms a considerable amount of dimer at higher protein concentrations. Detailed mass tables can be found in Table S 2.
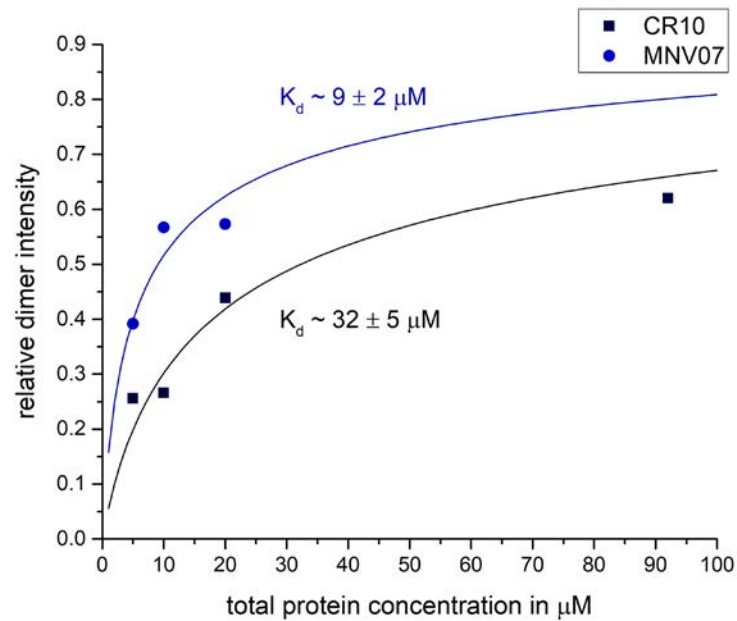
Figure 20: P dimer dissociation constants ($K_d$) for both murine norovirus strains. The relative dimer intensity was calculated based on monomer and dimer peak areas in native spectra (Figure 19) summed over all charge states The $K_d$ was then determined by global non-linear least squares fitting as described in chapter 5.4.2.
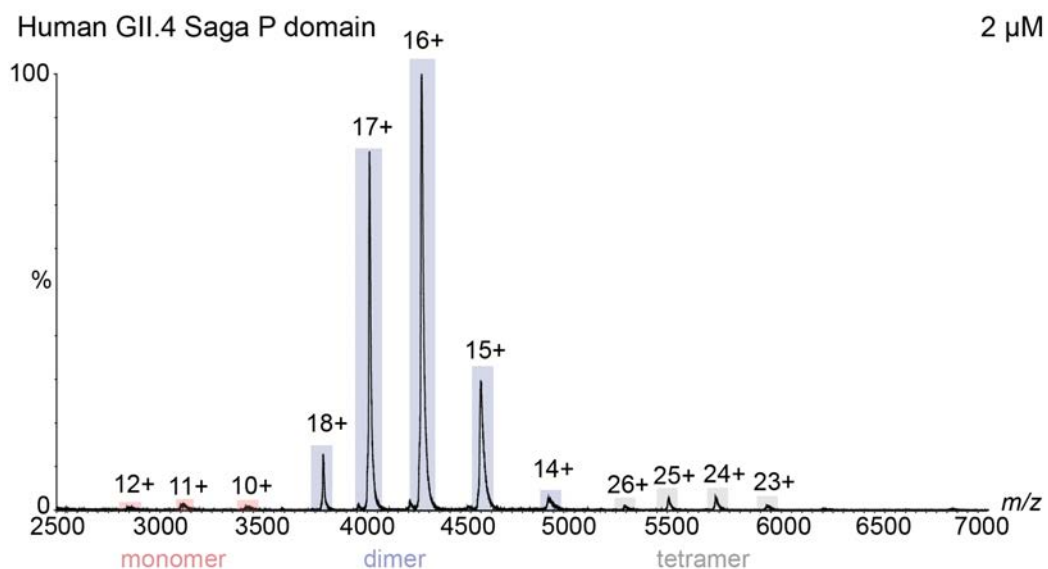


Figure 21: Native MS spectrum for human norovirus GII.4 Saga P domain at 2 μM monomer concentration. Even at this very low concentration, the P domain forms almost exclusively dimers.

### 3.3.2   MURINE P DIMER DISSOCIATION IS STRAIN AND LIGAND DEPENDENT

The native MS results confirm the existence of a monomeric and dimeric state of murine P domains under native-like conditions. However, the observation that even closely related MNV P domains show largely different $K_d$s is surprising and raises the question if these differences can be pinpointed to specific amino acid changes in the dimer interface.

MNV07 and CR10 P domains differ in only 15 amino acids, mostly located in the P2 domain (Figure 22). Changes in amino acids at the dimer interface are minimal, so that reduced dimer stability in CR10 cannot be easily assigned to a certain amino acid substitution based on the crystal structure. However, as described in chapter 3.2.1, deamidation of a single residue in the P2 domain is able to largely increase P dimer dissociation in human GII.4 noroviruses. Therefore, it can be suspected that small amino acid changes could also have strong effects on murine P domain dynamics. However, which amino acids could potentially induce these changes cannot be easily deduced from a static crystal structure. In comparison to human norovirus P domains, which are almost exclusively dimeric independent of the protein concentration, it is possible that MNV P domains in general are more dynamic and have smaller contact areas between the monomers in the dimeric state.
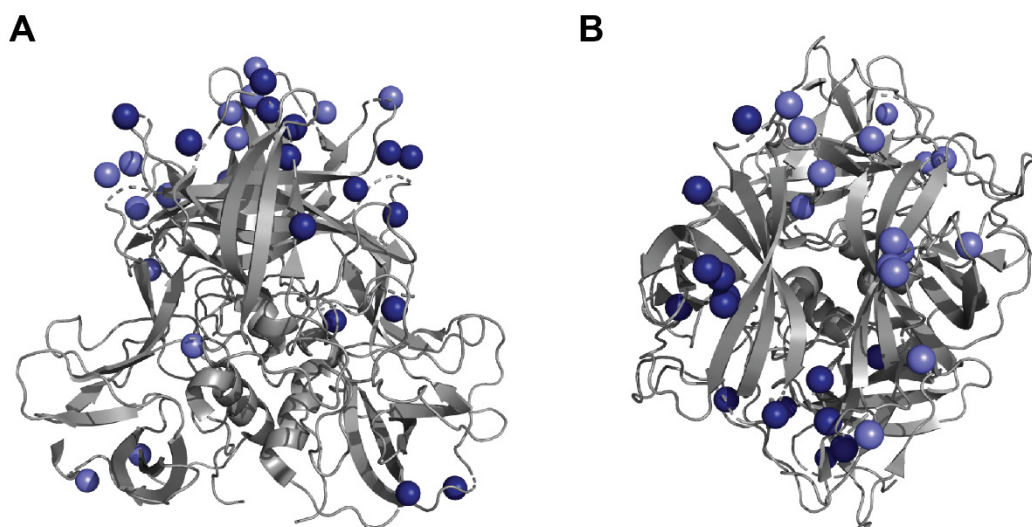


Figure 22: Amino acid differences between MNV 07 and MNV CR10 mapped to the MNV CR10 structure (pdb 6H6M) in front (A) and top (B) view. Light and dark blue dots indicate the amino acid differences between the MNV07 and CR10 strains in the individual monomers.

In recent NMR measurements with murine P domains of the closely related strain MNV1_CW (98 % sequence identity to CR10), a $K_d$ of around 30 μM was obtained for P dimer dissociation (Creutznacher et al., in preparation), which is in the same range as the $K_d$ obtained for the CR10 strain in native MS. Furthermore, the murine P domain rapidly degrades at neutral pH and shows decreased thermal stability at increasing pH. As observed in native MS, MNV P domains only dimerize at higher protein concentrations, in contrast to human P domains, which dimerize independently of the protein concentration. Interestingly, GCDCA binding strongly stabilizes MNV1_CW dimers, so that no monomers are observed anymore independent of the protein concentration. Furthermore, ligand binding increases the stability against thermal denaturation in solution (Creutznacher et al., in preparation).

In human norovirus P dimers binding of GCDCA to the MNV binding site (Figure 10) is sterically hindered and has so far not been observed. Based on the recent NMR and native MS data as well as crystal structures from literature, it can be hypothesized that murine P domains interact transiently with each other under ligand-free conditions and only form stable P dimers when bile acid is bound. This theory is also supported by the observation of ongoing monomer-dimer exchange on the time scale of the chromatographic separation in ligand-free MNV1_CW P dimers, which is absent after GCDCA binding (Creutznacher et al., in preparation).

It would be interesting to measure GCDCA-bound and unbound MNV P dimers in native MS as well, to confirm the switch of the oligomeric state observed in NMR and to determine the dimer stability in both states, e.g. by analyzing the P dimer dissociation behavior at increasing collisional activation. In the same way, MNV strain-dependent differences in GCDCA binding and dimer stabilization could be analyzed.

## 3.4 HDX-MS analysis of protein dynamics in lipid bilayers

To develop an HDX-MS protocol for proteins that interact with membranes, a collaboration project with the group of Prof. Michael Schlierf at the Technical University Dresden was set up. The project included screening of optimal HDX-MS conditions for the membrane protein Mistic as well as measurement of its dynamics in lipid vesicles of different compositions and in detergent micelles.

Parts of the following chapter are in preparation for publication in:
*Krainer, G., Batet, M., Dülfer, J., Anadamurugan, A., Textor, M., Frotscher, E., Hartmann, A., Uetrecht, C., Keller, S., Schlierf, M. (in preparation)*

### 3.4.1 ADAPTATION OF THE BOTTOM UP HDX-MS PROTOCOL

The standard bottom-up HDX-MS workflow includes labeling of the protein in deuterated buffer, followed by rapid mixing with quench buffer at pH 2.5 and flash freezing of the mixed sample. In doing so, the pH and temperature are reduced to 2.5 and -80 °C, respectively, to quench the exchange reaction as well as to unfold the protein and make it accessible for protease cleavage. Protease cleavage usually takes place on-column as part of the LC-MS step. Cleaved peptides are trapped and desalted, separated on an analytical column and then sprayed into the mass spectrometer for analysis. For membrane-associated proteins however, this process is not fully suitable and needs some optimization to gain optimal protein sequence coverage and prevent the introduction of large amounts of lipids and detergents that can deteriorate LC-MS performance (Martens and Politis, 2020, Trabjerg et al., 2018). Therefore, a lipid depletion step was introduced into the HDX-MS workflow and quench buffer and protease digestion conditions were screened for maximal sequence coverage of Mistic.

Before HDX, Mistic stored in LDAO detergent micelles was reconstituted into palmitoyloleoylphosphatidylcholine (POPC) and dioleoylphosphatidylcholine (DOPC) LUVs. The quality of the reconstitution was checked by DLS measurements, which showed uniform size distributions of lipid vesicles before and after protein reconstitution as well as no aggregation (Figure S 22).

After reconstitution and overnight dialysis to remove residual LDAO, the protein was diluted in undeuterated HDX buffer and mixed with different quench buffers to screen for optimal digestion conditions. Most studies investigating Mistic

(un)folding and structural dynamics are performed in detergent micelle environments (Broecker et al., 2014, Frotscher et al., 2018, Hartmann et al., 2015, Jacso et al., 2013, Krainer et al., 2018, Psachoulia et al., 2006, Roosild et al., 2005). The protein is known to partially unfold under high concentration urea treatment (Broecker et al., 2014, Jacso et al., 2013) and is less stable in non-ionic detergents like DDM than in zwitterionic detergents like DPC and LDAO, with greatest stability in the latter one (Hartmann et al., 2015, Krainer et al., 2018). Therefore, DDM and DPC were added to the quench buffer to test whether detergent introduction can facilitate unfolding of Mistic from membrane vesicles or if presence of urea only is sufficient.
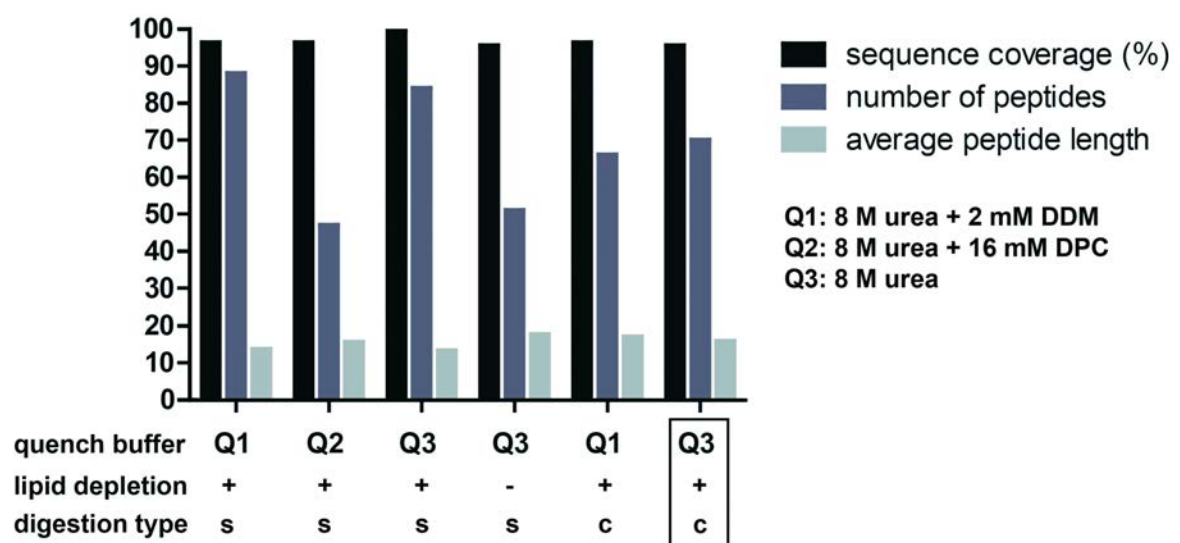


Figure 23: Mistic peptide identification in POPC vesicles with different glycine quench buffers (Q1-3, pH 2.4, mixed 1:1 with protein solution), (+/-) lipid depletion and digested either in solution (s) or on column (c). Sequence coverage and average peptide length are nearly identical in all conditions, but the number of identified peptides is strongly reduced in Q2 and without lipid depletion. The box indicates the combination of conditions that was used in the following HDX-MS experiments with Mistic. With this workflow, sequence coverage was almost identical in POPC, DOPC and LDAO (Figure S 23).

However, addition of detergents did not provide any benefit in terms of peptide number and sequence coverage (Figure 23). Detergents and lipids usually do not interfere with peptide peaks in the mass spectrum (Hebling et al., 2010), but they reduce column lifetime and require extensive and frequent cleaning of the mass spectrometry source. Therefore, the quenched samples were run through a filter with lipid absorbing $ZrO_2$ beads, which reduced the amount of lipids, but also the

peptide signal intensity. In most conditions, the number of peptides was higher in samples that were digested in solution compared to digestion on column; however, to reduce back exchange and avoid peak overlaps from pepsin autocleavage peptides in deuterated spectra, cleavage on column was preferred. The adapted HDX-MS workflow also showed increased back exchange and experimental error compared to the standard workflow (Figure S 24), mainly due to additional sample handing time during the lipid depletion step. This means that the adapted workflow has reduced sensitivity for smaller deuteration changes compared to the standard workflow, but still enables detection of larger conformational differences for Mistic in lipid and detergent environments.

### 3.4.2  FULLY DEUTERATED (FD) CONTROL PREPARATION

A fully deuterated control is a quality control sample that provides information about the back exchange of every peptide in the HDX workflow. This control is usually produced by labeling of the protein for 12-24 hours at room temperature in presence of a strong denaturant that ensures unfolding and therefore complete labeling of all backbone amides (Masson et al., 2019). While it is not strictly necessary to have this control in comparative experiments where only the relative differences of deuterium uptake between two states are compared (as in chapters 3.1 and 3.2), it is necessary if information about the absolute deuterium uptake is desired. This can be the case if deuterium uptake should provide information about structural features, e.g. folded and unfolded protein areas or the orientation or overall topology of a protein in the membrane. In these cases, the deuterium uptake of each peptide needs to be normalized with the uptake of the respective fully deuterated control.

For membrane proteins, however, preparation of a fully deuterated control can be challenging because of the high hydrophobicity and strong interactions with lipids and detergent that hamper sufficient unfolding at HDX conditions. This could lead to underestimation of deuterium uptake in still folded regions in the fully deuterated control and consequently an uptake overestimation of the respective regions in the normalized data. Therefore, it is important to ensure sufficient (or at least best possible) unfolding of the protein in the fully deuterated control (Wales et al., 2013).

MISTIC UNFOLDING UNDER UREA TREATMENT

Fortunately, folding and unfolding of Mistic has been studied extensively by different techniques (Broecker et al., 2014, Frotscher et al., 2018, Hartmann et al., 2015, Jacso et al., 2013, Krainer et al., 2018, Psachoulia et al., 2006, Roosild et al., 2005). As most of these studies are performed with Mistic in LDAO, this condition was also used to prepare the fully deuterated control. It has been shown that introduction of 6-6.5 M urea leads to disruption of the detergent micelle (Jacso et al., 2013) and partial unfolding of helices (Broecker et al., 2014, Krainer et al., 2018). Full unfolding of helices is only reached when LDAO is depleted (Broecker et al., 2014, Jacso et al., 2013, Krainer et al., 2018)(Figure 24). It was observed that helices 3 and 4 are the most stable and build a scaffold for the recruitment of helix 2 during refolding, while helix 1 is the least stable and is recruited last (Jacso et al., 2013).
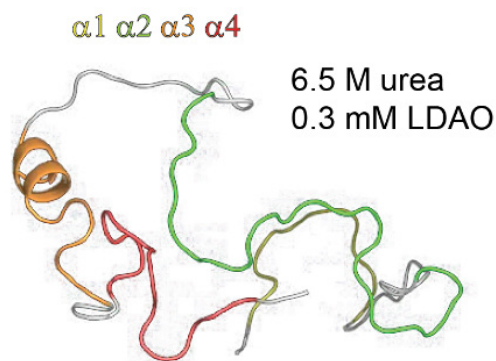


Figure 24: Snapshot of the folding trajectories of Mistic as obtained from molecular dynamics simulations. The snapshot is taken at 6.5 M urea and 0.3 mM LDAO (similar to the LDAO concentration in FD controls). Adapted with permission from Jacso et al. (2013), Copyright (2013) American Chemical Society.

To test optimal conditions for the fully labeled control, Mistic in LDAO was diluted to 480 µM, which is far below the critical micelle concentration (CMC, 1.84 mM, as given in Broecker et al. (2014)), to induce breakup of micelles. The protein was then directly labeled 24 h with deuterated buffer or preincubated 2 h with 6 M urea before the 24 h deuterium labeling step (carried out at 48 µM LDAO in both conditions). In the first condition, Mistic should be unstructured and interact transiently with LDAO instead of being folded and embedded in the micelle (Broecker et al., 2014). The second condition should result in increased unfolding and less interactions with LDAO, and therefore highest deuterium incorporation.

For comparison, intact Mistic in LDAO micelles (6 mM LDAO) was labeled for one week at room temperature, serving as "maximum deuterated" comparison in the folded, micelle embedded state.

Surprisingly, deuteration values of both FD controls (with and without urea incubation) were identical in the range of the experimental error. A possible explanation could be that the dilution of urea during the deuterium labeling step resulted in partial protein refolding, despite the absence of detergent micelles under labeling conditions. Furthermore, protein aggregation has been observed under low detergent/low urea conditions (Broecker et al., 2014), but could not be observed in the present HDX-MS experiment. However, aggregates could have been filtered out during the filtering step for lipid depletion. Deuteration of FD controls was higher than in folded Mistic labeled for one week, which confirms a more unfolded state with less contacts to LDAO molecules in the FD control compared to the micelle embedded state.

## NORMALIZATION OF DEUTERIUM UPTAKE

To get further insights about the topology and orientation of Mistic in the lipid bilayer, deuteration data needed to be normalized to compare the absolute deuterium uptake of certain protein regions relative to each other. As full unfolding and therefore full deuteration of Mistic under FD labeling conditions could not be granted, normalization with the FD control was compared to normalization with the number of exchangeable peptide backbone hydrogens (so-called "fractional uptake") which does not take into account the back exchange of the individual peptide. Relative differences between certain regions of the protein were almost identical with both normalizing approaches, except for the *N*-terminal region and helix (H) 1, which exhibited a very low fractional uptake, but very high FD-normalized uptake.

Analysis of back exchange in the *N*-terminus of the protein showed that this region has a 13 % higher back exchange compared to all other regions of the protein, without having a much higher intrinsic exchange rate compared to other regions of similar length (average overall BE 54 %). As explained before, this phenomenon could either be explained by high stability and insufficient unfolding of this region, which leads to artificially low deuteration in the FD control, or increased back exchange in these peptides during the measurement. Taking into account the results from Jacso et al. (2013), helix 1 is supposed to be the least stable of all four helices and unfolds first under urea treatment. It is therefore

very likely that the low FD control deuteration is indeed caused by increased back exchange. Consequently, data was finally normalized with the FD control to account for different levels of back exchange in certain regions of the protein.

This example shows that it is important to carefully screen for unfolding conditions for preparation of a proper FD control, when information about the overall structural features of the protein is desired. This is particularly challenging for membrane proteins, as they tend to interact strongly with the lipid environment and often aggregate when the lipid environment is depleted to induce unfolding. The degree of unfolding in presence of denaturing agents can for example be screened by CD spectroscopy, fluorescence measurements, FRET or NMR (Horne and Radford, 2016, Jefferson et al., 2018). Furthermore, pH and incubation time can be adjusted to allow for maximum labeling, but too harsh condition should be treated with caution to avoid falsely high deuterium incorporation into side chains (Masson et al., 2019).

### 3.4.3 MISTIC CONFORMATIONAL DYNAMICS IN POPC, DOPC AND LDAO

To gain further insights into Mistic conformation in different lipid environments, an HDX time course from 30 seconds up to 1 hour was measured for Mistic reconstituted in POPC and DOPC vesicles, as well as Mistic in LDAO detergent for comparison. Based on the FRET measurements (Krainer et al., in preparation), the hypothesis was that Mistic conformation would be strongly different in both lipids, being more compact in POPC and more stretched in DOPC (personal communication Michael Schlierf and Marta Batet). Thus, large HDX differences were expected all over the protein structure. However, deuterium uptake is very similar for Mistic in both lipids, indicating an overall highly similar structure.

Statistically significant differences are only found in three parts of the protein (residues 25-40, 52-67 and 74-93), with higher exposure of these regions in DOPC compared to POPC (Figure 25A and B). In LDAO, residues 25-40 and 74-93 have significantly higher deuterium uptake than in both lipids, whereas residues 52-67 are more protected in LDAO (Figure 25A and C).

The regions showing significant deuteration differences are identical in all three environments, indicating that regions that undergo dynamic structural changes are conserved, but the degree of these dynamics varies between the detergent and lipid environments.
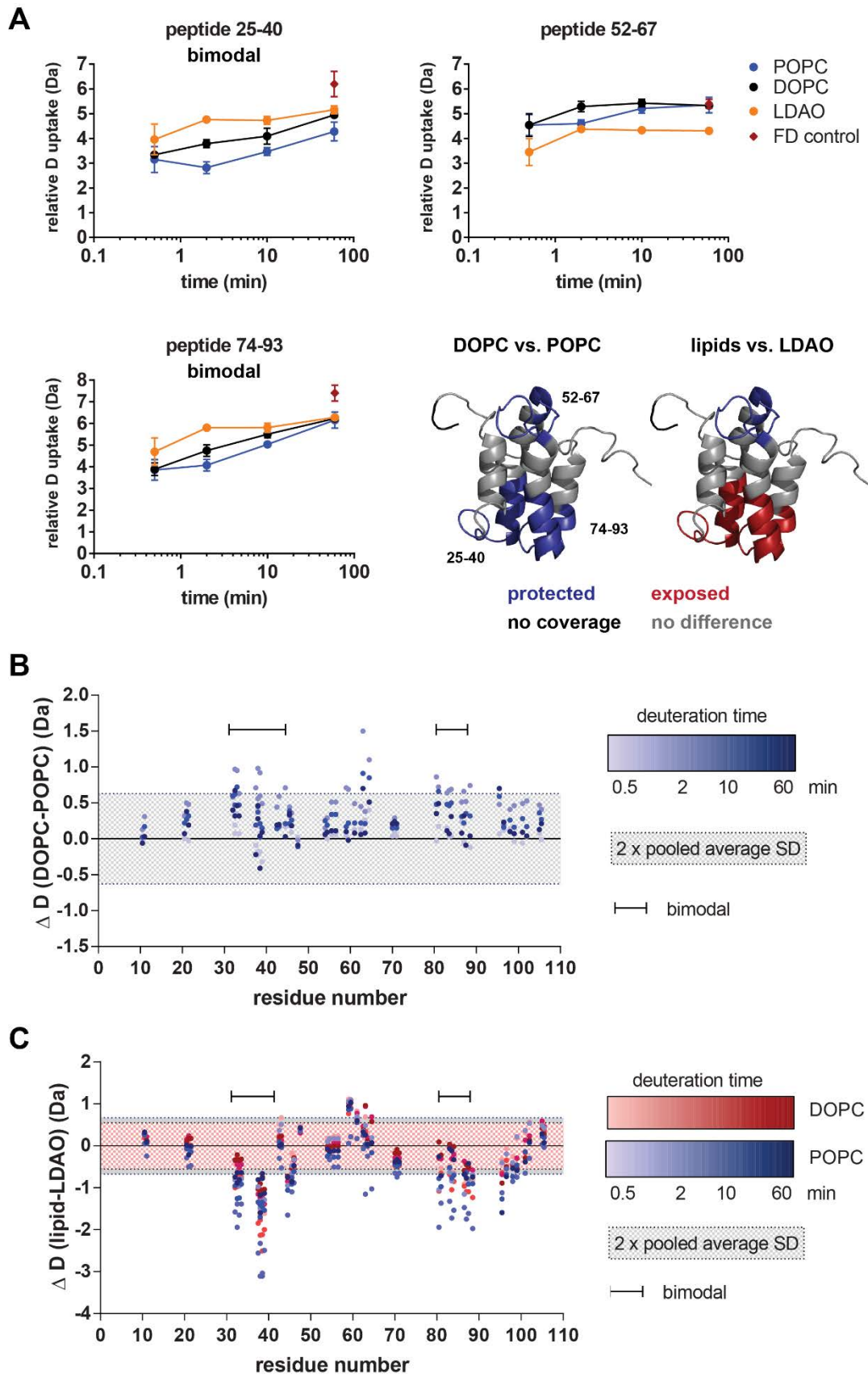
Figure 25: Significant deuteration differences of Mistic in POPC and DOPC lipid vesicles and LDAO detergent based on the centroid of the peak distribution.

(A) Uptake plots of peptides representative for the three regions where significant deuteration differences are observed. Error bars represent the standard deviation from three technical replicates. The structure represents the protein structure in LDAO detergent (pdb 1YGM). Significant differences were defined as passing a Students t-test ($p < 0.05$) and $\Delta D > 2x$ pooled average standard deviation of the dataset. (B) Difference plot for comparison of Mistic deuteration in POPC and DOPC. The grey area represents the threshold for a significant deuteration difference. (C) Difference plot for comparison of Mistic deuteration in LDAO and lipid vesicles (DOPC: red, POPC: blue). The colored area represents the threshold for a significant deuteration difference between DOPC-LDAO (red) and POPC-LDAO (blue, slightly higher than for DOPC). Differences around residue 100 did not pass the Students t-test and are therefore not considered significant.

Interestingly, residues 25-40 and 74-93 show bimodal peak distributions in LDAO as well as in lipid vesicles (Figure 26 A). In LDAO, two individual peak distributions are visible, while in lipids the two peak distributions are close to each other and result in one broad peak distribution. Binomial fitting of the individual peak distributions reveals that the first peak distribution corresponds to a very low deuterated conformation, e.g. due to strong interactions with detergents or lipids. A relative intensity shift from the first to the second peak distribution is observed over time, indicating conformational dynamics that induce mixed EX1/2 kinetics (Figure 26 C)(Guttman et al., 2013).

This relative intensity shift is more pronounced in LDAO than in lipids and might be caused by slightly different conformational dynamics in detergents and lipid vesicles. A comparison of the deuterium uptake of the two conformations reveals very similar deuteration of the more protected conformation (first peak distribution) in all three environments (Figure 26B). The deuteration of the more exposed conformation, however, is constant over time in LDAO and as high as in the FD control, whereas it is lower in lipids and only gets close to the FD control deuteration level after one hour.

To get more insights about the topology and the orientation of Mistic individual helices to each other in lipid vesicles, the absolute deuteration of certain parts of the protein structure was compared. For this purpose, the observed deuterium levels were normalized with the FD control to account for different levels of back exchange as described in chapter 3.4.2. The FD normalized deuteration percentage was then plotted onto the Mistic NMR structure using a color gradient with 10 % color steps.
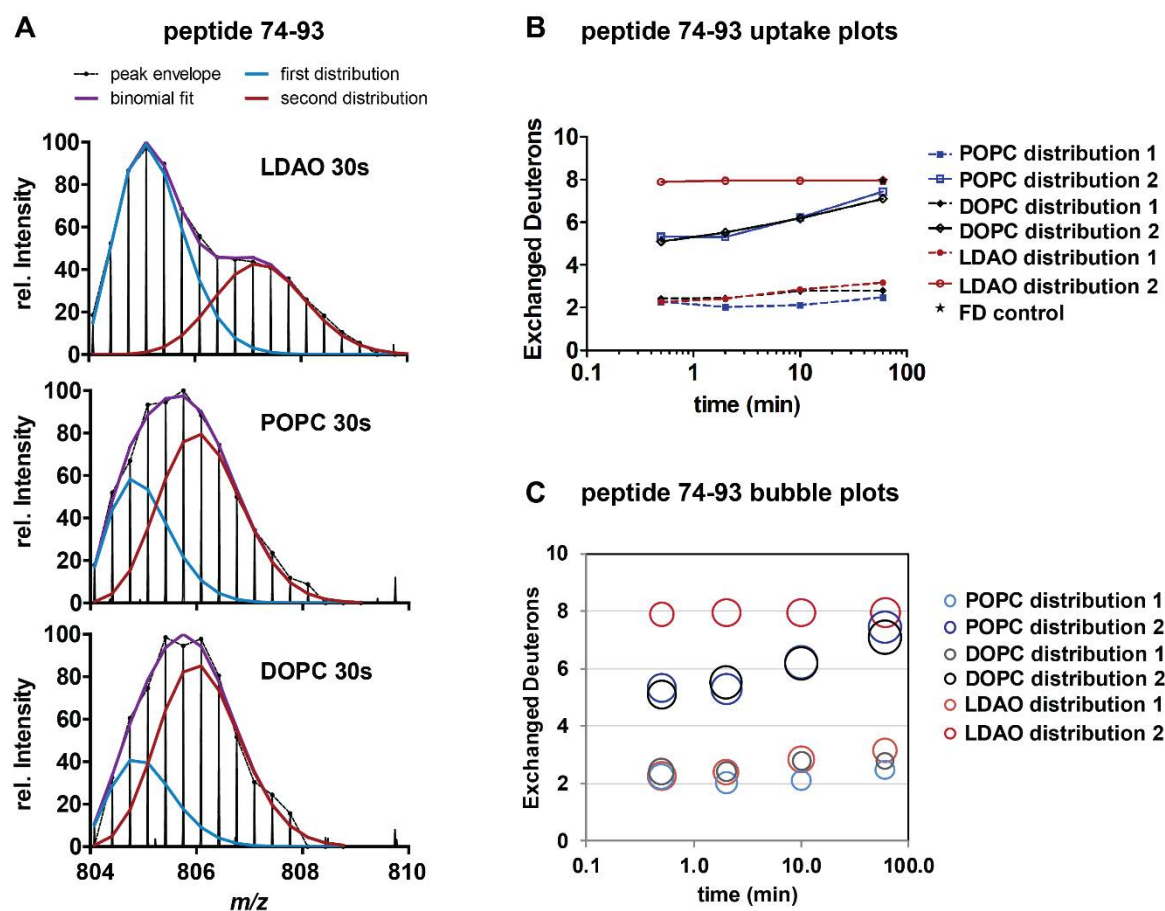
Figure 26: Exemplary analysis of bimodality in peak distributions of peptide 74-93. (A) Binomial fitting of the individual peak distributions of Mistic in LDAO, POPC and DOPC after 30 s of deuteration. (B) Deuterium uptake of individual peak distributions at every time point. The FD control is indicated as asterisk at the last time point. (C) Relative intensities and deuterium uptake of the individual peak distributions over time. The size of the bubble represents the relative intensity of peak distributions 1 and 2 at each time point. Plots represent the average of three replicate measurements.

Figure 27 exemplarily shows Mistic deuteration in POPC after 2 min of labeling in both the exposed and protected conformation (see Figure S 26 for all states and time points). The *N*-terminus, helix 1, the loop region between helix 2/3 and the *C*-terminus are highly deuterated which suggests an exposed location with high solvent accessibility. The middle part of helix 2 around the kink is lower deuterated and the upper part of helix 4 is very low deuterated which points towards more lipid interactions and/or more conformational rigidity compared to the higher deuterated regions. The protein regions between helix 1/2 and 3/4, which adopt two conformations, are either very low deuterated (protected state)

or medium/highly deuterated (exposed state) in comparison to the other protein regions.
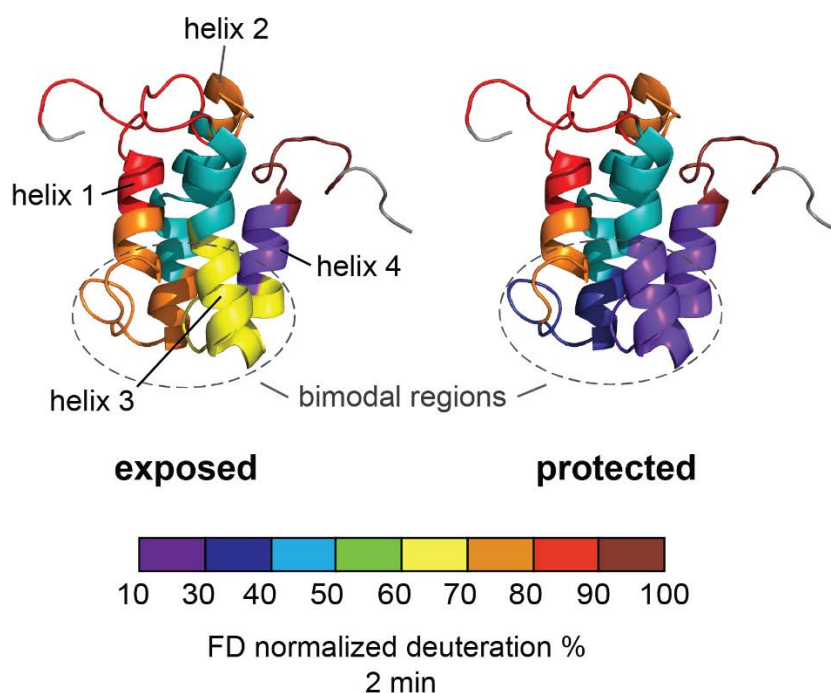


Figure 27: Mistic deuteration in POPC after 2 min of labeling. The percentage reflects the FD normalized deuterium uptake averaged over several overlapping peptides in each region. Peptides in encircled regions show bimodal peak distributions, resulting from an exposed and a more protected conformation in these areas.

### 3.4.4    MODEL FOR MISTIC MEMBRANE TOPOLOGY AND DYNAMICS

Differential HDX-MS results show that three regions of the protein exhibit significant deuterium uptake differences between lipids and detergent, but also between the two lipids. These regions are located in the H1/H2 loop and the *N*-terminal part of H2, the H2/H3 loop and parts of H3 and H4 including the loop between H3 and H4. In comparison to DOPC, uptake in these regions is reduced in POPC. Other protein regions did not show significant deuteration differences in POPC and DOPC which points towards a similar conformation, in contrast to the first assumption based on FRET measurements. In LDAO the H2/H3 loop region is more protected than in lipid vesicles, but the other two regions show a higher deuterium uptake compared to lipid vesicles.

## STRUCTURAL DYNAMICS IN DETERGENT MICELLES AND LIPID VESICLES

Furthermore, the H1/H2 loop and the *N*-terminal part of H2 as well as parts of H3 and H4 including the loop between H3 and H4 show clear bimodal peak distributions in LDAO and lipid vesicles, which indicate conformational dynamics that induce mixed EX1/2 kinetics. For the H3/H4 region, the deuteration of the higher deuterated peak distribution in LDAO is identical to the FD control while it is much lower in POPC and DOPC. This high deuteration could be either the result of higher conformational flexibility or incomplete folding of this region in LDAO. The degree of folding in LDAO was not measured before labeling, but the protein was produced using a reliable and reproducible protocol optimized to ensure complete folding of Mistic in LDAO. The protein preparation used in HDX-MS behaved identical to other Mistic preparations during all protein purification steps (personal communication Marta Batet). Furthermore, literature suggests that Mistic is usually stable and completely folded in LDAO as long as no denaturant is present (Broecker et al., 2014, Frotscher et al., 2018, Jacso et al., 2013, Krainer et al., 2018). Therefore, it is likely that the high deuteration is caused by structural dynamics in this area that leads to solvent accessibility similar as in the FD control. Furthermore, a drift of H4 from the protein core has also been seen in MD simulations of Mistic in LDAO (Psachoulia et al., 2006) which would result in higher solvent accessibility.

In contrast, deuteration of the lower deuterated peak distribution is only slightly different in LDAO, POPC and DOPC. This could mean that the conformation of the protected state is similar in these environments, but the degree of dynamics that leads to a structural opening strongly depends on the environment. In lipid vesicles, the exposure to deuterated solvent is less pronounced and only leads to increased deuteration over time, which could be the result of tighter packing or more stable interactions in lipid vesicles than in the more artificial detergent environment. However, the observed dynamics could have an influence on lipid bilayer stability, e.g. by introducing larger defects that lead to leakage of solvent into the bilayer and thus increase in deuteration.

## PROTEIN ORIENTATION AND MEMBRANE DEFECTS

The overall deuteration of Mistic in all environments is much higher than expected for transmembrane proteins that are usually deeply buried in the membrane and thus shielded from contact with deuterated solvent (Pan and Konermann, 2010). Such a high deuteration can therefore only be achieved if

either large parts of the protein are located at the outside of the membrane vesicle or membrane defects allow a leakage of solvent into the membrane where it is then able to deuterate the most accessible parts of the protein.

It has been suggested that, in spite of its classical alpha-helical structure, Mistic does not adopt a membrane-spanning topology (Textor, 2016). Molecular dynamics simulations as well as the proteins biophysical properties do not support a peripheral localization of Mistic and rather suspect an orientation in parallel with the membrane (Sansom et al., 2008, Textor, 2016). Furthermore, Trp13 in H1 is suggested to be located in the hydrophobic core of the membrane (Textor, 2016), which would be in contrast to a peripheral localization.

Interestingly, it has been shown that membrane defects are important for Mistic association with lipid vesicles as well as detergent micelles (Debnath et al., 2011, Krainer et al., 2018, Textor, 2016). Several studies also showed periplasmic location of the *C*-terminus (Roosild et al., 2005, Textor, 2016), which matches with the high deuteration in HDX-MS. Taken together, it can be assumed that the low deuterated protein regions are in contact with lipid/detergent while the high deuterated regions have less contacts and are more exposed to the deuterated solvent. Taking all these observations into account, a model of Mistic supposed membrane topology can be created (Figure 28).

## IMPLICATIONS ON POSSIBLE BIOLOGICAL FUNCTIONS

It has been suspected before that helix pairs H1/H2 and H3/H4 could serve as a sensor of membrane thickness by changes in distance to each other based on the lipid environment (Textor, 2016). In this theory, the loop between H1 and H2 would act as a flexible linker, which is supported by the observed HDX differences in this region. However, the H1/H2 and H3/H4 pairs are thought to be relatively fixed in their positions two each other, which is in contrast to the dynamics seen in HDX. HDX-MS differences of Mistic in POPC and DOPC reveal that the protected "closed" conformation is more exposed to deuterated solvent in DOPC, while the exposure of the "open" conformation is identical in both lipids.

This indicates that the lipid environment has an influence on the overall conformation of the protein, which is in line with FRET data (Krainer et al. in preparation). Protein dynamics, however, seem to result in very similar conformations in both lipid vesicles irrespective of the "closed" starting conformation. This means that if the "open" conformation has some biological function, this would be independent of the properties of the lipid environment.
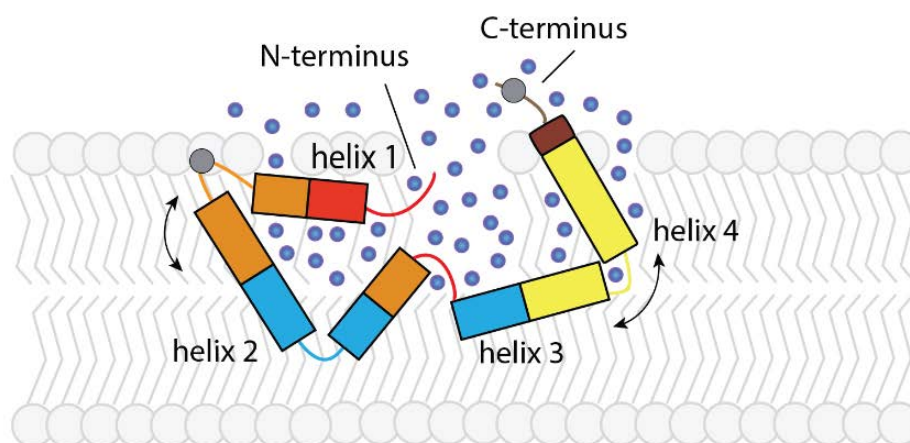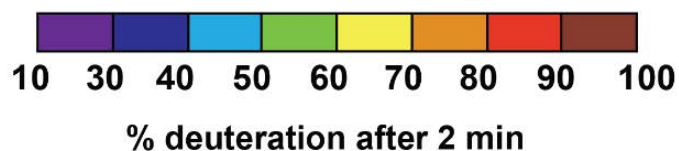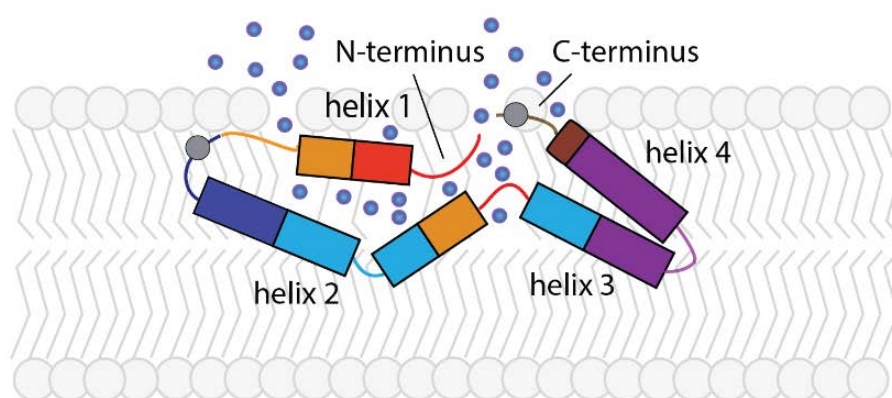
Figure 28: Model of Mistic overall topology and conformational dynamics in lipid vesicles based on FD-normalized HDX data. Two regions between H1/H2 and H3/H4 show bimodal peak distributions, which indicate the presence of two conformations (A: open, B: closed). In the open state more membrane perturbations are induced, which leads to increased penetration of solvent molecules (blue dots) into the lipid bilayer and thus higher deuteration of accessible parts of the protein. Grey dots represent the position of FRET labels in accompanying single-molecule FRET experiments.

However, the lipid properties could influence kinetics of these conformational switches, as seen by a larger shift towards the exposed conformation in lipids than in LDAO.

Taken together, HDX-MS data points towards a more complex biological function of Mistic than being a sensor of membrane properties alone. The observed dynamics could facilitate the interaction with fusion proteins to perform its chaperone function and promote overexpression of target proteins. In this context, it is important to note the *N*-terminal helix 1 does not seem to be required to promote overexpression as other Mistic homologs lack this helix (Roosild et al., 2006). In fact, H1 is the only helix that does not show significant differences in HDX between POPC and DOPC and could therefore be less influenced by different lipid properties.

In summary, HDX-MS could add some information about Mistic conformation and dynamics in different environments, but further work is required to fully uncover its conformational dependency on lipid properties, mechanisms of action and biological function. It could also been shown that protein dynamics in lipid vesicles are different from detergents, suggesting that only measurements in more native-like environments can provide a real picture of Mistic native conformation and dynamics.

## 4    OUTLOOK

The objective of this project was to set-up an HDX-MS workflow for the identification of strain-specific structural changes in norovirus P dimers upon glycan binding. The implemented HDX-MS standard workflow allows characterization of protein-ligand interactions and dynamics of soluble proteins in a robust and reproducible manner. The standard workflow has also been adapted for the investigation of protein-membrane interactions, which will be important in the future to map the occurring structural changes in the norovirus capsid under more native-like conditions by HDX-MS.

HDX-MS measurements on norovirus P dimers indicate a strain-specific response of the P dimer to glycan binding. In addition, N373 deamidation in GII.4 norovirus strains has large effects on glycan recognition and protein dynamics. Further studies are required to shed some light on the potential role of this deamidation in the infection process. Even though HDX-MS experiments with norovirus P dimers in solution could already give insights into the different structural dynamics induced by glycan binding, they only provide a limited picture of the protein dynamics and interactions in the viral capsid. Therefore, future research will gradually move to VLPs and intact virions, as well as displaying glycans on surfaces or lipids to more closely mimic the natural infection context. In the following subchapters a brief overview of future research aims, the necessary HDX-MS protocol adjustments as well as potential challenges will be presented.

### 4.1.1    STRAIN DEPENDENT DIFFERENCES ON THE VLP LEVEL

The most important question would be whether the changes in protein dynamics upon glycan binding observed in P dimers are comparable in VLPs. Structural dynamics in the P domain could be influenced by the S domain and the VP2 protein and could be propagated throughout the capsid inducing changes far away from the glycan binding site. Furthermore, it would be interesting to investigate if deamidation also occurs in VLPs and, if so, how it influences the dynamics and stability of the capsid.

This may require an adaptation of the HDX-MS protocol, to account for increased peptide numbers and potentially higher stability of VLPs under acidic conditions. However, the method benefits from the limited size of the individual capsid proteins, which occur in multiple copies, so that the absolute number of peptides

observed is the same for single VP1 subunits and VLPs reducing overlap of peptide signals. Preliminary results from first peptide mapping experiments with in-house produced human norovirus VLPs using the P dimer HDX protocol yielded 93-100 % VP1 sequence coverage, in the same range as for P dimers (Figure 29). Therefore, it can be supposed that HDX-MS experiments for glycan binding to VLPs in solution will not require large methodological adjustments.
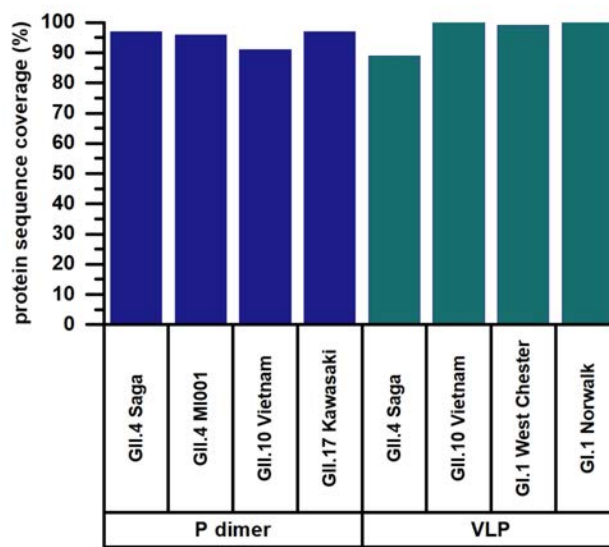


Figure 29: Human norovirus P dimer and VLP sequence coverage using the HDX-MS protocol for P dimers.

### 4.1.2   MEMBRANE PRESENTATION OF GLYCANS

So far, HDX-MS experiments have focused on binding of free glycans in solution, which does not completely resemble the situation *in vivo*. It would be therefore of high interest to analyze the effect of membrane association and binding to membrane-attached glycans. Membrane attachment could alter glycan specificity as some parts of the glycan may become sterically inaccessible. Furthermore, glycans displayed on a surface or lipid bilayer will only interact with parts of a viral capsid. Therefore, HDX-MS experiments can show if structural changes only occur at the glycan-interacting face of the capsid or if a local change is transmitted through the capsid, resulting in a global structural response to glycan binding. Glycan-display on surfaces will also allow probing multivalency effects similar to the cellular environment.

To perform these experiments glycolipids will need to be attached to a surface, which will be executed by a collaborator from the current research group. After binding to lipid-attached glycans and deuterium labeling, VLPs must be removed from the surface during the quench step. Here, quench conditions have to be screened in order to wash off particles optimally without introducing large amounts of background material into the sample. If considerable amounts of glycolipids get washed-off as well, they can be depleted in a similar way as in the HDX-MS workflow for membrane proteins.

### 4.1.3   INFECTIOUS NOROVIRUS PARTICLES

The ultimate goal would be to analyze structural dynamics in infectious viral particles in presence and absence of glycans, both in solution and presented on lipids or surfaces. This would give a complete picture of the structural response to glycan binding in a more native-like environment, which takes several factors like influence of all viral proteins and the genome, avidity and membrane interactions into account. When a HDX-MS workflow for infectious particles is established, the influence of environmental conditions and co-factors, such as pH changes, temperature, glycans, bivalent cations, bile acids and the murine protein receptor, on structural dynamics of the capsid can be investigated. These results can help to elucidate which structural changes are required to prime for host cell entry and provide valuable insights for drug and vaccine development. MNV can be grown in cell culture and is available from collaborators. In contrast, sufficient amounts of infectious human norovirus particles can so far only be obtained from patient stool samples. Therefore, initial HDX-MS method development for infectious particles needs to be done with MNV and later adapted to human norovirus, if samples become available. However, there are usually some challenges involved with the measurement of intact virions. It must be ensured that the virus is fully inactivated during the quench step, which necessitates the screening for an inactivation procedure that is compatible with HDX-MS. The most common inactivation protocols for MNV include heat inactivation, high organic disinfectants or other chemicals (Kamarasu et al., 2018), which would largely increase back exchange and be incompatible with either pepsin activity or MS analysis. Therefore, different quench buffer formulations need to be screened for complete MNV inactivation and HDX-MS compatibility. Preliminary results suggest that quench buffers containing 6 M

urea or 6 M guanidine hydrochloride fully inactivate MNV (Figure 30A), while 0.2 M TCEP alone is not sufficient. Particles inactivated with urea or guanidine hydrochloride containing quench buffers are also accessible for pepsin digestion (Figure 30B). Further MS measurements are required to decide which quench buffer formulation provides the best sequence coverage and if other sample preparation steps are necessary to reduce spectral complexity.
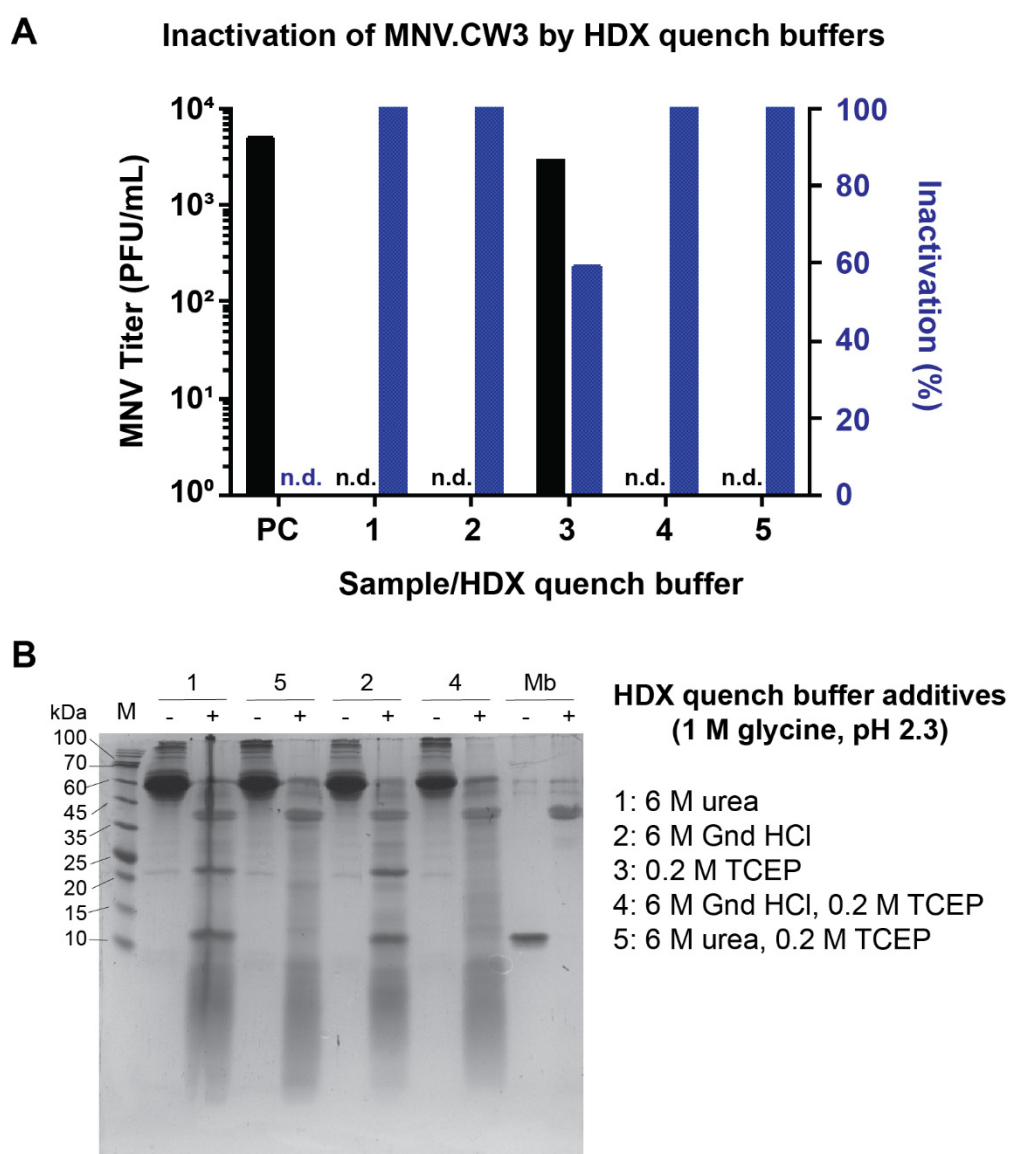


Figure 30: MNV inactivation and pepsin digestion. Infectious MNV.CW3 particle were mixed with equal amounts of different quench buffers and frozen in liquid nitrogen. (A) Sample was adjusted to pH 8 and a plaque assay with BV2 cells was performed to determine the MNV titer after mixing with quench buffers (black bars). Blue bars represent the inactivation relative to the positive control (PC) without quench buffer. (B) Fully inactivated MNV samples 1, 2, 4, and 5 were digested for 30 seconds in-solution (+ with, - without pepsin) and analyzed by SDS-PAGE. A digest of 10 µg myoglobin (Mb) served as positive control.

# 5 MATERIAL AND METHODS

## 5.1 Sample preparation

### 5.1.1 NOROVIRUS P DOMAINS

Norovirus P domains were synthesized and purified as described in Mallagaray et al. (2019). Briefly, *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, two His-tags, an HRV 3C cleavage domain, and the P domain. Due to the cloning strategy, the sequences of GII.4 Saga 2006, GII.4 MI001 and GII.17 Kawasaki 308 P domains contain an extra GPGS sequence preceding K225 (R228 for GV MNV07 and CR10), whereas GII.10 Vietnam 026 contains a GPG sequence preceding S224 (Table 6).

Table 6: Norovirus P domains. Molecular weight refers to the monomeric P domain and was calculated using the ProtParam tool (Wilkins et al., 1999).

| Norovirus strain | GenBank ID | VP1 amino acids (N terminal additions) | Molecular weight (kDa) |
|---|---|---|---|
| GII.4 Saga 2006 | AB447457 | 225-530 (GPGS) | 34.1 |
| GII.4 MI001 | KC631814 | 225-530 (GPGS) | 34.1 |
| GII.10 Vietnam 026 | AF504671 | 224-538 (GPGS) | 34.7 |
| GII.17 Kawasaki 308 | LC037415 | 225–530 (GPGS) | 34.1 |
| GV MNV07 | AET79296 | 228-530 (GPGS) | 33.4 |
| GV CR10 | ABU55613 | 228-530 (GPGS) | 33.4 |

Transformed cells were grown for 3 h at 37 °C. Overexpression was induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) at an $OD_{600}$ value of 1.5. Incubation was continued at 16 °C for 48 h. Cells were lysed in 75 mM phosphate buffer, 100 mM NaCl, pH 7.3 (human P domains) or 20 mM sodium acetate, 100 mM NaCl, pH 5.3 (murine P domains), 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 16/600 200 pg column (GE Healthcare) in 20 mM sodium phosphate buffer (pH 7.3) or 20 mM sodium acetate, 100 mM NaCl (pH 5.3) buffer for human and murine norovirus P domains, respectively. Protein purity and dimer concentration were monitored by SDS-PAGE and ultraviolet absorption. 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) to prevent further deamidation and eluted using a linear salt gradient.

Wild type P dimer samples were stored at 5 °C in the following buffers until analysis: GII.10 Vietnam: 25 mM Tris-HCl, 300 mM NaCl, pH 7.3; GII.17 Kawasaki, GII.4 Saga and GII.4 MI001 (the latter two as wild type N373): 20 mM sodium acetate, 100 mM NaCl pH 4.9. To create spontaneously deamidated (partially N373iD) GII.4 MI001 and GII.4 Saga P dimer (Mallagaray et al., 2019), wild type GII.4 MI001 P dimer was stored in 25 mM Tris-HCl, 300 mM NaCl, pH 7.3 and GII.4 Saga P dimer was stored in in 75 mM sodium phosphate buffer, 100 mM NaCl, pH 7.3 at 4 °C for several months. Murine P dimer samples were stored in 20 mM sodium acetate, 100 mM NaCl, pH 5.3.

## 5.1.2   MISTIC

Protein production and purification was done as illustrated in Hartmann et al. (2015). In brief, Mistic was produced from a previously described pET-30 EK/LIC expression vector containing a kanamycin resistance gene (Merck, Darmstadt, Germany) (Broecker et al., 2014, Jacso et al., 2013). The plasmid was co-transformed into the *Escherichia coli* expression strain BL21(DE3) (Agilent Technologies) with the plasmid pEvol PylRS containing the engineered Amber suppressor tRNA/synthetase system carrying a chloramphenicol resistance marker. Recombinant protein production was carried out according to the following published procedures (Tyagi and Lemke, 2013) with minor modifications. Briefly, cells were grown at 37 °C with shaking at 220 rpm in Terrific Broth medium supplemented with 25 µg/mL kanamycin and 30 µg/mL chloramphenicol. The cells were induced at $OD_{600}$ = 0.6 with 0.2 % arabinose and

1 mM IPTG and cultivated for a further 24 h at 37 °C. The protein was purified in LDAO (Sigma Aldrich, St. Louis, USA) by a combination of immobilized metal affinity chromatography, followed by His-tag removal by enterokinase cleavage, anion-exchange chromatography, and removal of enterokinase and uncleaved fusion protein by reverse immobilized metal affinity chromatography (Broecker et al., 2014, Jacso et al., 2013).

### 5.1.3 GLYCANS

Table 7: Overview of glycans. Methyl α-L-fucopyranoside and HBGA B trisaccharide were purchased from Carbosynth. Galactose was purchased from Sigma Aldrich.
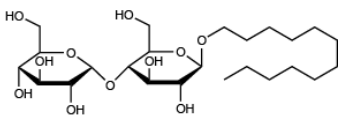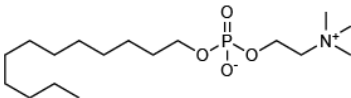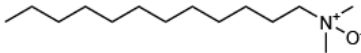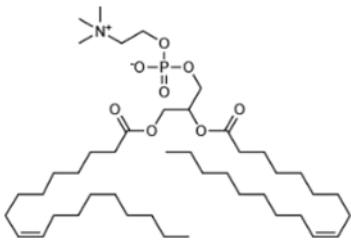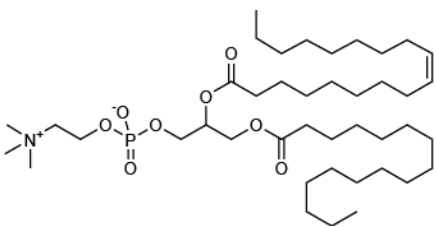
| Name (abbreviation) | Glycan nomenclature | Chemical structure |
|---|---|---|
| Methyl α-L-fucopyranoside (fucose) | α-L-Fuc-(1,O)-CH3 |  |
| HBGA B trisaccharide (HBGA B tri) | α-D-Gal-(1,3)-[α-L-Fuc-(1,2)]-β-D-Gal-(1,O)-CH3 |  |
| Galactose | D-(+)-Galactose |  |

### 5.1.4 LIPIDS AND DETERGENTS

Detergent micelles are widely used in membrane protein research to facilitate water solubility and maintain protein stability in absence of a native hydrophobic environment. The amphiphilic detergent structure consists of a polar head group that provides solubility and a nonpolar hydrocarbon moiety that mimics the

hydrophobic environment for the protein. At concentrations above the CMC (Preston, 1948), detergent molecules rapidly form micelles that are a prerequisite for the solubilization of membrane proteins. Similar to detergents, phospholipid molecules consist of a polar head group, but the nonpolar tail contains two acyl chains, which leads to a cylindrical shape of the molecule. Instead of micelles, these molecules form vesicles out of lipid bilayers, which mimic the natural lipid environment better than detergents. Table 8 gives an overview of the detergents and lipids used in this work.

Table 8: Overview of detergents and lipids

| Name (abbreviation) | Other names | Chemical structure |
|---|---|---|
| **Detergents** | | |
| *n*-dodecyl-β-D-maltoside (DDM) | |  |
| n-dodecyl phosphocholine (DPC) | C12-PC |  |
| N,N-dimethyldodecyl-amine N-oxide (LDAO) | Lauryldimethylamine *N*-oxide |  |
| **Lipids** | | |
| 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) | Dioleoyl-phosphatidylcholine, 18:1 ($\Delta^9$-*cis*) PC |  |
| 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) | Palmitoyloleoyl-phosphatidylcholine, 16:0-18:1 PC |  |

## 5.1.5    VESICLE PREPARATION AND MISTIC RECONSTITUTION

LUV preparation as well as Mistic reconstitution was performed prior to HDX-MS experiments. The desired lipid (Avanti Polar lipids) was first dissolved in a 2:1 chloroform/methanol mixture, then dried under gentle nitrogen flow, followed by vacuum drying for more than 1.5 h. The lipid cake was hydrated using reconstitution buffer (50 mM Tris, 50 mM NaCl, pH 7.4), yielding a 20 mM lipid concentration. The lipid suspension was vortexed for 5 min, followed by ten freeze-thaw cycles. After being incubated for 1 h while shaking, the solution was extruded 41 times through a 100 nm polycarbonate filter using a Mini-Extruder (Avanti Polar Lipids) as indicated in the manufacturer's protocol (AvantiPolarLipids, 2020).

The unimodal vesicle size distribution of the vesicles was confirmed by DLS (Wyatt technology). Reconstitution of Mistic was performed by gradually replacing detergent micelles with lipid vesicles. 25 µL of 63 µM Mistic, solubilized in 6 mM LDAO, was mixed with 25 µL of 40 mM lipid in vesicle suspension, resulting in 50 µL 10 mM lipid vesicles decorated with LDAO at sub-solubilizing detergent concentrations (3 mM) and a 31.5 µM concentration of the protein (1269:1 lipid:protein ratio). The mixture was incubated at room temperature for 45 min during which the sample was diluted stepwise to reach a LDAO free protein-vesicle concentrate far below the critical micellar concentration (CMC LDAO = 1.5 mM). In other words, every 15 min 50 µL of buffer were added until 200 µL were reached. After the stepwise dilution, the sample contained 7.8 µM Mistic, 5 mM lipid and 0.75 mM LDAO. To quantitatively remove LDAO from LUVs, samples were dialyzed two times against 50 mM Tris and 50 mM NaCl pH 7.4 in a dialysate volume of 1000 x sample volume with a membrane cut-off of 3.5 kDa. (first exchange after 2 h, second exchange overnight). The final sample for HDX-MS contained ~200 µL of 7.8 µM Mistic and 5 mM lipid.

## 5.2 HDX-MS

### 5.2.1 SAMPLE PREPARATION AND LABELING

Samples were 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 then frozen in liquid nitrogen.

EXPERIMENT-SPECIFIC METHODOLOGICAL DETAILS
The following experiment-specific adjustments were made to the general sample preparation workflow described above:

**Chapter 3.1**
Protein: GII.4 Saga P dimer (wild type and deamidated)
Ligands: HBGA B trisaccharide, fucose, galactose
Reference: Mallagaray et al. (2019)

Wild type and deamidated P dimer (50 pmol) were mixed with glycans at tenfold of the final concentration (final: 10 mM HBGA B trisaccharide, 100 mM fucose, 100 mM galactose as negative control) and directly diluted in deuterated buffer. After 1 min, 10 min, 1 h and 8 h the exchange reaction was quenched as described above.

**Chapter 3.2**
Protein: GII.10 Vietnam 026 P dimer, GII.17 Kawasaki 308 P dimer, GII.4 Saga P dimer (partially deamidated), GII.4 MI001 P dimer (wild type and partially deamidated)
Ligands: HBGA B trisaccharide, fucose
Reference: Dülfer et al. (in preparation)

P dimer (30-50 pmol) was mixed with glycans at tenfold of the final concentration (final: 10 mM HBGA B trisaccharide, 100 mM fucose) and directly diluted in deuterated buffer. After 1 min, 10 min, 1 h and 8 h (15 s, 1 min, 10 min for partially deamidated GII.4 Saga P dimer) the exchange reaction was quenched as

described above. As a FD control, P dimers were diluted 1:9 in 99 % deuterated 20 mM Tris buffer with 150 mM NaCl and 6 M urea at pH 7, labeled for 24-72 h at room temperature, and then quenched as described above.

**Chapter 3.4.1**

Protein: Mistic in POPC, DOPC and LDAO

Ligands: -

Reference: -

50 pmol Mistic (6.4 µL of 7.8 µM) were mixed with 43.6 µL of undeuterated Tris buffer (the same one as used for reconstitution). This would result in ≈ 87 % deuterium labeling when deuterated buffer is used in the HDX experiment. For screening of quench and digestion conditions two undeuterated samples were prepared for every combination of lipid (POPC and DOPC) and screening condition.

**Quench buffers:**

Q1: 2 M glycine, 8 M urea, 2 mM DDM, pH 2.4

Q2: 2 M glycine, 8 M urea, 16 mM DPC, pH 2.4

Q3: 2 M glycine, 8 M urea, pH 2.4

**Screening conditions:**

1. **On-column digestion without lipid removal:**

   50 µL of each quench buffer was added to the sample and directly frozen in liquid nitrogen

2. **In-solution digestion without lipid removal:**

   47 µL of each quench buffer was added to the sample, the sample was placed on ice and 3 µL of 334 µM pepsin solution (in 2 M glycine, pH 2.4) was added (protein:pepsin = 1:200). After 2 min of digestion time, samples were frozen in liquid nitrogen.

3. **On-column digestion with lipid removal:**

   50 µL of each quench buffer was added to the sample, transferred to a filter with lipid removal beads (HybridSPE-Phospholipid beads, Merck) and centrifuged for one minute at 1 °C. The filter was removed and the flow through directly frozen in liquid nitrogen.

4. **In-solution digestion with lipid removal:**

47 µL of each quench buffer was added to the HDX sample, the sample was placed on ice and 3 µL of 334 µM pepsin solution (in 2 M glycine, pH 2.4) was added to the sample (protein:pepsin = 1:200). After 1 min of digestion time, samples were transferred to a filter with lipid removal beads and centrifuged for one minute at 1 °C. The filter was removed and the flow through directly frozen in liquid nitrogen.

Furthermore, the Mistic stock solution in LDAO was diluted to 50 pmol in Tris buffer and then mixed with Q3 quench buffer for comparison.

**Chapter 3.4.2**

Protein: Mistic in POPC, DOPC and LDAO

Ligands: -

Reference: Krainer et al. (in preparation)

50 pmol Mistic in POPC, DOPC and LDAO was diluted 1:9 in 99 % deuterated 50 mM Tris buffer (pH 7.4, 50 mM NaCl, 25 °C, additionally 3 mM LDAO for LDAO sample) to start the exchange reaction. After 30 s, 2 min, 10 min and 1 h the exchange reaction was quenched by 1:1 addition of ice-cold quench buffer (1 M glycine, 8 M urea, pH 2.1), which decreased the pH to 2.3. The samples were transferred to a filter with lipid removal beads (HybridSPE-Phospholipid beads, Merck) and centrifuged 30 s at 1 °C. The filter was removed and the flow through directly frozen in liquid nitrogen. All time points were performed in triplicate.

Two FD controls were prepared by the following protocols respectively:

1.) 123 µM Mistic in 6 mM LDAO was diluted to 10 µM in 50 mM Tris undeuterated and LDAO-free buffer with 50 mM NaCl at pH 7.4. The protein was then further diluted 1:10 in 99 % deuterated buffer (50 mM Tris, 50 mM NaCl, pH 7.4) and labeled for 24 h at room temperature. The exchange reaction was quenched by 1:1 addition of ice-cold quench buffer (1 M glycine, 8 M urea, pH 2.1). The samples were transferred to a filter with lipid removal beads and centrifuged 30 s at 1 °C. The filter was removed and the flow through directly frozen in liquid nitrogen.

2.) 123 µM Mistic in 6 mM LDAO was diluted to 10 µM in 50 mM Tris undeuterated and LDAO-free buffer with 50 mM NaCl and 6 M urea at pH 7.4 and incubated for 2 h at room temperature. The protein was then further diluted 1:10 in 99 % deuterated buffer (50 mM Tris, 50 mM NaCl, pH 7.4) and labeled for 24 h at room temperature. The exchange reaction was quenched by 1:1 addition of ice-cold quench buffer (1 M glycine, 8 M urea, pH 2.1). The samples were transferred to a filter with lipid removal beads and centrifuged 30 s at 1 °C. The filter was removed and the flow through directly frozen in liquid nitrogen.

For comparison, 50 pmol Mistic in 6 mM LDAO was diluted 1:9 in 99 % deuterated 50 mM Tris buffer (pH 7.4, 50 mM NaCl, 25 °C, 3 mM LDAO) and labeled for one week at room temperature. The exchange reaction was quenched by 1:1 addition of ice-cold quench buffer (1 M glycine, 8 M urea, pH 2.1). The samples were transferred to a filter with lipid removal beads and centrifuged 30 s at 1 °C. The filter was removed and the flow through directly frozen in liquid nitrogen.

## BACK EXCHANGE CONTROLS

25 pmol Angiotensin I and Bradykinin peptides (both Alfa Aesar) were labeled for 24 h in 20 mM Tris pH 7 with 6 M guanidine hydrochloride and quenched with 300 mM phosphate buffer, pH 2.3. Two replicates were directly frozen in liquid nitrogen and two replicates were transferred to a filter with lipid removal beads and centrifuged 30 s at 1 °C. The filter was removed and the flow through directly frozen in liquid nitrogen.

## 5.2.2 HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC) AND MS

### HPLC SETUP

An Agilent Infinity 1260 HPLC system (Agilent Technologies) equipped with two pumps, a solvent degasser, a column oven and two valves was used to perform protein digestion and peptide separation for HDX-MS experiments. To maintain a low temperature under quench conditions, the whole HPLC system is placed into a custom-modified fridge with a closable opening on the left site for manual sample injection. The system is equipped with an in-house packed pepsin column (IDEX guard column with an internal volume of 60 µL, Porozyme immobilized

pepsin beads, Thermo Scientific) in the 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).

The LC flow path was designed similar to standard bottom-up HDX-MS setups (Wang et al., 2016). The sample is loaded manually into the injection loop and then injected into the system by a manual switch of the injection valve to the inject position. For protein digestion as well as peptide desalting and trapping, the isocratic pump is in line with the pepsin column and the trap column (Figure 31 A). After trapping, the second valve is automatically switched to the analytical position, which puts the trap column in line with the reverse-phase analytical column (Figure 31 B). While peptides are eluted from the trap and separated on the analytical column by an organic solvent gradient, the pepsin column is washed and equilibrated for the next run. After each sample, a specific wash method is run for cleaning and equilibration of the trap and analytical column.

## LC-MS METHOD FOR HDX-MS

The HPLC system was cooled to 2 °C and pepsin digestion of deuterated samples was performed at a flow rate of 75 µL/min for 3 min in solvent A. Peptides were eluted and separated on the analytical column using a 7 min gradient of 8-40 % solvent B (solvent A: 0.4 % formic acid (FA) in water, solvent B: 0.4 % FA in acetonitrile (ACN)) at 150 µL/min. MS was performed on an Orbitrap Fusion Tribrid instrument (Thermo Scientific) in positive ESI MS-only mode (for detailed settings see Table S 3) calibrated on the same day using Pierce™ LTQ Velos ESI Positive Ion Calibration Solution (Thermo Scientific). All HDX-MS measurements were executed in three technical replicates. Specific settings for the individual experimental setups can be found below.

## LC-MS METHOD FOR PEPTIDE AND PTM IDENTIFICATION

The HPLC system was cooled to 5 °C and pepsin digestion of undeuterated samples was performed at a flow rate of 75 µL/min for 3 min in solvent A. Peptides were eluted and separated on the analytical column using a 27 min elution gradient of 8-40 % solvent B. MS was performed on an Orbitrap Fusion Tribrid instrument (Thermo Scientific) in positive ESI data-dependent MS/MS acquisition mode with higher-energy collisional dissociation (HCD) (for detailed settings see Table S 4) calibrated on the same day using Pierce™ LTQ Velos ESI Positive Ion Calibration Solution (Thermo Scientific).

**A**
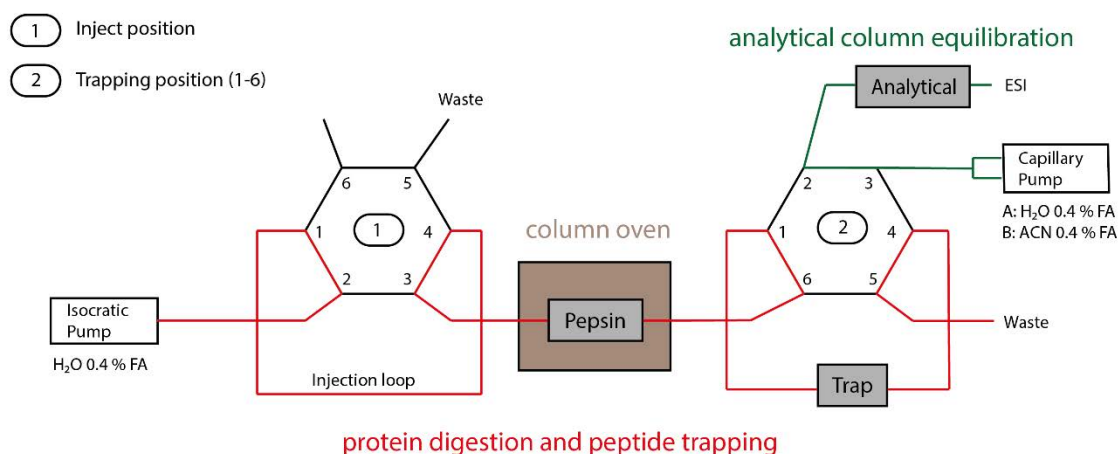


**B**
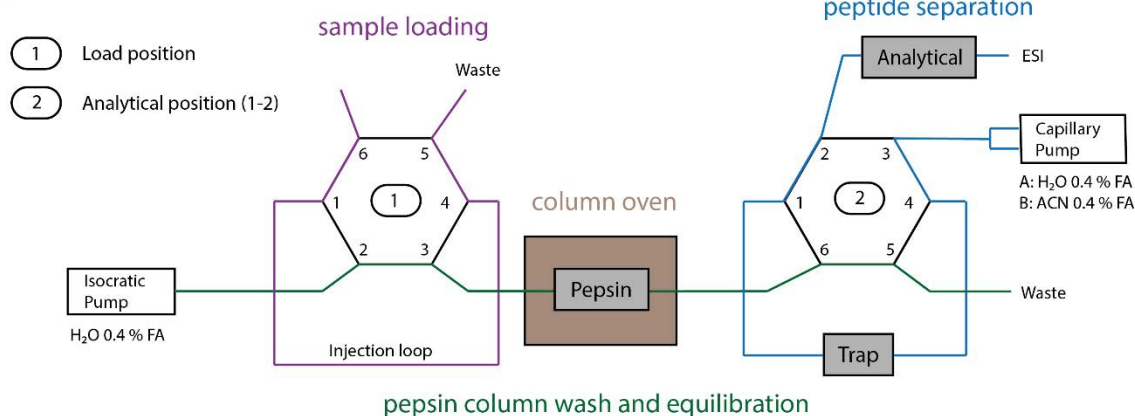


Figure 31: LC flow path for HDX-MS experiments. (A) Protein digestion and peptide trapping mode. After sample loading in "load" position (panel B), the manual injection valve (1) is turned to the "inject" position. The second valve (2) is in "trapping" position. The isocratic pump is in line with the pepsin and trap column. The protein solution is injected into the LC-system, digested in the pepsin column and the resulting peptides are trapped and desalted in the trap column. During this step, the analytical column is equilibrated. (B) Peptide separation mode. Valve 2 is switched to "analytical" position and the capillary pump is put in line with the trap and analytical column. Peptides are eluted from the trap and separated on the analytical column by an organic solvent gradient. Following separation, peptides are ionized in the ion source of the mass spectrometer by ESI. In parallel, the pepsin column can be washed and re-equilibrated for the next run.

## EXPERIMENT-SPECIFIC METHODOLOGICAL DETAILS

The following experiment-specific adjustments were made to the general LC-MS workflow described above:

### Chapter 3.1

Protein: GII.4 Saga P dimer (wild type and deamidated)

Ligands: HBGA B trisaccharide, fucose, galactose

Reference: Mallagaray et al. (2019)

Measurements were executed as described above with the following deviations:

HPLC system temperature for HDX measurements: 0 °C

Solvent A: 0.23 % formic acid in water

Solvent B: 0.23 % formic acid in acetonitrile

The wild type and deamidated datasets with B trisaccharide were acquired on two consecutive days and all time points were performed in triplicates. The wild type dataset with B trisaccharide, fucose and galactose contains single measurements performed on the same day.

### Chapter 3.2

Protein: GII.10 Vietnam 026 P dimer, GII.17 Kawasaki 308 P dimer, GII.4 Saga P dimer (partially deamidated), GII.4 MI001 P dimer (wild type and partially deamidated)

Ligands: HBGA B trisaccharide, fucose

Reference: Dülfer et al. (in preparation)

All time points were performed in three technical replicates, apart from the 8 h time point of GII.10 Vietnam with fucose, which only represents a single measurement. The triplicate measurement of GII.4 MI001 P dimer was influenced by peptide carry over, which overlaid with the lower deuterated peak distribution and resulted in a falsely high intensity. Therefore, a separate single-replicate measurement with additional pepsin column washing (2 M urea, 2 % ACN, 0.4 % FA, pH 2.5) between sample injections was performed and only deuteration differences, which are present in both datasets, are considered real.

**Chapter 3.4.1**

Protein: Mistic in POPC, DOPC and LDAO

Ligands: -

Reference: -

Online digestion was performed at varying flow rates. For proteins digested in-solution, the pepsin column was removed. In between the samples, wash runs were performed with high percentage of solvent B on the analytical and trap column. The pepsin column was washed by injection of a wash solution (2 M urea, 2 % ACN, 0.4 % FA, pH 2.5).

**Chapter 3.4.2**

Protein: Mistic in POPC, DOPC and LDAO

Ligands: -

Reference: Krainer et al. (in preparation)

The pepsin column was washed by injection of a wash solution (2 M urea, 2 % ACN, 0.4 % FA).

## 5.2.3   ORBITRAP FUSION MASS SPECTROMETER

The Orbitrap Fusion mass spectrometer (Thermo Scientific) used for HDX-MS experiments is equipped with three mass analyzers (Quadrupole, linear ion trap and Orbitrap, so-called "Tribrid" architecture) that can be combined in different ways depending on the type of experiment. The high resolution provided by the Orbitrap mass analyzer together with the various modes of peptide fragmentation allow detection and identification of peptides with high confidence. In the following paragraphs, the instrument setup (Figure 32) and functions of the main instrument parts will be described in more detail.

### H-ESI SOURCE

The Orbitrap mass spectrometer is equipped with a metal capillary that is connected to the LC system and introduces the sample solution into the heated electrospray (H-ESI) ion source in positive-ion polarity mode. There, the solution is ionized as described in introductory chapter 1.1.2. The heated auxiliary gas supports solvent evaporation. The gas-phase sample ions then enter the mass

spectrometer through the ion transfer tube and are captured and focused by an ion guide (S-lens). They are then transferred into the quadrupole mass filter by the active beam guide, which reduces noise by preventing neutral ions from entering the quadrupole (Eliuk and Makarov, 2015).
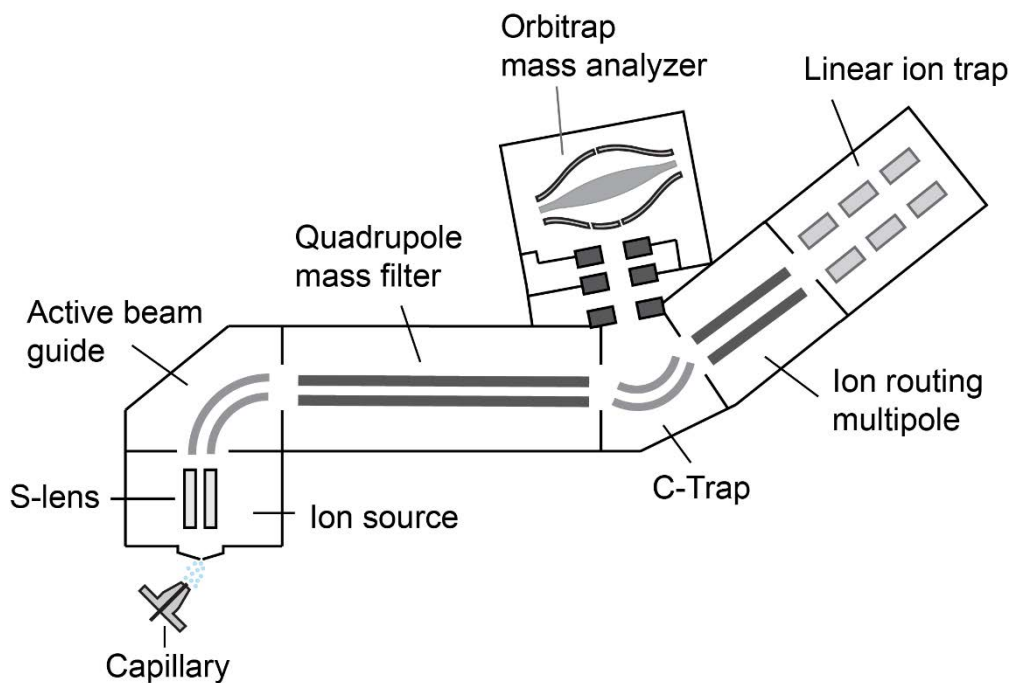


Figure 32: Orbitrap Fusion instrument setup based on Eliuk and Makarov (2015). The capillary transports the sample solution from the LC system into the ion source region, where it is transferred into the gas-phase and enters the mass spectrometer. The generated sample ions are focused by the S-lens and transferred into the quadrupole via the active beam guide. The quadrupole can be used for ion transmission or as a mass filter. Ions can then cycle back and forth between the Orbitrap mass analyzer and the linear ion trap via the C-Trap and the ion routing multipole depending on the measurement mode (e.g. MS-only or MS/MS) and experimental setup.

## QUADRUPOLE MASS FILTER

The quadrupole mass analyzer is able to discriminate and filter ions of different $m/z$. It consists of four parallel metal rods that are connected with each other so that when a potential is applied, ions oscillate in the electrical field of the quadrupole. Depending on the applied radio frequency (RF) and direct current (DC) potential, only ions with a specific $m/z$ will have stable trajectories, while other ions collide into the rods and are filtered out. Quadrupoles can also be used in RF-only mode, which allows all ions to pass through, thus acting as ion

transmitter to other parts of the MS instrument (Chernushevich et al., 2001, Haag, 2016).

## C-TRAP AND ORBITRAP MASS ANALYZER

The Orbitrap is an ion trap mass analyzer that consists of an outer barrel-like electrode and an inner spindle-like electrode, where ions are trapped in orbital motion in the electric field (Zubarev and Makarov, 2013). Ions of a large $m/z$ range are first focused in a curved linear trap (C-trap) that contains a bath gas for collisional cooling of the ions (Nolting et al., 2017). For injection, RF voltages in the C-Trap are rapidly ramped down and DC voltages are applied which leads to the release of ion packets into the Orbitrap (Eliuk and Makarov, 2015). The injection takes place at a peripheral position (different from the center of the spindle) which causes a coherent axial oscillation of ions in the electric field of the Orbitrap (Perry et al., 2008). By ramping up the potential of the inner electrode during injection, ions get squeezed towards the inner electrode and start to cycle around it on elliptical trajectories (Makarov, 2000)(Figure 33 A).
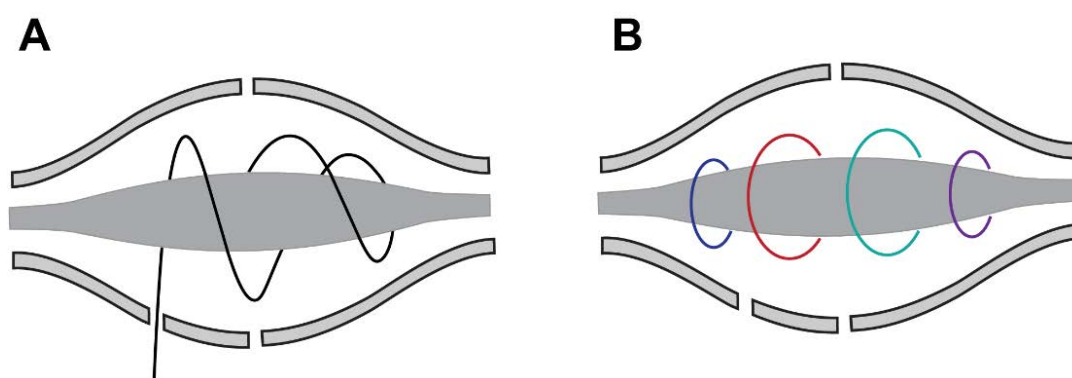


Figure 33: Ion motion in the Orbitrap mass analyzer. (A) Ions are injected at a position with some distance to the center of the central spindle electrode and start to cycle around it on elliptical trajectories. (B) When the field becomes static, ions with the same axial frequencies and thus same $m/z$ spread into rings and oscillate around the spindle electrode. The axial oscillations induce a current in the outer electrodes that is detected and transformed into a mass spectrum.

For detection, the field becomes static, so that ions with different rotational frequencies but the same axial frequency spread into rings that oscillate along the inner spindle electrode (Figure 33 B). The axial motion is independent of all initial parameters of the ions except for the $m/z$ (Perry et al., 2008). The axial oscillations of ion rings are then detected by the image current they induce on the outer electrode. The image current in the time domain is then Fourier-

transformed into the frequency domain and converted into a mass spectrum (Zubarev and Makarov, 2013).

## ION ROUTING MULTIPOLE

The ion routing multipole can serve several functions, e.g. ion storage, ion transfer and fragmentation. It works in a similar way as the quadrupole in ion transmission (RF-only) mode, which allows storage and stable passage of ions between the ion trap and Orbitrap mass analyzer. It can also act as collision cell for fragmentation by HCD (Eliuk and Makarov, 2015). This term is used due to the fact that it uses higher energies than those achievable for CID in the linear ion trap. In HCD the ions pass through the C-trap, enter an acceleration gap and get fragmented in the ion routing multipole in a similar way as in CID. Fragment ions are then trapped and cooled inside the cell and sent to a mass analyzer. It provides the advantage of collecting all fragments without low-mass cut-off as in the ion trap (Nolting et al., 2017, Olsen et al., 2007).

## LINEAR ION TRAP

Depending on the measurement method, the linear ion trap can be used for ion storage, selection, fragmentation or detection. It consist of a set of quadrupole rods with a two-dimensional RF field for radial- and stopping potentials at the end of the electrodes for axial ion confinement (Douglas et al., 2005). Depending on the electrical parameters applied, ions of a certain $m/z$ range can be stored in the ion trap. In order to select ions and act as a mass analyzer, ions of specific $m/z$ can be selectively ejected from the trap by varying the RF potential and subsequently detected on a detector plate outside the trapping volume. The detector signal is measured as function of the RF amplitude which correlates directly with the $m/z$ of the detected ion (Nolting et al., 2017). Due to its ability to accumulate ions over time, ions can be detected with high sensitivity, but the resolution and mass accuracy is lower than in the Orbitrap analyzer. In the Orbitrap Fusion instrument, the linear ion trap consists of a high-pressure cell for precursor ion isolation and a low-pressure cell for improved scan speed and mass accuracy. The ion trap can also be used to generate fragment ions by CID or ETD.

## 5.2.4 DATA ANALYSIS AND STATISTICS

### IDENTIFICATION OF PEPTIDES AND PTMS

Precursor and fragment ions were searched and matched against a local protein database just containing the protein of interest in MaxQuant (Max-Planck-Institute of Biochemistry, version 1.5.7.0) (Tyanova et al., 2016). The digestion mode was set to "unspecific" and *N*-terminal acetylation, deamidation, oxidation and disulfide bond formation were included as variable modifications with a maximum number of five modifications per peptide. Peptides between 5 and 30 amino acids length were accepted. The MaxQuant default mass tolerances for precursor (4.5 ppm) and fragment (20 ppm) ions defined for the Thermo Orbitrap instrument were used for data search. The minimum score for successful identifications was set to 20 for unmodified and 40 for modified peptides. Thermo Xcalibur 3.0.63 software (Thermo Scientific) was used for manual data inspection. For norovirus P domain peptides carrying a deamidation, chromatographic peak areas where calculated in Xcalibur to obtain a wild type/deamidated peptide ratio.

### EXPERIMENTAL DESIGN AND STATISTICAL RATIONALE

The rationale for experimental design and data analysis is based on HDX-MS community-recommendations (Masson et al., 2019). In brief, sample quality was assessed with native MS and HDX-MS conditions were optimized for maximum sequence coverage and detection sensitivity. Labeling time points were chosen to cover 3-4 orders of magnitude. Three independent labeling reactions were performed for each time point if not stated otherwise and the level of back exchange was assessed with a fully deuterated protein control as well as a mix of deuterated model peptides. Details about the peptide identification method, statistical analysis with Student's t-test and color mapping procedure are given in the individual methods sections. Fragmentation spectra for identification of deamidated norovirus P domain peptides are given in the supplement (Figure S 14 - Figure S 20). All HDX-MS data has been manually inspected and exchange differences in bimodal peak distributions have been validated by binomial fitting.

## CALCULATION OF DEUTERIUM CONTENT

DeutEx software (obtained from Petr Man, Prague) was used to determine the deuterium uptake via centroid analysis. To convert Thermo raw data into the required format for DeutEx, peaks were picked in the open-source MZmine 2.31 software (Pluskal et al., 2010) and peak lists were exported with a custom-built export function that was integrated in MZmine. HX-Express2 (Guttman et al., 2013, Weis et al., 2006a) was used for deconvolution of bimodal peak distributions by binomial fitting.

## STATISTICAL TESTS

For all datasets containing triplicate data, a two-sided Student's t-test was applied for each peptide and labeling time point for assessment of statistical significant differences between two states. The Student's t-test was either performed in an Excel spreadsheet (Microsoft) or GraphPad Prism (GraphPad Software, Inc., version 5.03). Additionally, a minimum threshold for deuteration differences ($\Delta D$) was calculated for each dataset based on published procedures (Arora et al., 2015, Hageman and Weis, 2019a, Houde et al., 2011). Only if a deuteration difference at a certain time point passes the Student's t-test and is larger than $\Delta D$, the difference is considered significant. Specific details about this procedure can be found in the experiment specific data analysis details. Furthermore, a significant deuteration difference based on the statistical criteria must be present in several overlapping peptides or at least several time points of one peptide to be considered significant.

## DATA REPORTING AND VISUALIZATION

Based on the HDX-MS community recommendations (Masson et al., 2019) detailed information about the experimental conditions, statistics and HDX-MS data quality parameters for all experiments are provided as HDX summary tables (Table S 5 - Table S 7). Sequence coverage, average peptide length and redundancy was calculated using an Excel spreadsheet. Peptide coverage maps, indicating the effective peptide coverage in each HDX experiment, were plotted with MS Tools (Kavan and Man, 2011) and can be found in supplement (Figure S 27 - Figure S 33). Bar graphs, deuterium difference plots and bimodal spectra were plotted with GraphPad Prism (version 5.03).

## EXPERIMENT-SPECIFIC DATA ANALYSIS DETAILS

The following experiment-specific adjustments were made to the general data analysis workflow described above:

**Chapter 3.1**

Protein: GII.4 Saga P dimer (wild type and deamidated)

Ligands: HBGA B trisaccharide, fucose, galactose

Reference: Mallagaray et al. (2019)

For comparison of triplicate data for HBGA B trisaccharide binding, a Student's t-test was used with the α-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 cut-off of $\Delta D > 0.64$ (99 % percentile calculated according to Arora et al. (2015) was used to account for possible day to day variation in the experimental conditions. HDX-MS of native GII.4 Saga P dimers with fucose and galactose was performed as single measurements.

**Chapter 3.2**

Protein: GII.10 Vietnam 026 P dimer, GII.17 Kawasaki 308 P dimer, GII.4 Saga P dimer (partially deamidated), GII.4 MI001 P dimer (wild type and partially deamidated)

Ligands: HBGA B trisaccharide, fucose

Reference: Dülfer et al. (in preparation)

For comparison of triplicate data, a two-sided Student's t-test using deuteration differences from centroid analysis was used with the α-value set to 0.05. A peptide was only considered to have a significant HDX difference if the peptide passed the t-test and $\Delta D$ exceeded 2 x the pooled average standard deviation (Hageman and Weis, 2019a, Houde et al., 2011) of the dataset either for several time points or for the same time point in overlapping peptides. For some peptides, deuteration of FD controls was lower than deuteration of the 8 h labeling time point. Therefore, datasets were not normalized to the absolute FD deuterium uptake and only relative differences between states are presented.

For comparison of the unbound wild type and deamidated MI001 P dimer, the ratio of the FD controls from both measurements was used for normalization. Additionally, a higher cut-off of $\Delta D > 0.42$ (99 % percentile calculated according to (Arora et al., 2015)) was used to account for possible day-to-day variation in the experimental conditions.

Deuterated spectra of peptides in certain protein regions showed bimodal peak distributions that led to lower deuteration values in centroid analysis. To validate the deuteration differences observed in centroid data analysis and to calculate relative intensities of both peak distributions, bimodal spectra of peptides representative for certain regions were analyzed by binomial fitting in HX Express (Guttman et al., 2013). To compare relative intensities of both distributions in different states, an average over all bimodal time points for both distributions in each state was calculated for several peptides. Averaged relative intensities of different peptides are presented as bar plots in Figure 16. Relative intensity differences of the first peak distribution in different states were analyzed with a two-sided Student's t-test for each pair of states in each experiment (p < 0.05).

**Chapter 3.4.2**

Protein: Mistic in POPC, DOPC and LDAO

Ligands: -

Reference: Krainer et al. (in preparation)

For comparison of triplicate data, a two-sided Student's t-test using deuteration differences from centroid analysis was used with the α-value set to 0.05. A peptide was only considered to have a significant HDX difference if the peptide passed the t-test and $\Delta D$ exceeded 2 x the pooled average standard deviation (Hageman and Weis, 2019a, Houde et al., 2011) of the dataset either for several time points or for the same time point in overlapping peptides.

Deuterated spectra of peptides in certain protein regions showed bimodal peak distributions. To validate the deuteration differences observed in centroid data analysis and to calculate relative intensities of both peak distributions, bimodal spectra of peptides representative for certain regions were analyzed by binomial fitting in HX Express (Guttman et al., 2013).

## 5.3 Native MS

### 5.3.1 LCT TOF MASS SPECTROMETER

The high-mass modified LCT mass spectrometer (Micromass) is equipped with an ion source for ESI, a hexapole for ion transmission and a Time-of-Flight (ToF) analyzer for ion detection (Figure 34). The modifications include an increased pressure in the source region, the use of argon in the hexapole and adjustments in the electronics for transmission of high mass ions, in a similar way as described for QToF instruments in van den Heuvel et al. (2006).
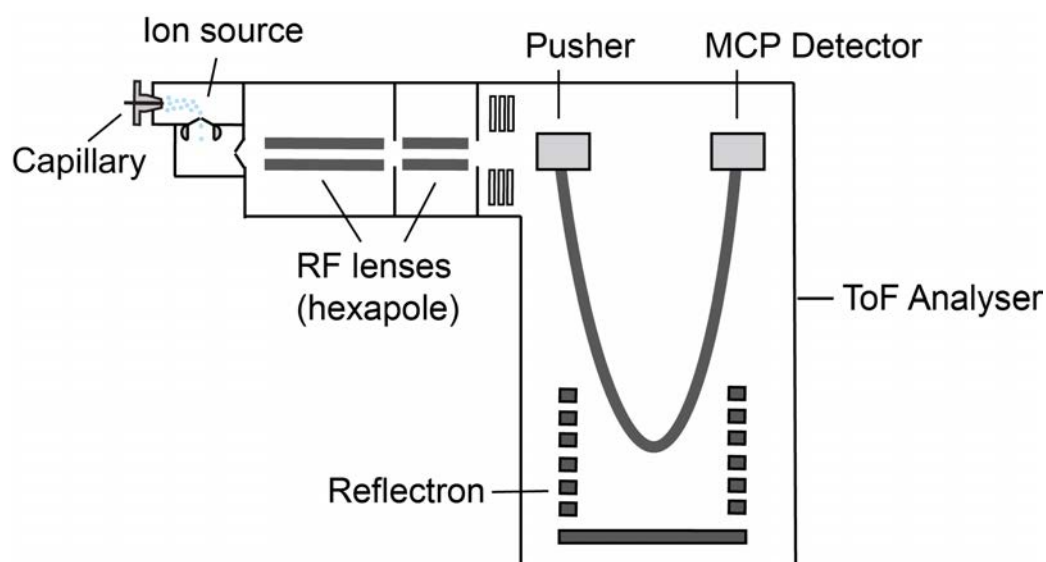


Figure 34: Schematic representation of the LCT mass spectrometer used for native MS measurements.

Ions generated in the ion source are transferred to the time of flight analyzer via the hexapole region. In the pusher, the ion beam is converted into ion packages by pulses applied orthogonal to the ions flight path. Ion packages are then pushed towards the reflectron, which reflects ions back to the detector. The use of a reflectron increases the path length in compact instrument setups and therefore improves separation of ions with different $m/z$ (Mamyrin et al., 1973). As ions travel from the pusher to the detector they are separated in mass according to their flight times, with ions of the highest $m/z$ arriving latest at the detector. Once an ion hits the multi-channel plate (MCP) detector, the signal is detected and its flight time is converted into an $m/z$ ratio.

## 5.3.2   SAMPLE PREPARATION AND MS MEASUREMENT

Proteins were subjected to buffer exchange to various concentrations of ammonium acetate at 4 °C via centrifugal filter units (13,000 x g, VivaSpin 500, MWCO 10,000 (Sartorius) or Micro Bio-Spin 6 columns (Bio-Rad) according to the manufacturer's protocol. Mass spectra were acquired at room temperature (25 °C) in positive ion mode on an LCT mass spectrometer modified for high mass (Waters, UK and MS Vision, the Netherlands) with a nano-electrospray ionization source. Gold-coated electrospray capillaries were produced in house for direct sample infusion. Capillary and sample cone voltages are given in the experiment-specific sections. The pusher was set to 100-150 µs. Pressures were 7 mbar in the source region and 6.2 x $10^{-2}$ to 6.5 x $10^{-2}$ mbar argon in the hexapole region. A spectrum of a 25 mg/mL cesium iodide solution from the same day was applied for calibration of raw data using the MassLynx software (Waters, UK).

### EXPERIMENT-SPECIFIC METHODOLOGICAL DETAILS

The following experiment-specific adjustments were made to the general sample preparation workflow described above:

**Chapter 3.2**

Protein: GII.10 Vietnam 026 P dimer, GII.17 Kawasaki 308 P dimer, GII.4 Saga P dimer (fully deamidated), GII.4 MI001 P dimer (wild type and partially deamidated)

Reference: Dülfer et al. (in preparation)

Native MS measurements were performed using 3 to 4.5 µM purified P dimers. Proteins were subjected to buffer exchange to 125 to 300 mM ammonium acetate at pH 7.5. Capillary and sample cone voltages were 1.20 kV to 1.35 kV and 150 to 240 V, respectively. Peak integration and calculation of oligomer fractions was performed with OriginPro 2016 software (Origin Lab Corporation).

**Chapter 3.3**

Protein: GII.4 Saga P dimer (wild type), murine CR10 P dimer, MNV07 P dimer

Reference: Creutznacher et al. (in preparation)

Native MS measurements were performed using 2 to 92 µM purified P domains. Proteins were subjected to buffer exchange to 150 mM ammonium acetate at pH 5.3. Capillary and sample cone voltages were 1.30 kV and 200 V, respectively. Peak intensities of oligomers were summed over the respective *m/z* region and relative intensities were calculated using Microsoft Excel. The dissociation constant $K_d$ was calculated using OriginPro 2016.

## 5.4   Mathematical methods

### 5.4.1   GLYCAN BINDING SITE OCCUPANCY

To choose a glycan concentration that leads to high P dimer binding site occupancy under HDX conditions, the binding site occupancy during equilibration and labeling conditions can be calculated using the law of mass action. The total protein and ligand concentration ($[P_{total}]$ and $[L_{total}]$) is the sum of the free components and the bound components:

Equation 6 $\qquad P_{total} = P_{free} + PL \qquad$ and $\qquad L_{total} = L_{free} + PL$

With a one-site binding model, the ligand binding can be described as follows:

Equation 7

$$P_{free} + L_{free} \underset{K_d}{\overset{K_a}{\rightleftharpoons}} PL$$
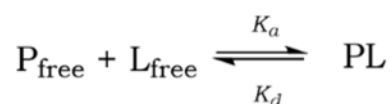
and the dissociation constant $K_d$ of this reaction is:

Equation 8 $\qquad K_d = \dfrac{[P]_{free}\,[L]_{free}}{[PL]}$

Substitution of $P_{free}$ using Equation 6 results in:

Equation 9 $\qquad K_d = \dfrac{([P_{total}]\text{-}[PL])\,([L_{total}]\text{-}[PL])}{[PL]}$

Solving this equation for [PL] gives a second order equation, where only the solution with negative sign before the square root is relevant in a chemical

context. The ligand occupied protein fraction ($[PL]/[P_{total}]$) can then be calculated as follows:

Equation 10
$$\frac{[PL]}{[P_{total}]} = \frac{(P_{total}+L_{total}+K_d) - \sqrt{(P_{total}+L_{total}+K_d)^2-4P_{total}L_{total}}}{2P_{total}}$$

To retrieve ligand-occupied wild type and deamidated protein fractions, relative peak areas of wild type and deamidated peptides were calculated based on the respective LC-MS chromatograms. Ligand-occupied fractions were calculated as described above with the respective $K_d$ for glycan binding to wild type and deamidated P domains. Then, conditional probabilities for all combinations of wild type/deamidated and occupied/unoccupied monomers in P dimers were calculated as exemplarily shown in Equation 11.

Equation 11

$$P_{monomer1}(\text{wild type} \cap \text{occupied}) \cdot P_{monomer2}(\text{wild type} \cap \text{occupied})$$
$$P_{monomer1}(\text{wild type} \cap \text{occupied}) \cdot P_{monomer2}(\text{wild type} \cap \text{unoccupied})$$
$$P_{monomer1}(\text{wild type} \cap \text{occupied}) \cdot P_{monomer2}(\text{deamidated} \cap \text{occupied})$$
$$\dots$$

## 5.4.2   DISSOCIATION CONSTANT ($K_D$)

Using a similar approach as described in chapter 5.4.1 the P dimer association and dissociation constants $K_a$ and $K_d$ can be calculated from monomer-dimer ratios obtained in native MS spectra at different total protein concentrations (Equation 12-Equation 14).

Equation 12
$$P_{total} = \text{Monomer (M) + Dimer (D)}$$

Equation 13

$$2\,M \underset{K_d}{\overset{K_a}{\rightleftharpoons}} D$$
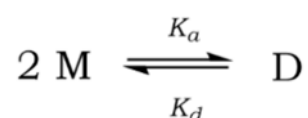
The dissociation constant $K_d$ is then given by

Equation 14
$$K_d = \frac{[M]^2}{[D]}$$

The relative dimer peak area from native MS spectra was calculated and plotted against the total protein concentration ($P_{total}$). The $K_d$ was then determined by global non-linear least squares fitting of Equation 15 (Thiele and Huttner, 1998) to the dataset using OrignPro 2016 software:

Equation 15
$$\frac{[D]}{P_{total}} = 1 - \frac{\sqrt{\frac{K_d}{2}\left(P_{total}+\frac{K_d}{8}\right) - \frac{K_d}{4}}}{P_{total}}$$

### 5.4.3   D/H BACK EXCHANGE (BE)

Based on HDX-MS community recommendations (Masson et al., 2019) D/H back exchange was calculated as follows:

Equation 16      $\text{back exchange (BE)} = \left(1 - \frac{m_{100\%} - m_{0\%}}{N \cdot D_{frac}}\right) \cdot 100$

The back exchange corrected deuterium uptake is then reflected by

Equation 17      $D_{corr} = \frac{m - m_{0\%}}{m_{100\%} - m}$

with

| | | |
|---|---|---|
| $m$ | = | observed peptide centroid mass |
| $m_{0\%}$ | = | undeuterated peptide centroid mass |
| $m_{100\%}$ | = | maximally labeled peptide centroid mass |
| $N$ | = | theoretical number of exchangeable peptide amides |
| $D_{frac}$ | = | fraction of D/H in the labeling buffer (0.9 in this case) |

## 5.5   Computational methods

### 5.5.1   SEQUENCE AND STRUCTURE ALIGNMENTS

GII.4 Saga, GII.4 MI001, GII.10 Vietnam and GII.17 Kawasaki, GV CR10 and GV MNV07 VP1 or P domain protein sequences were aligned with T-Coffee (Tommaso et al., 2011) and visualized with Jalview version 2.11.0 (Waterhouse et al., 2009). Phylogenetic trees were created in Jalview with BLOSUM62 and neighbor joining. GII.4 Saga (pdb 4X06), GII.10 Vietnam (pdb 3ONY) and GII.17 Kawasaki (pdb 5F4O) as well as GII.4 Saga (pdb 4X06) and GV CR10 (pdb 6H6M) P dimer crystal structures were superimposed in PyMOL.

### 5.5.2   HOMOLOGY MODELING

The SWISS-MODEL template library (Waterhouse et al., 2018) (Bertoni et al., 2017) SMTL version 2019-10-24, PDB release 2019-10-18 was searched with BLAST (Camacho et al., 2009) and HHBlits (Remmert et al., 2011) for evolutionary related structures matching the GII.4 MI001 P dimer target sequence. Based on the search results the GII.4 Farmington Hills P dimer structure (pdb 4OOV, 94 % sequence identity) was used for model building. Models were built based on the target-template alignment using ProMod3. The global and per-residue model quality has been assessed using the QMEAN scoring function (Benkert et al., 2011). The resulting GMQE (Global Model Quality Estimation) was 0.99 and QMEAN was 0.57 indicating very good accuracy and quality of the model structure (Figure S10).

## 5.6   Data visualization

Glycan, lipid and detergent structures as well as chemical reaction schemes were drawn with ChemDraw version 19.1 (PerkinElmer).

PyMOL™ Molecular Graphics System, Version 2.4.0a0 (Schrödinger) was used to view and modify protein structures. Structure files were downloaded from the Protein Data Bank (PDB) at rcsb.org (Berman et al., 2000).

Figures were prepared with Adobe Illustrator 2019 (Adobe Inc.).

## 5.7   Data availability

Full HDX data tables as well as MS raw data and peptide identification results of published datasets have been deposited to the ProteomeXchange Consortium (Deutsch et al., 2017) via the PRIDE (Vizcaino et al., 2016) partner repository.

**Chapter 3.1**
Dataset identifier: PXD011914
Reference: Mallagaray et al. (2019)

**Chapter 3.2**
Dataset identifier: PXD019884
Reference: Dülfer et al. (in preparation)

## 6 SUPPLEMENT

## 6.1 Supplemental figures

### 6.1.1 NOROVIRUS
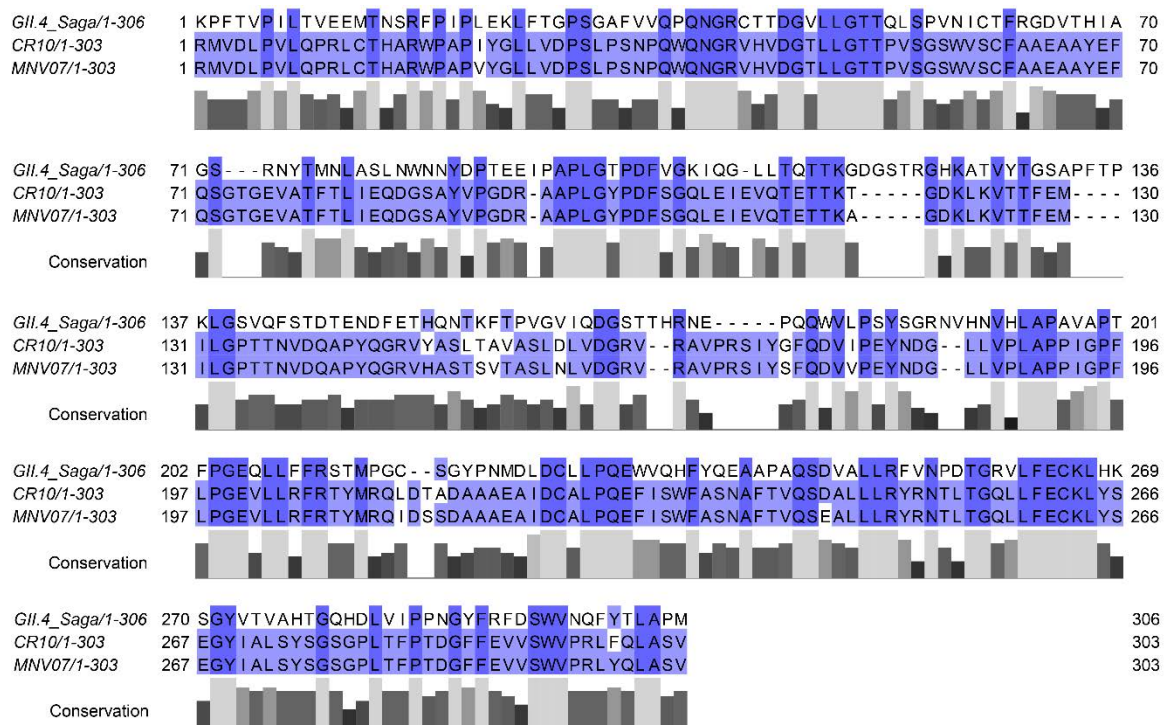


Figure S1: Sequence alignment of GII.4 Saga, GV CR10 and GV MNV07 P domain sequences. Color-coding represents sequence identity (blue), similarity (light blue) or difference (white). Bars below the sequence indicate the degree of amino acid conservation among the strains at the respective position (low (black) to high (light grey)).
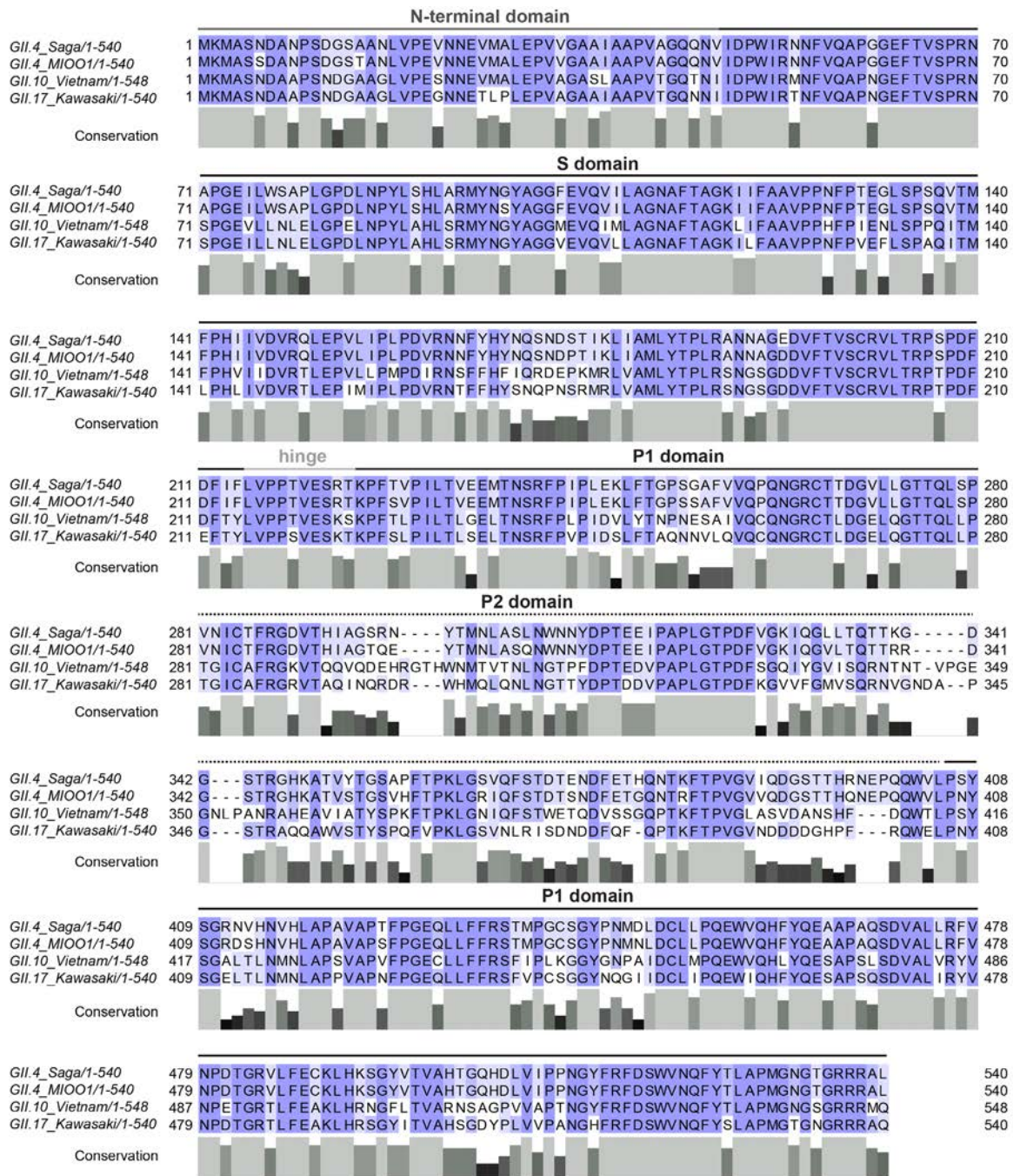
Figure S2: Sequence alignment of GII.4 Saga, GII.4 MI001, GII.10 Vietnam and GII.17 Kawasaki VP1 sequences. Color coding represents sequence identity (blue), similarity (light blue) or difference (white). Bars below the sequence indicate the degree of amino acid conservation among the strains at the respective position (low (black) to high (light grey)).
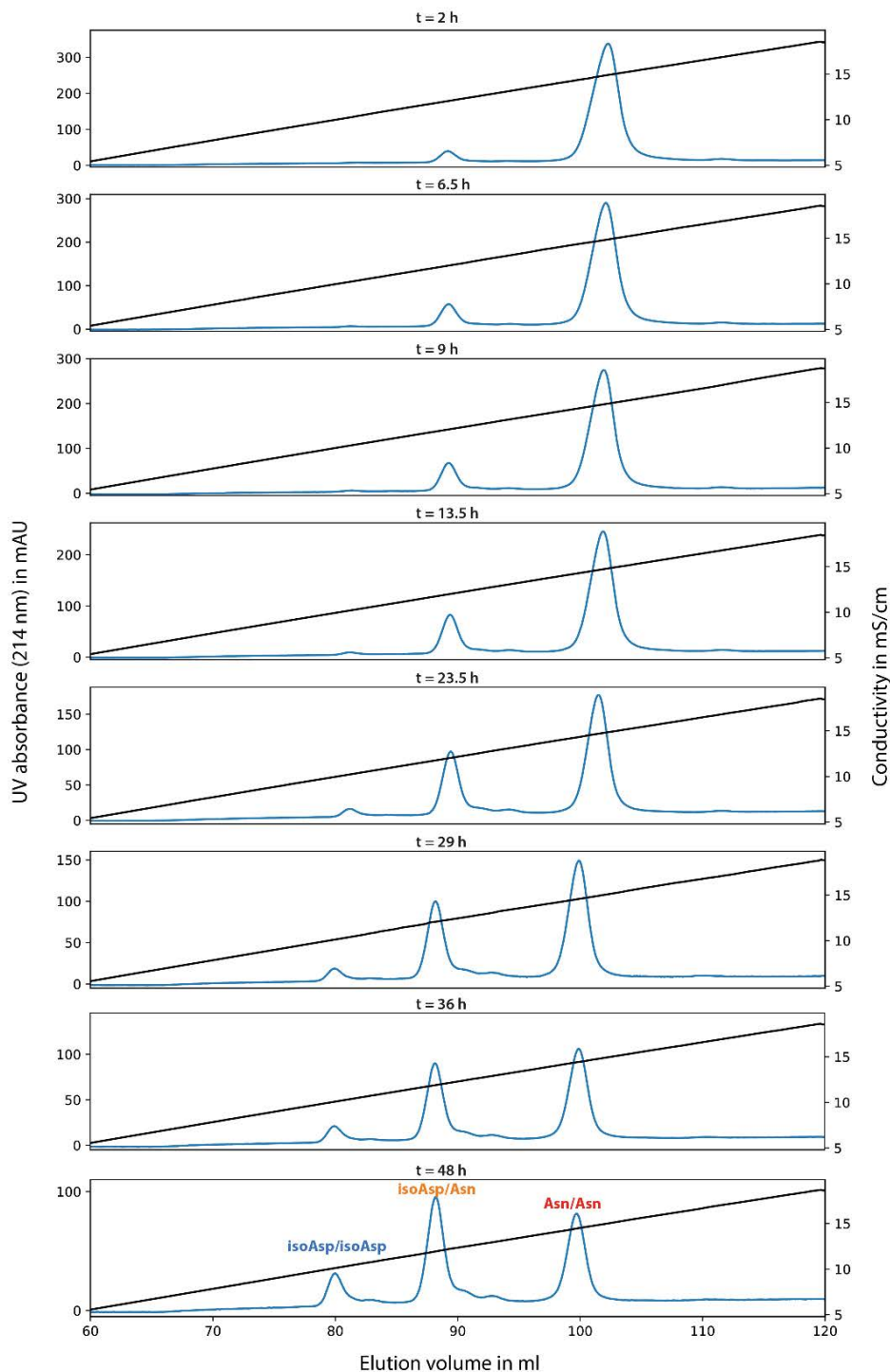
Figure S3: Ion exchange chromatograms of eight P dimer aliquots that were incubated up to 48 h at 37°C. UV absorbance has been measured at 214 nm (blue). Protein species elute with increasing conductivity (black). Three major fractions were identified, corresponding to iDiD, iDN and NN species, respectively. Note that the overall absorbance decreases due to some protein precipitation at prolonged incubation times. Precipitates were removed by centrifugation prior to analysis. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.
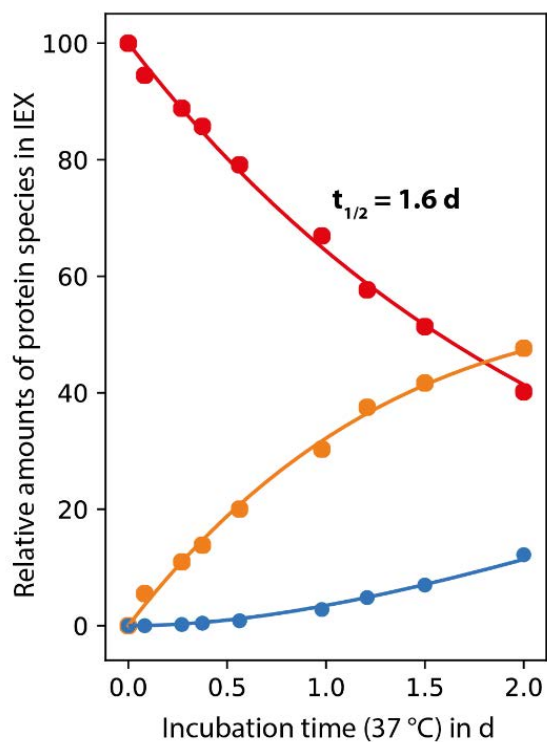
Figure S 4: Deamidation rate of P dimers incubated in 75 mM sodium phosphate buffer, 100 mM NaCl, pH 7.3 for up to 48 h at 37°C. Relative amounts of the different protein species were estimated from their respective peak integrals in ion exchange chromatograms (Figure S3): fully (iDiD, blue), half-deamidated (iDN, orange), and non-deamidated (NN, red) protein species. An exponential decay model was fitted to the decrease of non-deamidated protein yielding a half-life of 1.6 d. Note that this analysis could be biased due to protein precipitation during incubation. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.
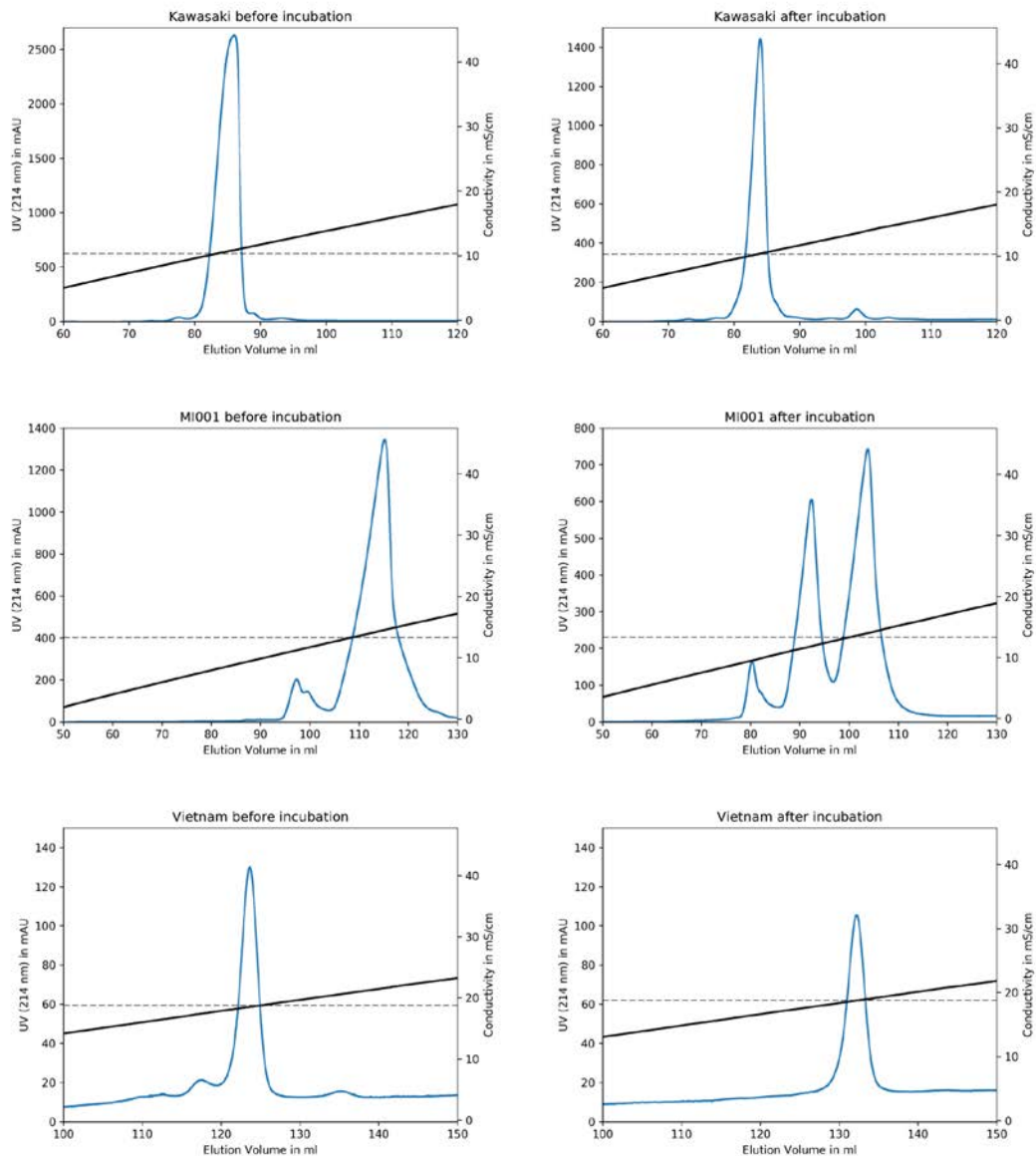
Figure S 5: Ion exchange chromatograms of GII.17 Kawasaki 308, GII.10 Vietnam O26 and GII.4 MI001 P dimers before and after incubation at elevated temperatures. UV absorbance has been measured at 214 nm (blue). The dashed line indicates the conductivity where the major peak elutes. Buffer batches changed over the course of the incubations and, therefore, elution volumes cannot be compared directly across the samples. The major peaks were separated and incubated at elevated temperatures to promote a potential deamidation reaction. Incubations conditions were as follows: 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, 500 mM NaCl (pH 7.3). Protein samples taken after the incubation were tested for deamidation by MS (Figure S 6 and Table S 1). In the case of GII.4 MI001, protein samples from both the first (iDiD) and the last peak (NN) were tested. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.
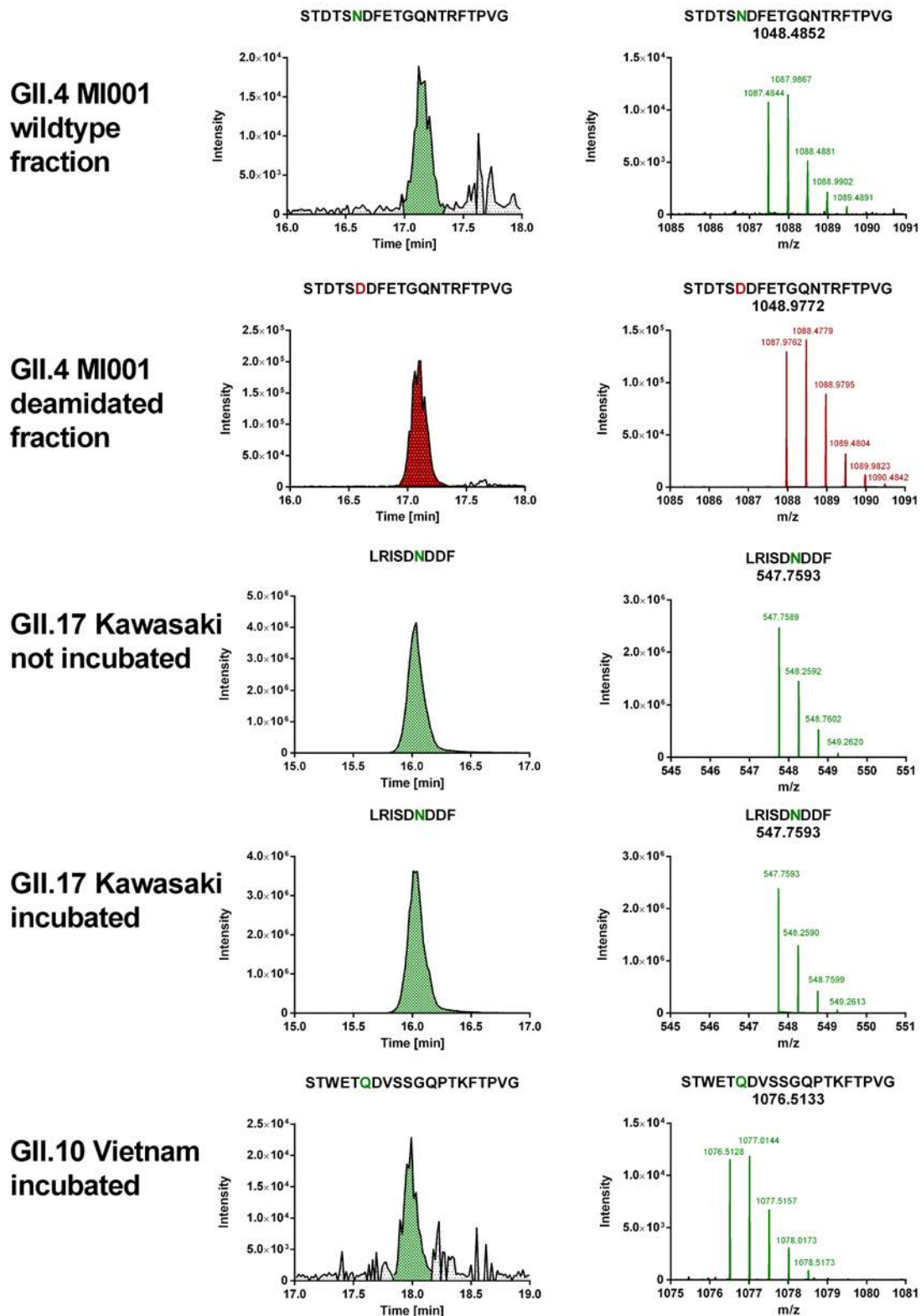
Figure S 6: Chromatograms and mass spectra of peptides ($z$=2) covering the potentially deamidated loops in GII.4 MI001, GII.17 Kawasaki and GII.10 Vietnam P dimers (Figure S 5). The theoretical monoisotopic mass is indicated below the peptide sequence in the mass spectrum. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.
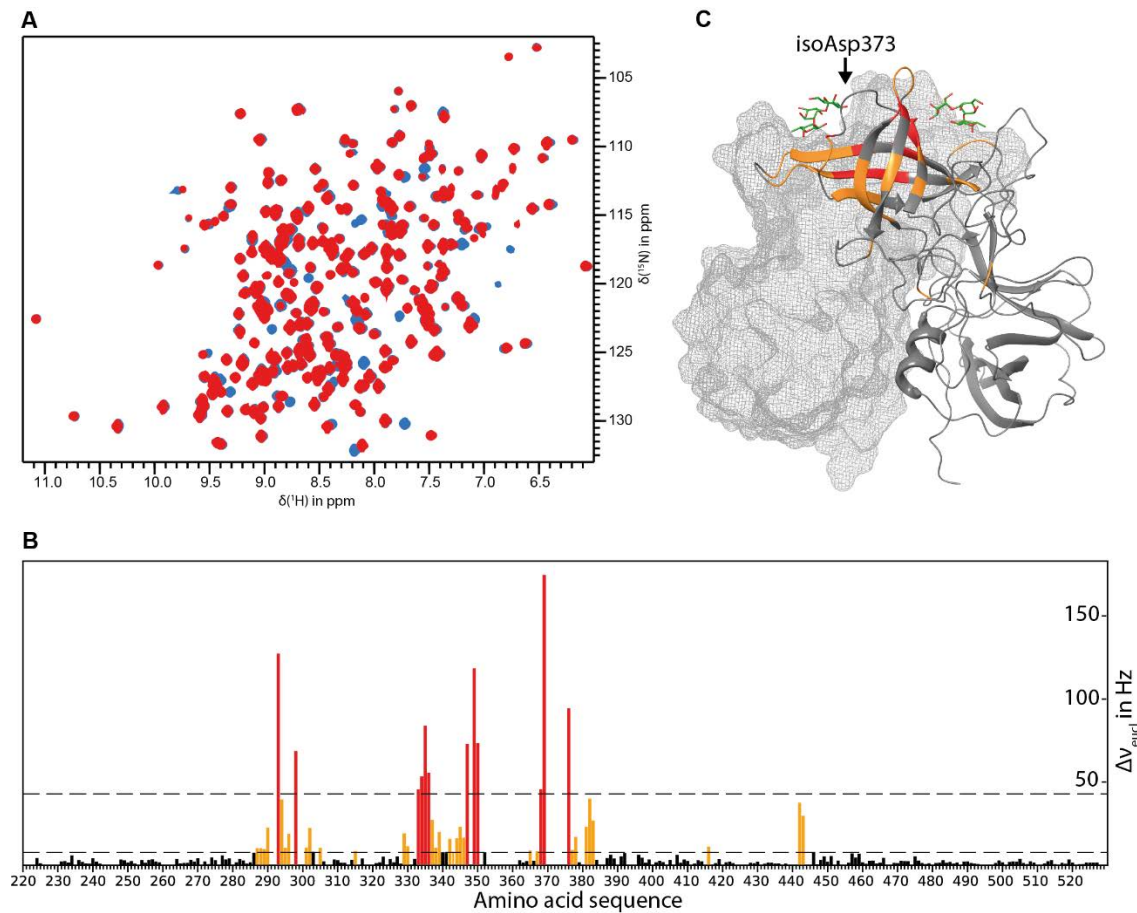
Figure S 7: NMR spectra reveal deamidation of N373 and formation of isoD373. (A) $^1$H,$^{15}$N TROSY HSQC of native Saga P dimers (red) and deamidated, post-translationally modified Saga P dimers (blue). (B) Chemical shift perturbations (CSPs) induced by irreversible deamidation of N373 with dashed lines at 2σ (red color) and at the experimentally determined significance threshold of 7.9 Hz (orange color). (C) CSPs mapped onto the crystal structure of P dimers complexed with B trisaccharide (pdb 4X06). Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.
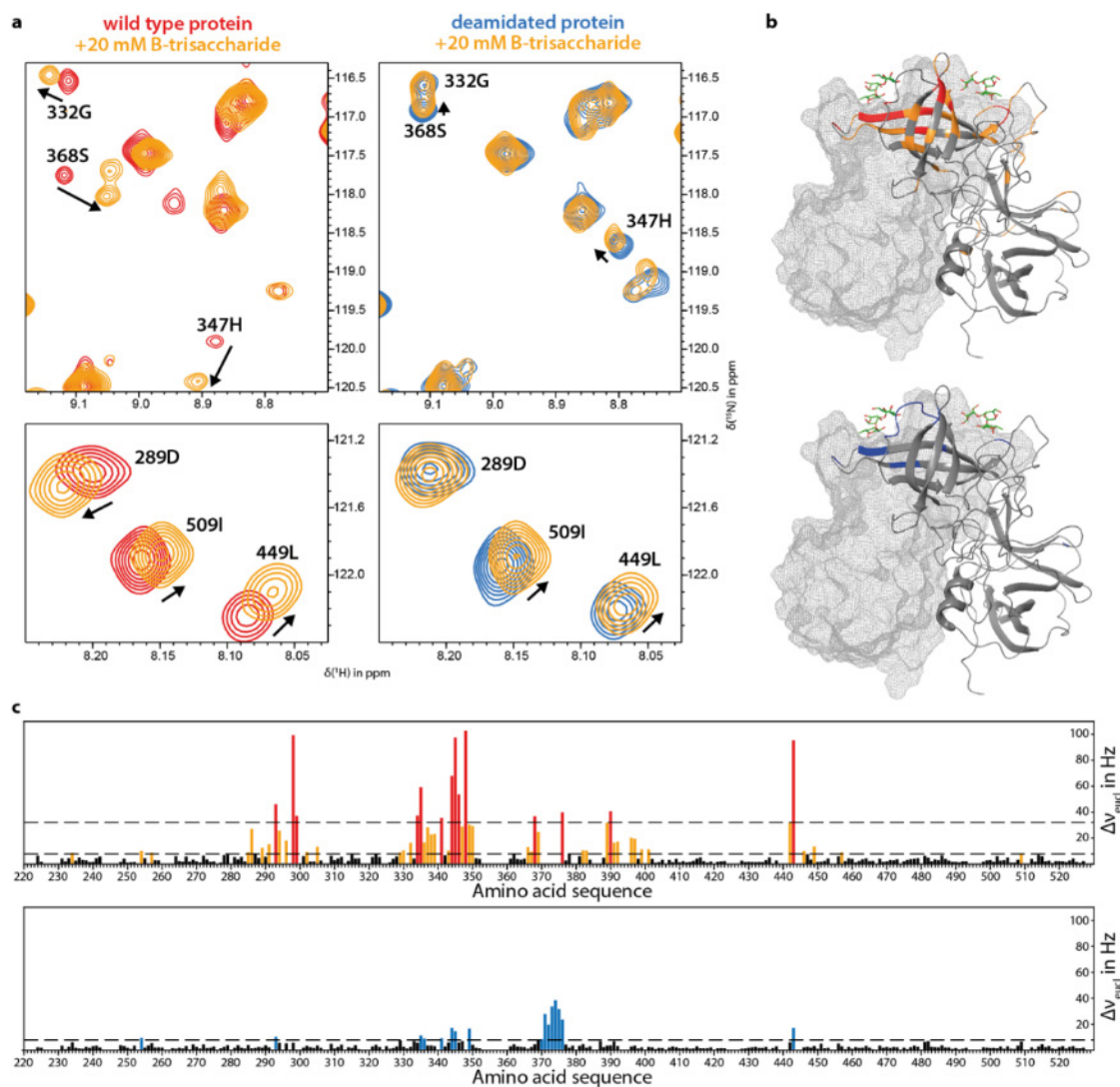
Figure S 8: Binding of blood group B trisaccharide to GII.4 Saga P dimers. a) $^1$H,$^{15}$N, TROSY HSQC spectra of native (red, left) and deamidated (blue, right) Saga P dimers in their free forms and in presence of 20 mM B trisaccharide (orange). Perturbations induced by ligand binding are indicated by arrows for selected amino acids. b) Mapping of 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 2σ threshold to indicate largest CSPs in native P dimers. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.
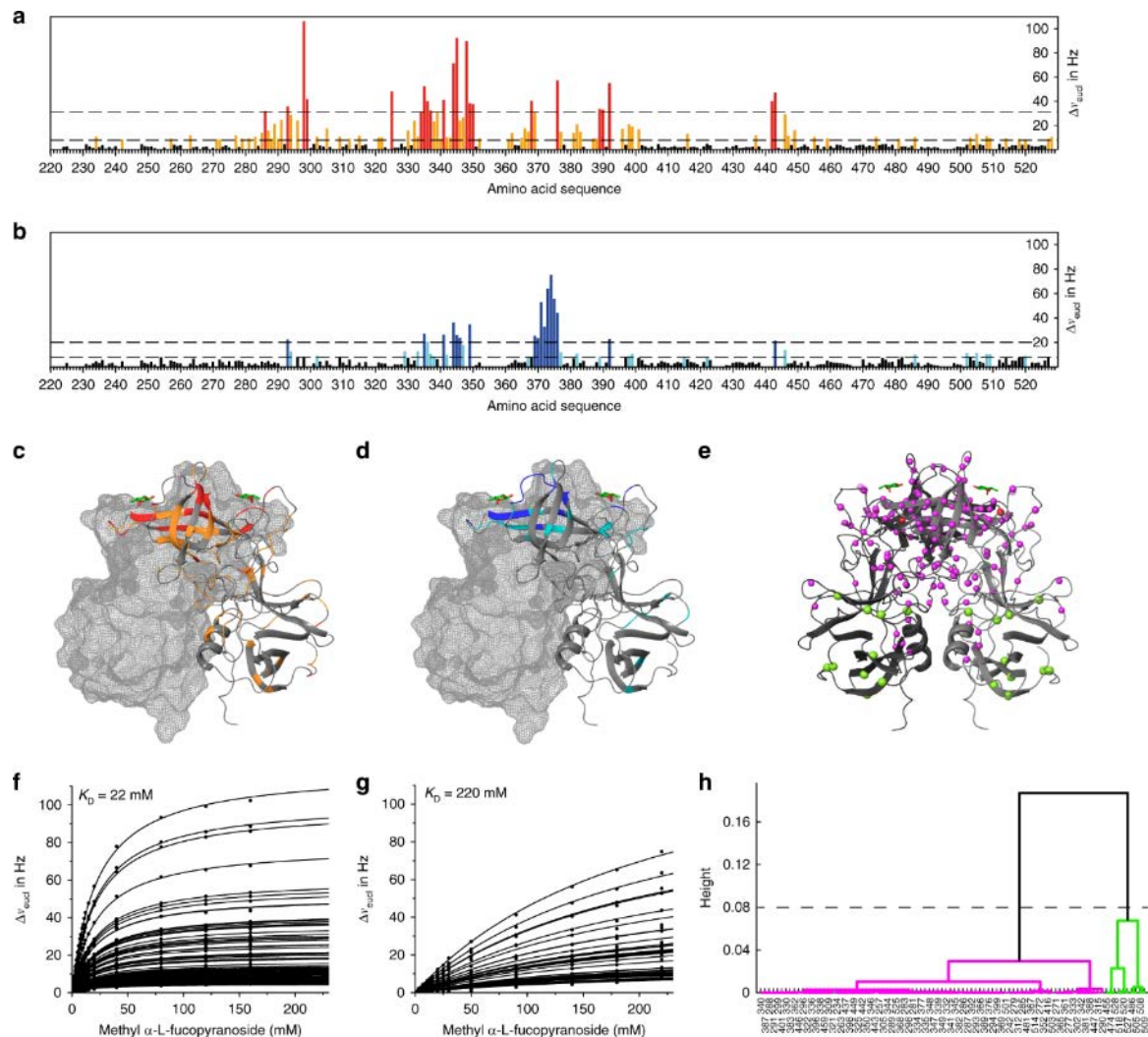
Figure S 9: Binding of methyl α-L-fucopyranoside to GII.4 Saga P dimers. a, b) Chemical shift perturbations (CSPs) of NH backbone signals of GII.4 Saga NN (a) and iDiD P dimers (b) in the presence of 160 and 220 mM methyl α-L-fucopyranoside, respectively. Dashed lines are at 2σ (red, blue) and at the 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. 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). 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). Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.
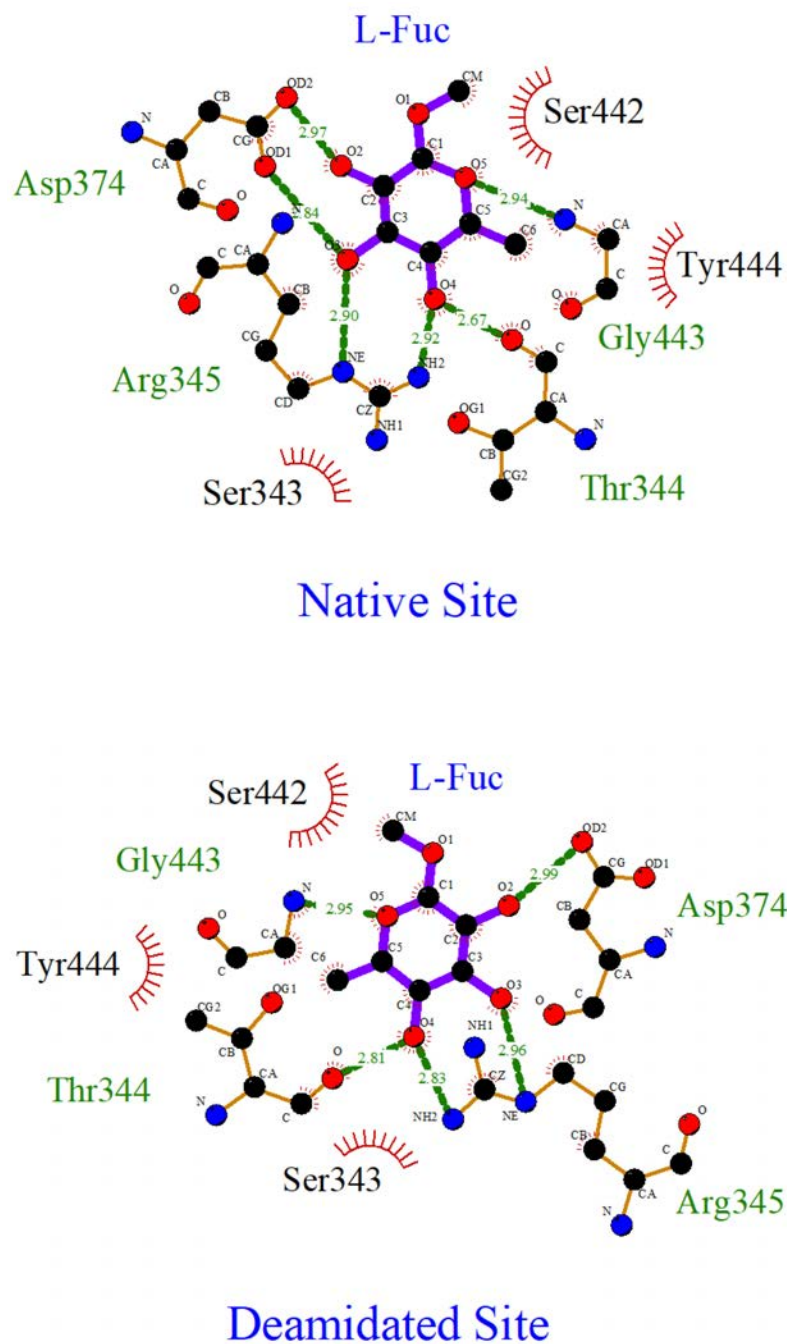
Figure S 10: Ligand interaction diagrams for methyl α-L-fucopyranoside bound to GII.4 Saga P dimers (pdb 6h9v) in the native and in the deamidated binding site. In the native site a bivalent hydrogen bond is formed between the side chain carboxyl group of Asp374 and the hydroxyl groups at positions C2 and C3 of methyl α-L-fucopyranoside. In the deamidated site, the Asp374 side chain carboxyl group makes only one hydrogen bond with the C2 hydroxyl group of methyl α-L-fucopyranoside. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.
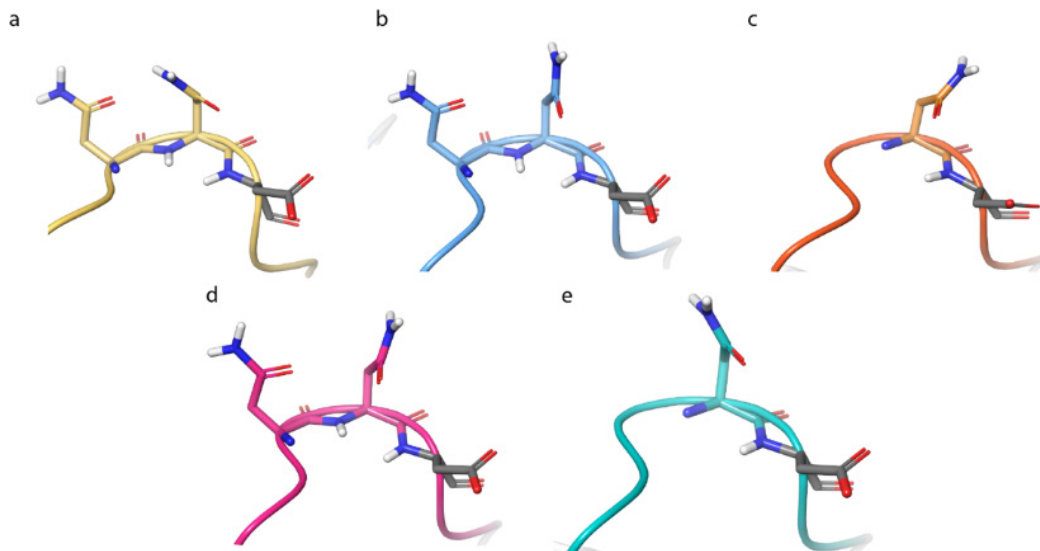
Figure S 11: Representative crystal structures from GII.4 norovirus strains containing Asn residues at position 372 and/or 373 in a reactive conformation. (a) VA387 (yellow, pdb 2obt), (b) CHDC2094 1974 (blue, pdb 5iyq), (c) NL 2004 (orange, pdb 3sld), (d) Farmington Hills 2004 (pink, pdb 4oov), and (e) Saga 2006 (turquoise, pdb 4wzl). Asp 374 is always depicted in grey. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.
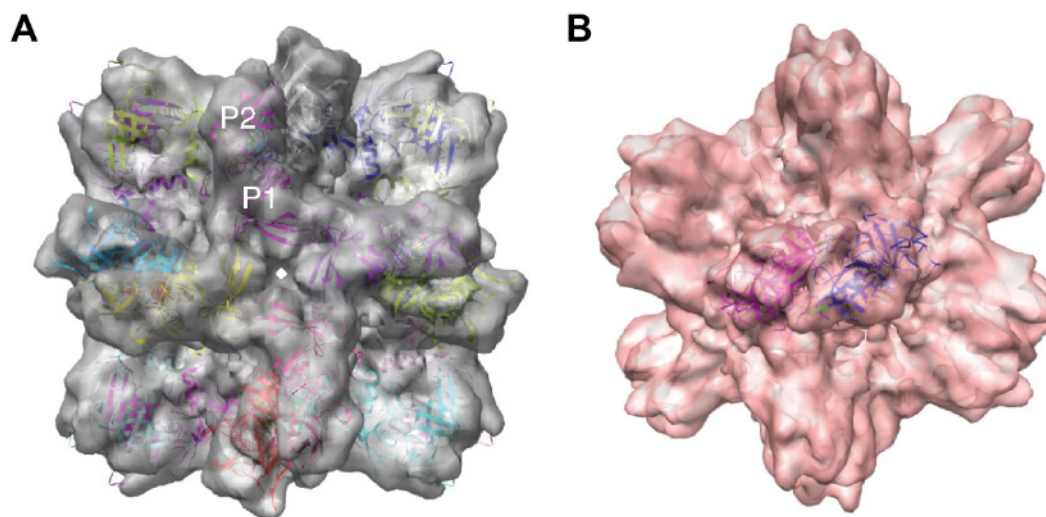


Figure S 12: Fitting of P dimer structures into P particles . (A) Structure and orientation of VA387 P dimers within the cryo-EM density map of P particles (transparent gray, four-fold view). (B) Location and structure of a single P dimer (monomers colored in blue and pink) in the P particle (two-fold top view). Adapted from Tan et al. (2008), Copyright (2008), with permission from Elsevier.
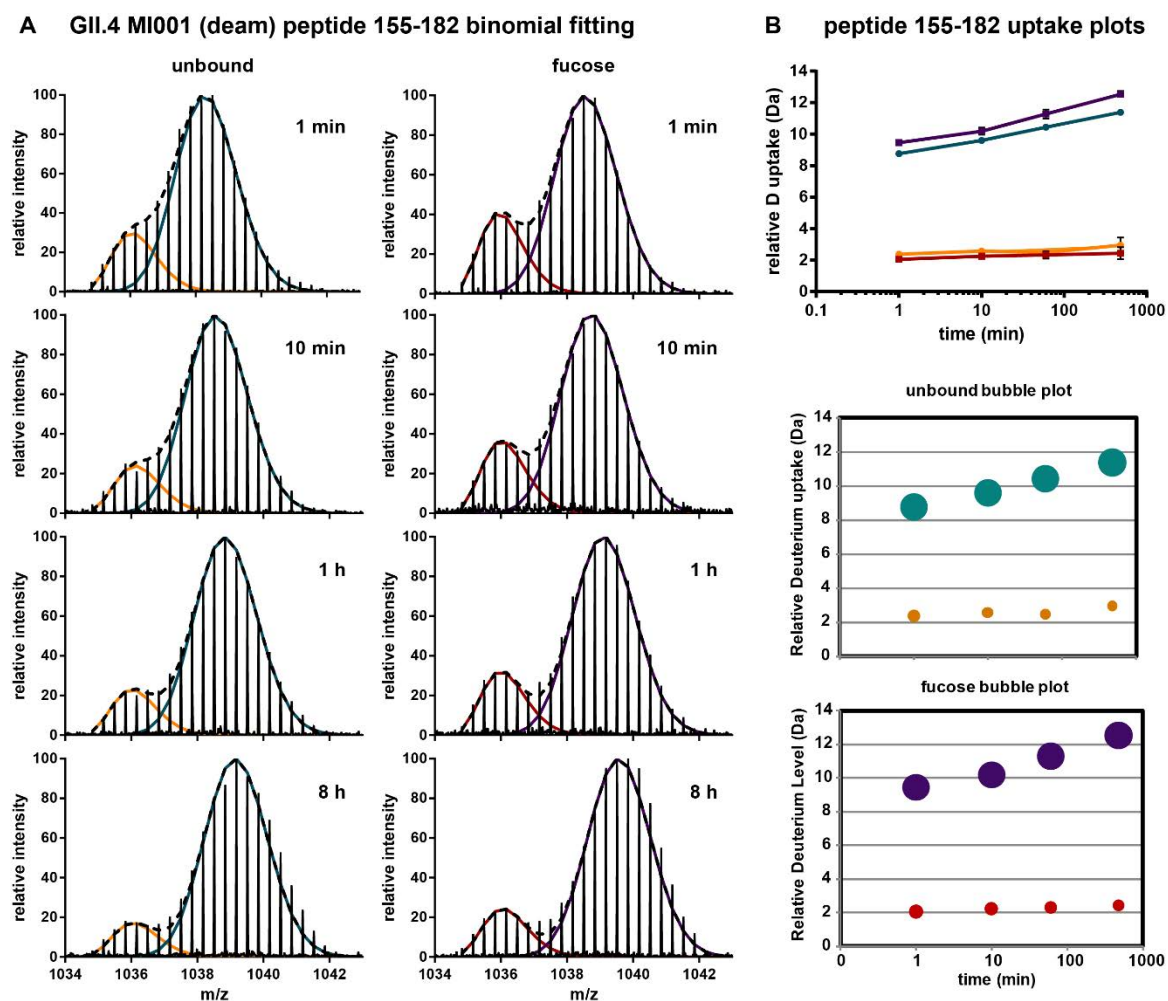
Figure S 13: Peptide HDX data analysis example for binomial fitting of bimodal spectra. (A) Binomial fitting with bimodal deconvolution was applied in HXExpress resulting in a low (orange, red) and a high (green, violet) deuterated peak distribution for every state and time point. (B) Deuteration values for the individual peak distributions can be visualized as uptake plot and significant differences between the unbound and ligand-bound state can be statistically analyzed. The relative deuterium uptake for each peak distribution can also be displayed as bubble plot, where the circle area corresponds to the relative intensity of the individual peak distribution at a given time point. Here, constant intensity ratios over time point towards the presence of two conformationally distinct protein populations.

**A**



**B**
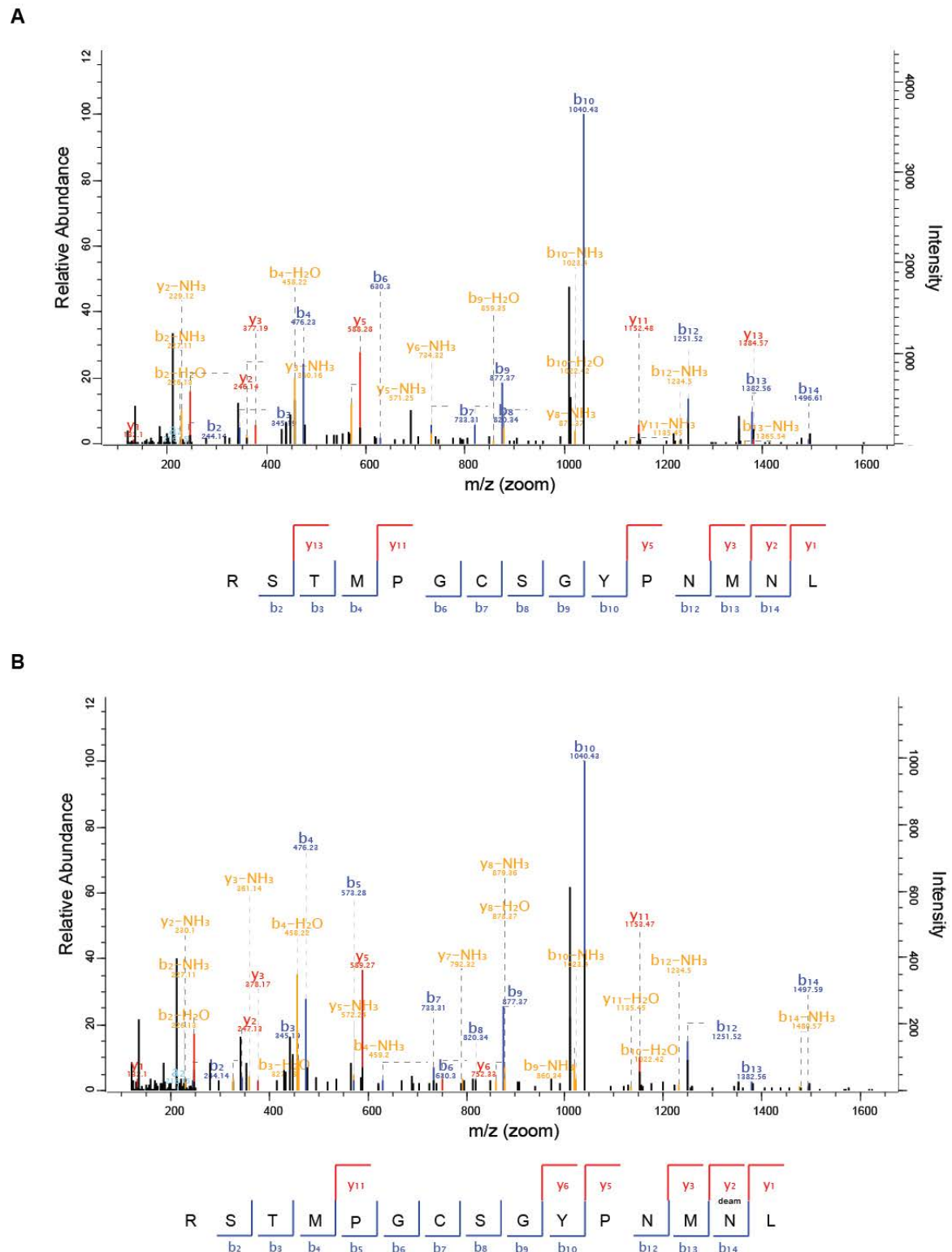


Figure S 14: Fragment spectra of wild type (A) and deamidated (B) peptide RSTMPGCSGYPNMN[448]L from GII.4 MI001 P dimer stored at pH 4.9.
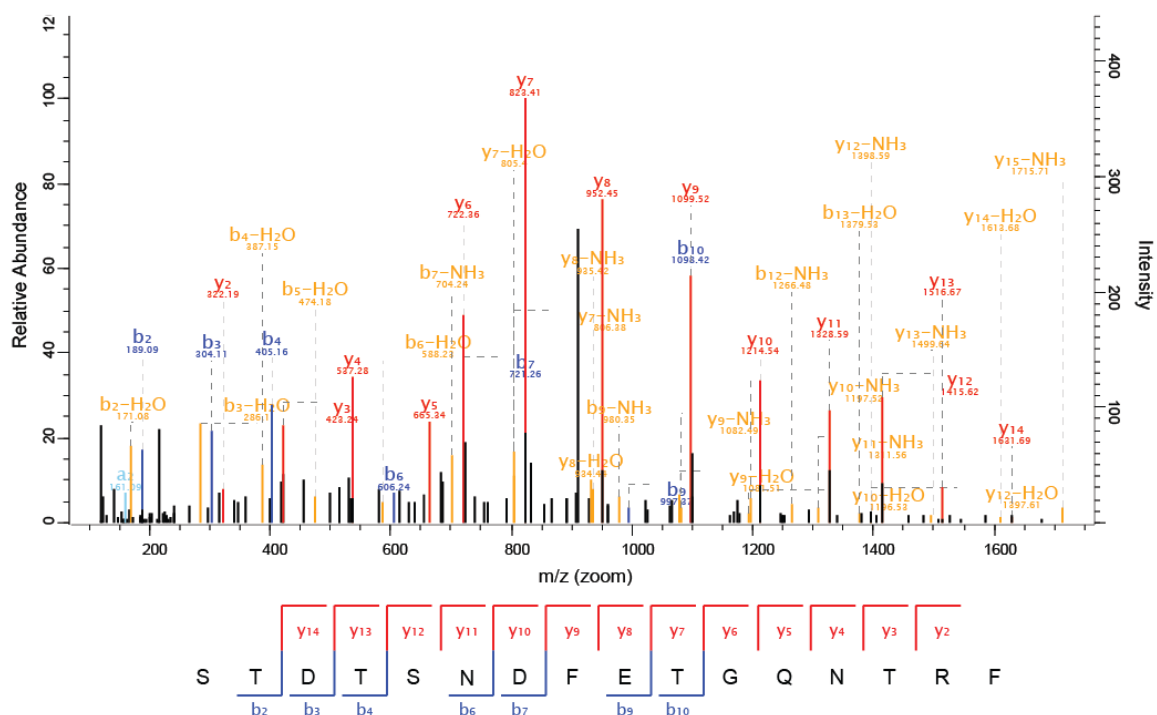
Figure S 15: Fragment spectra of wild type peptide STDTSN[373]DFETGQNTRF from GII.4 MI001 P dimer stored at pH 4.9. No peptide carrying a deamidation at N373 could be detected.
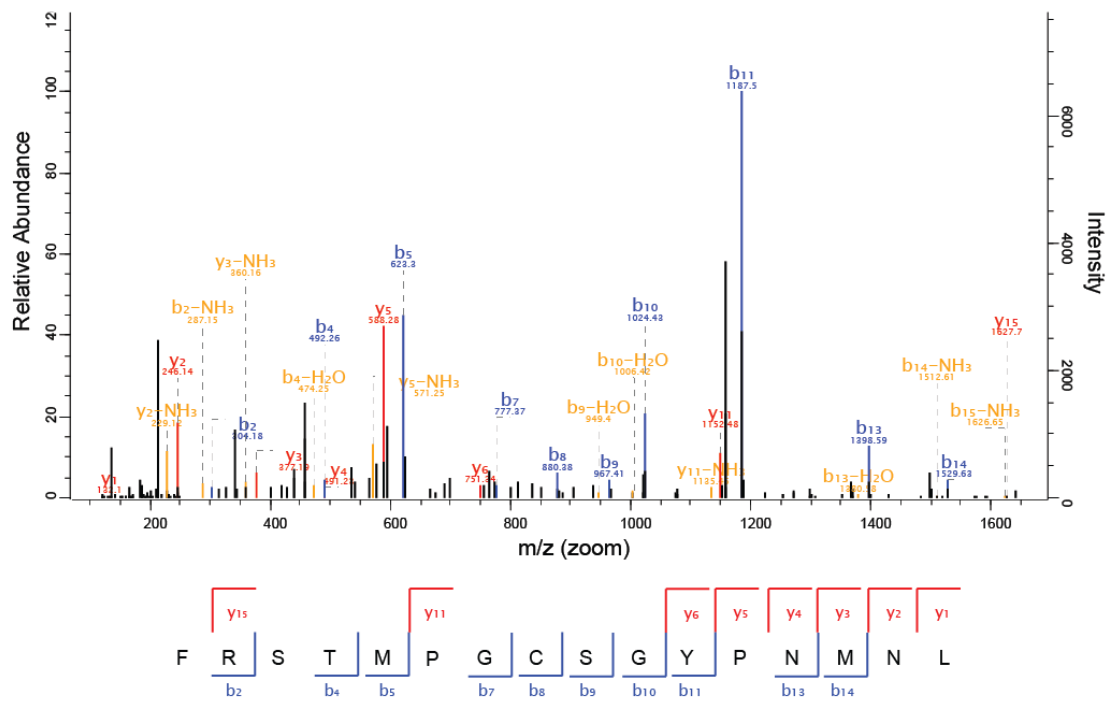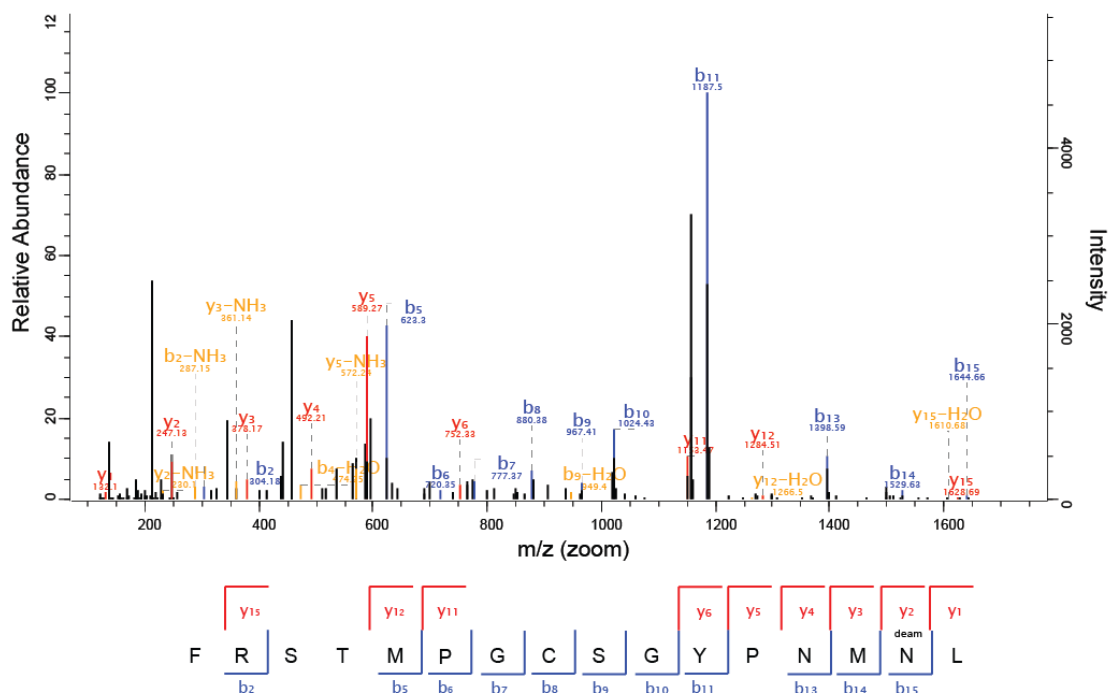
Figure S 16: Fragment spectra of wild type (A) and deamidated (B) peptide FRSTMPGCSGYPNMN[448]L from GII.4 MI001 P dimer stored at pH 7.3.

**A**
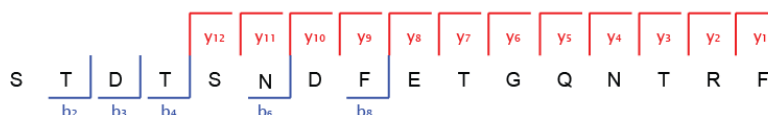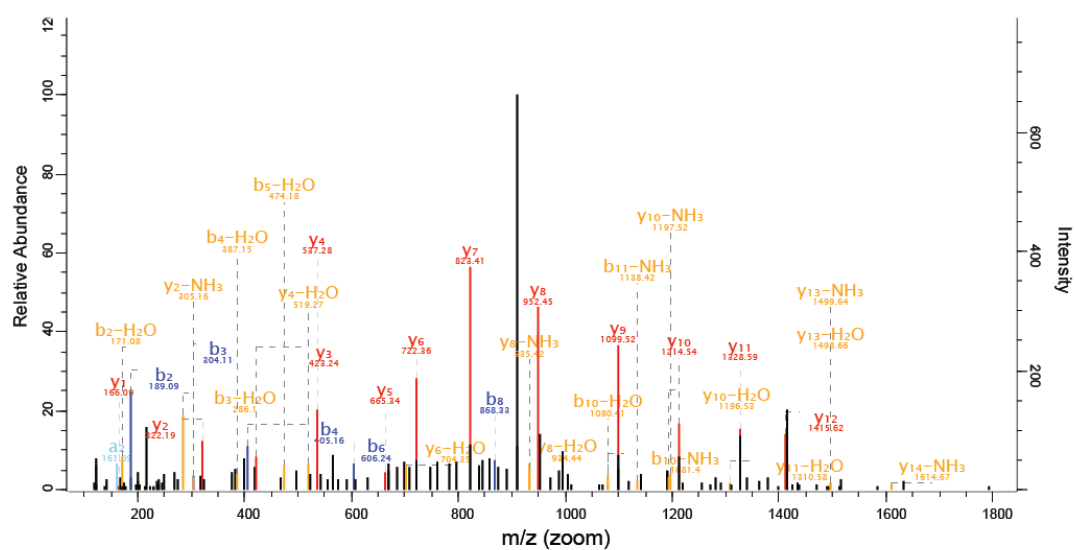


**B**
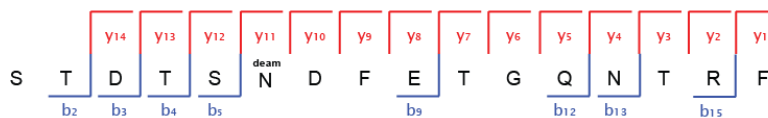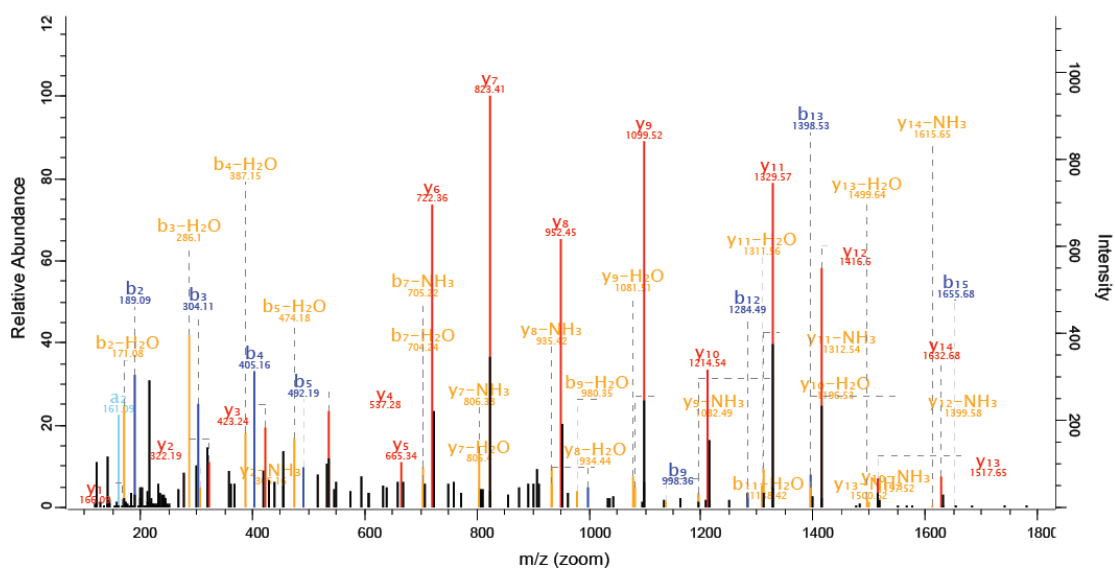


Figure S 17: Fragment spectra of wild type (A) and deamidated (B) peptide STDTSN[373]DFETGQNTRF from GII.4 MI001 P dimer stored at pH 7.3.
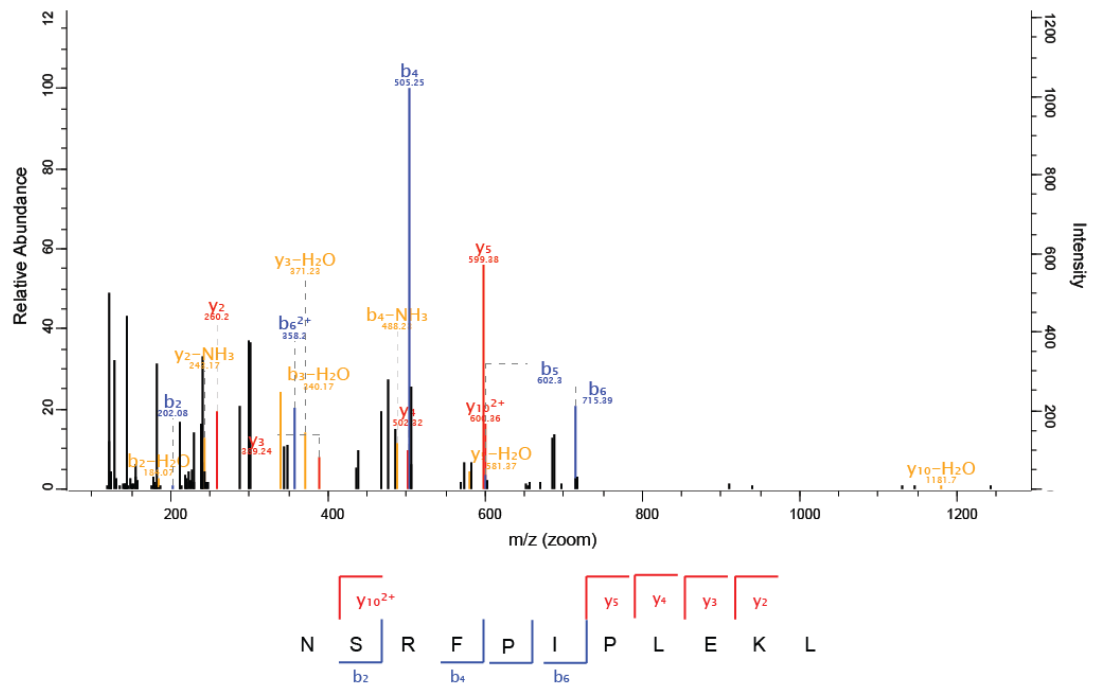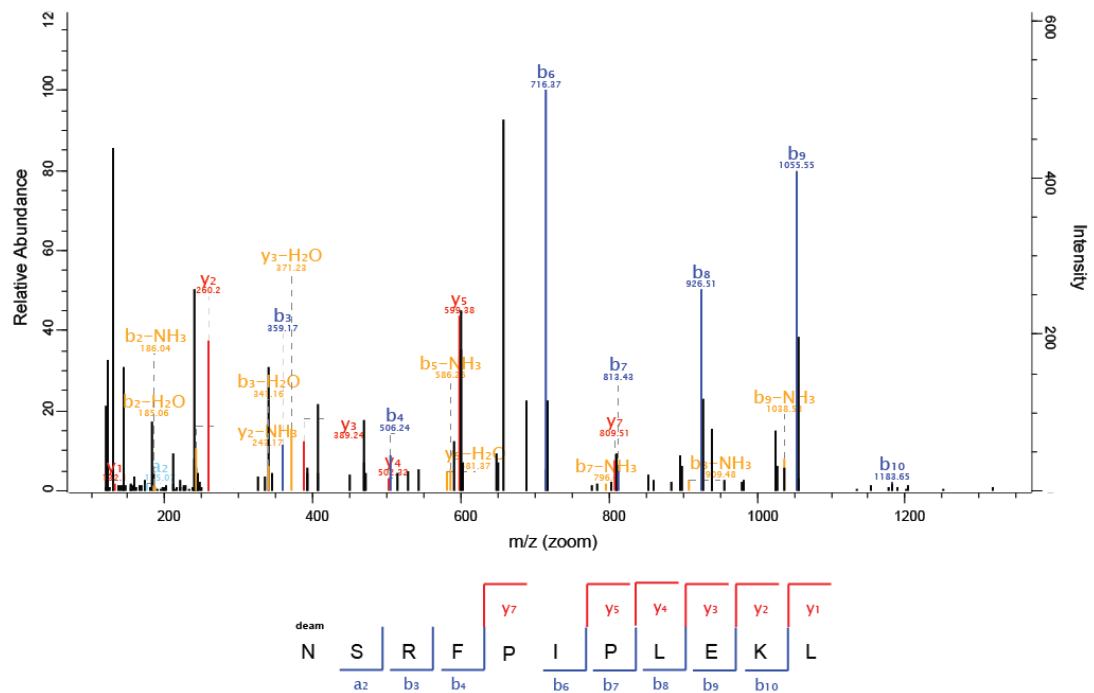
Figure S 18: Fragment spectra of wild type (A) and deamidated (B) peptide N[239]SRFPIPLEKL from GII.4 MI001 P dimer stored at pH 7.3.
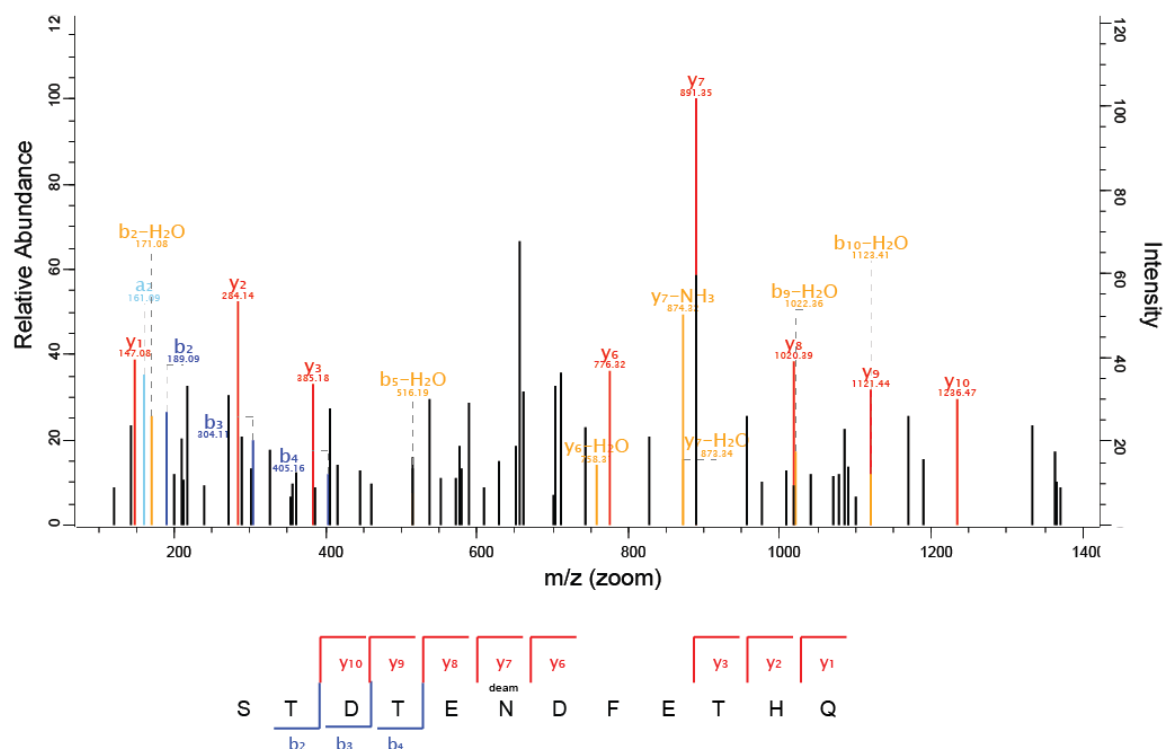
Figure S 19: Fragment spectra of deamidated peptide STDTSN[373]DFETHQ from GII.4 Saga P dimer stored at pH 7.3. Due to the low abundance wild type N373 peptides did not get selected for fragmentation in the mass spectrometer and could therefore not be identified in the automated peptide search in MaxQuant. Therefore, MS data was searched manually for wild type peptide precursors, which led to the identification of approximately 12 % N373 wild type peptides.

**A**



**B**
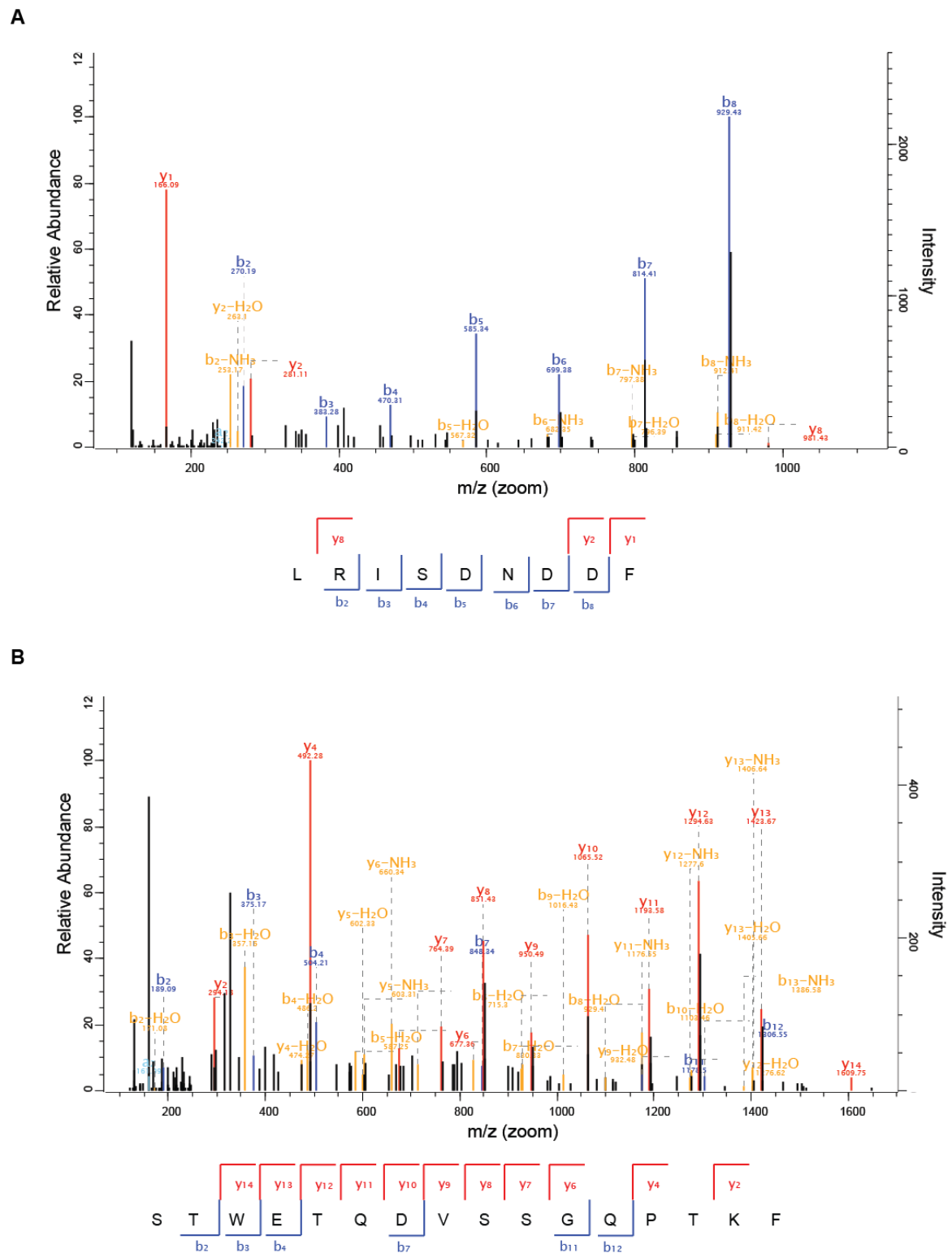


Figure S 20: Fragment spectra of peptides LRISDND[377]DF from GII.17 Kawasaki P dimer (A) and STWETQ[384]DVSSGQPTKF from GII.10 Vietnam P dimer (B). Both strains retain their wild type sequence at the N373 equivalent positon, even after months to years of storage.
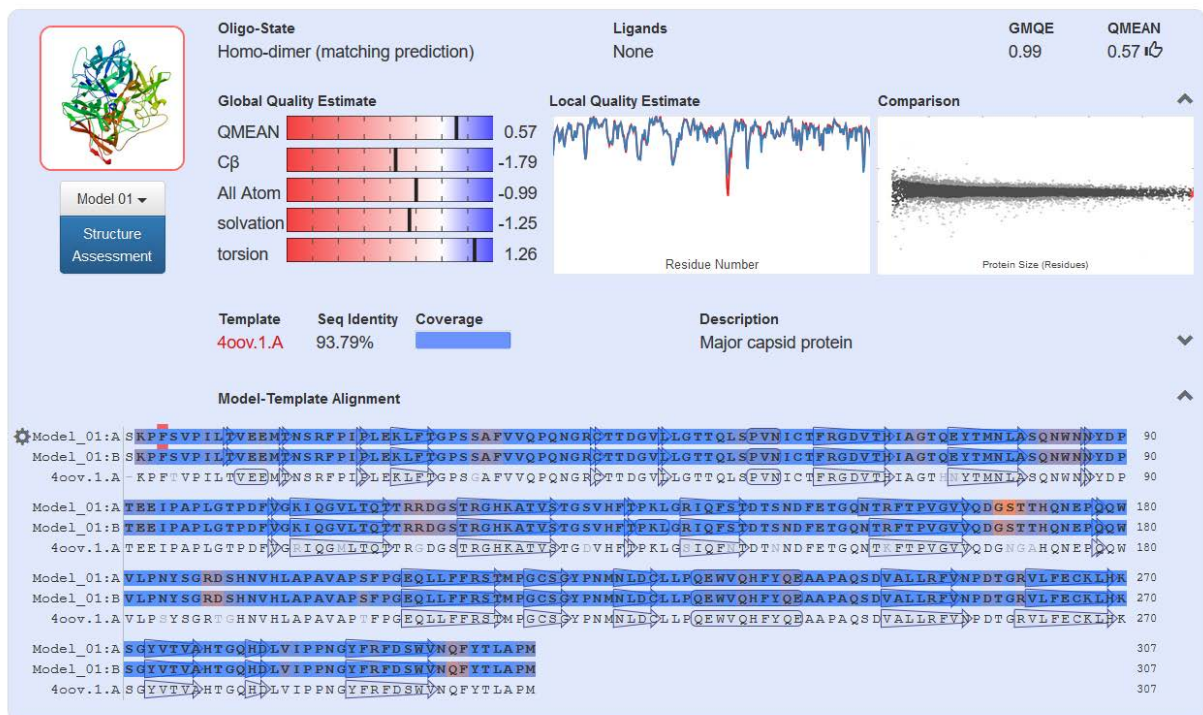
Figure S 21: SWISS MODEL result for the GII.4 MI001 P dimer homology model indicating good model accuracy.

## 6.1.2   MISTIC
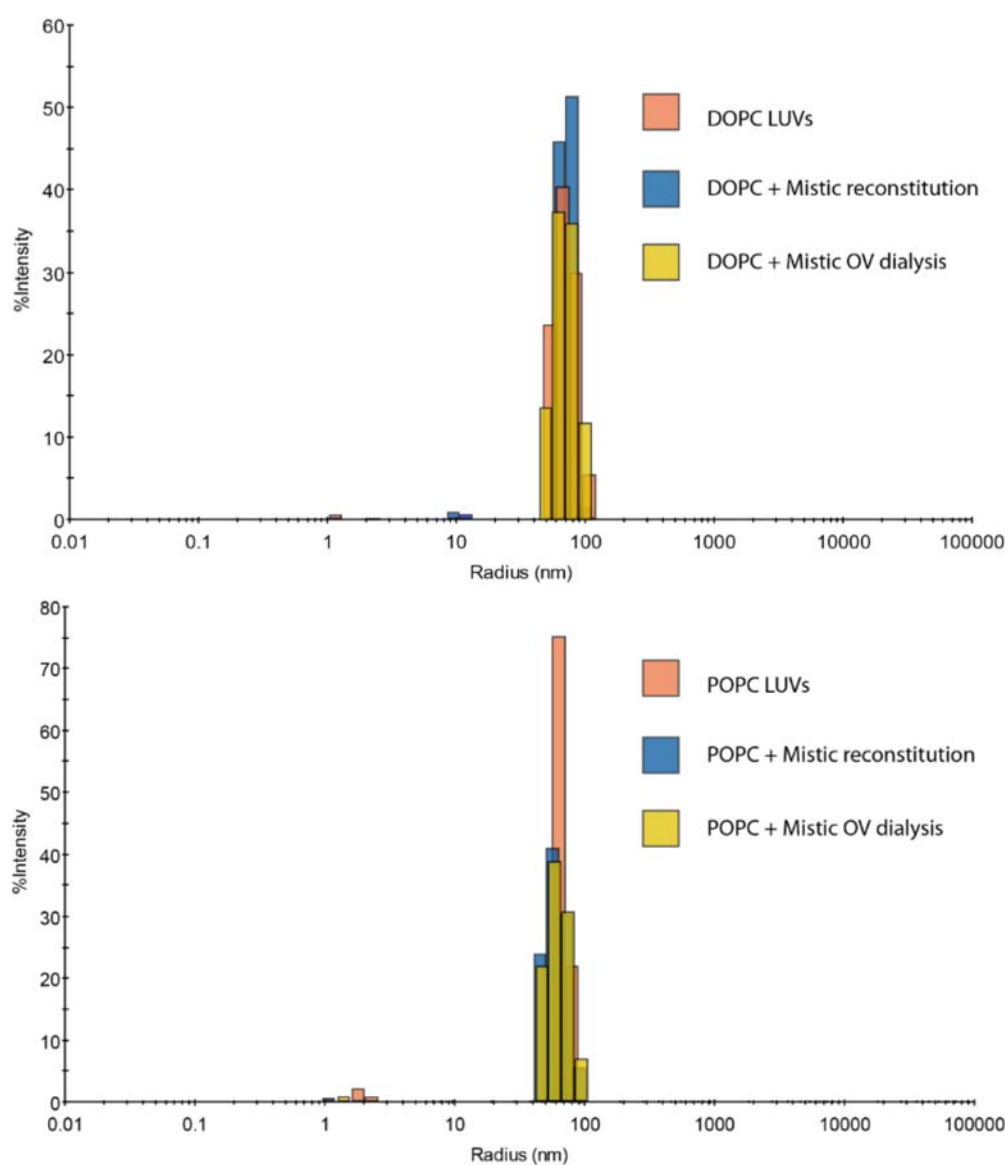


Figure S 22: DLS measurements of DOPC and POPC LUVs before (orange) and after Mistic reconstitution (blue) as well as after overnight dialysis (yellow) show uniform size distributions and no aggregation.
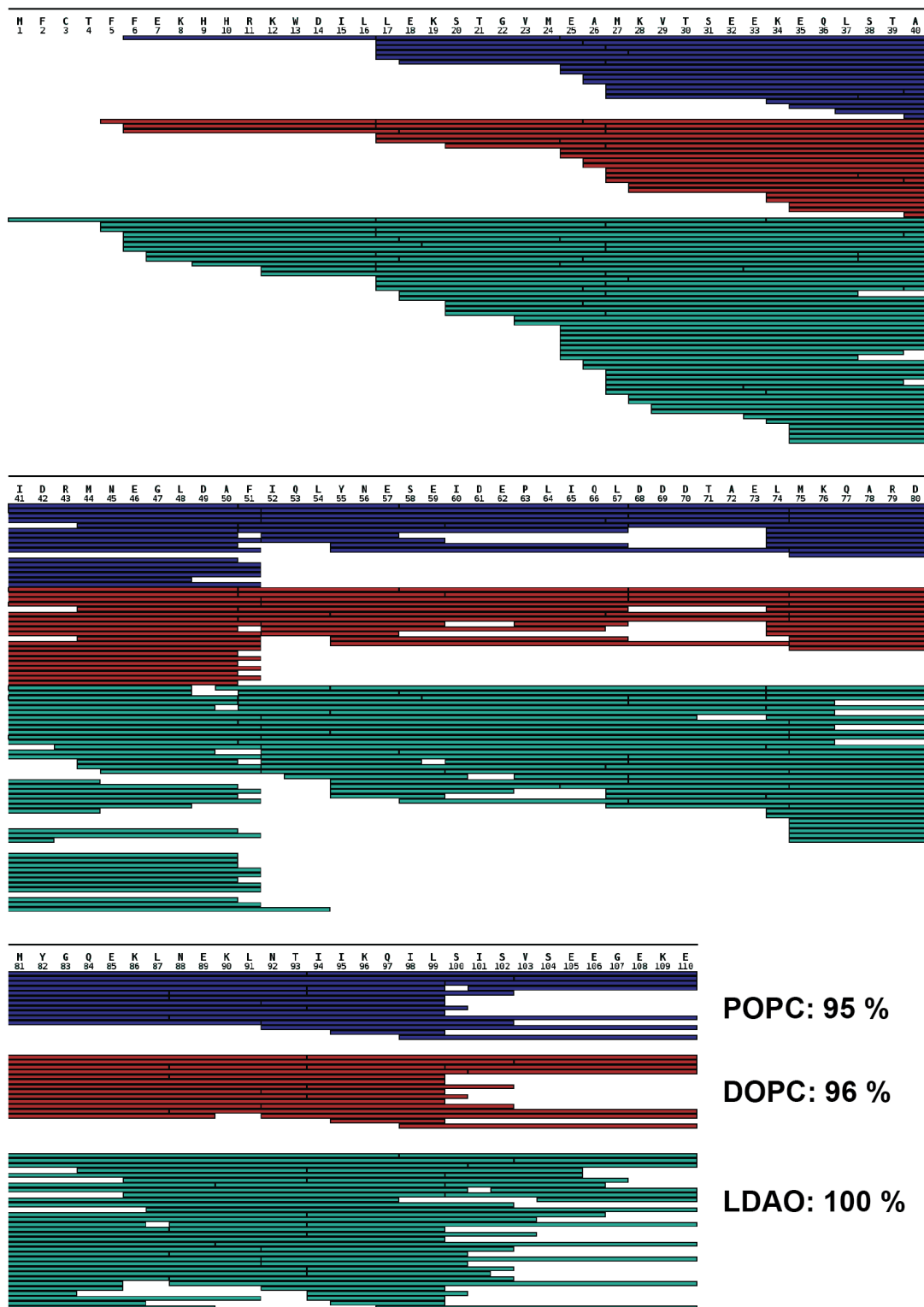
Figure S 23: Coverage maps for Mistic in POPC and DOPC vesicles or LDAO detergent quenched with Q3, lipid depletion and digested on column with pepsin (see chapter 3.4.1). Sequence coverage is almost identical for all three states, even though the number of peptides in higher in LDAO.
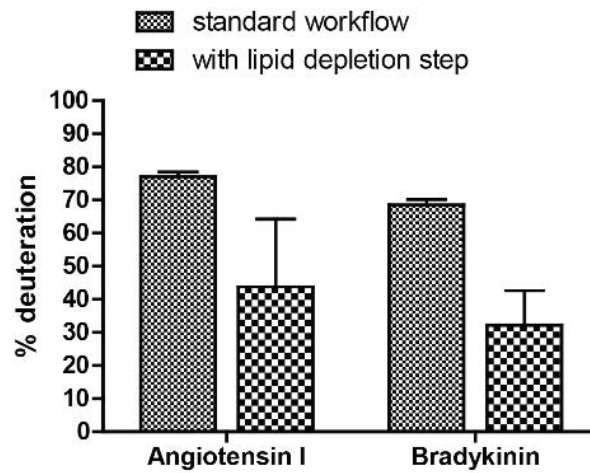
Figure S 24: Assessment of back exchange from fully deuterated Angiotensin I and Bradykinin peptides in the standard HDX-MS workflow and with additional lipid depletion. Error bars indicate the standard deviation (N=2). Introduction of a lipid depletion step strongly increases the back exchange and the experimental error.

Average model peptide back exchange in both HDX-MS setups:

| | | |
|---|---|---|
| Standard workflow: | Angiotensin I: 23 ± 1 % | Bradykinin: 32 ± 2 % |
| Lipid depletion: | Angiotensin I: 59 ± 25 % | Bradykinin: 66 ± 11 % |

# mistic peptide 74-93
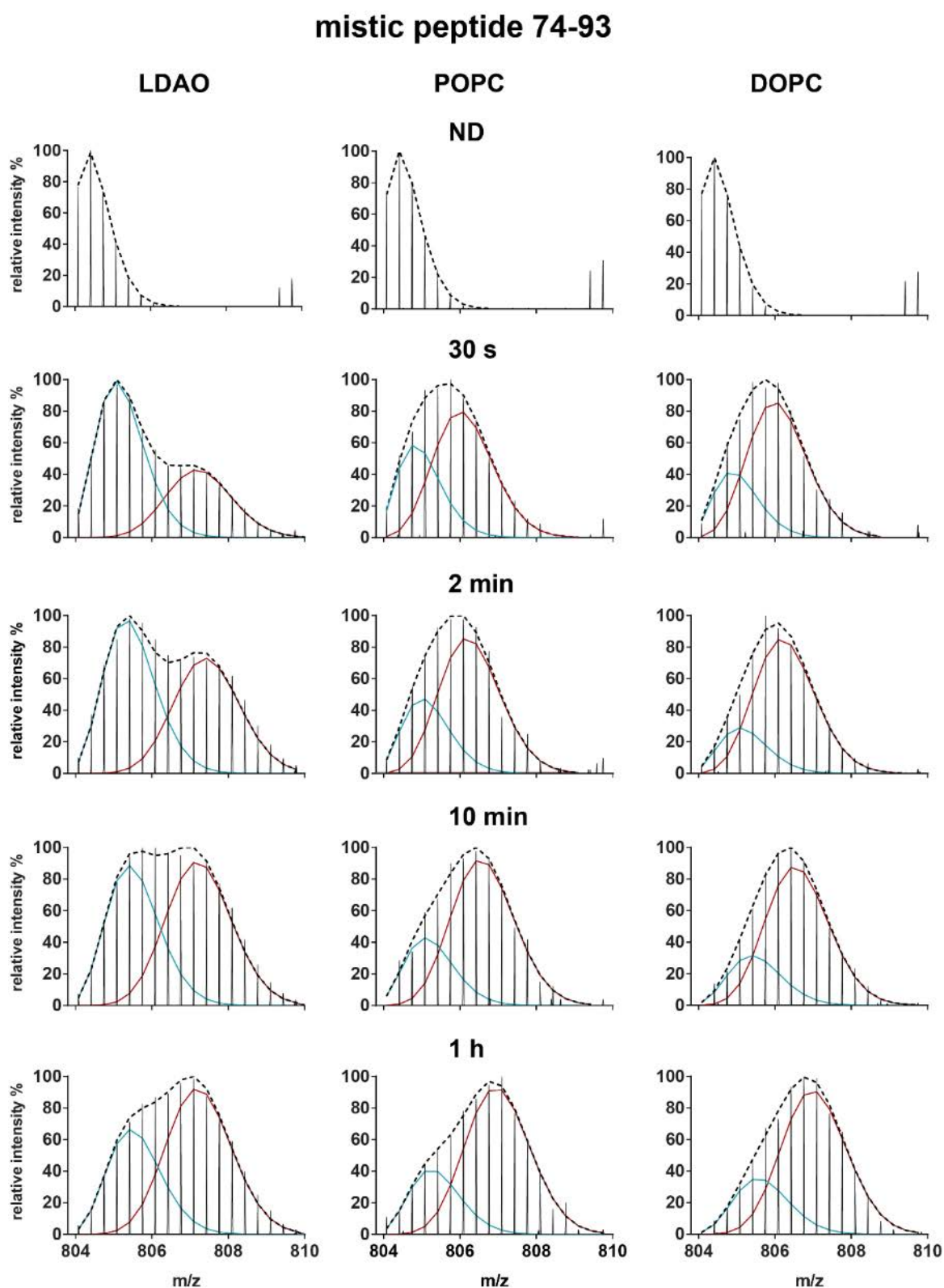


Figure S 25: Binomial fitting of the individual peak distributions of Mistic peptide 74-93 in LDAO, POPC and DOPC at all time points and without deuterium labeling (ND) for comparison. The two individual peak distributions are indicated in blue and red, the dashed line represents the sum of both components at a given *m/z* value.
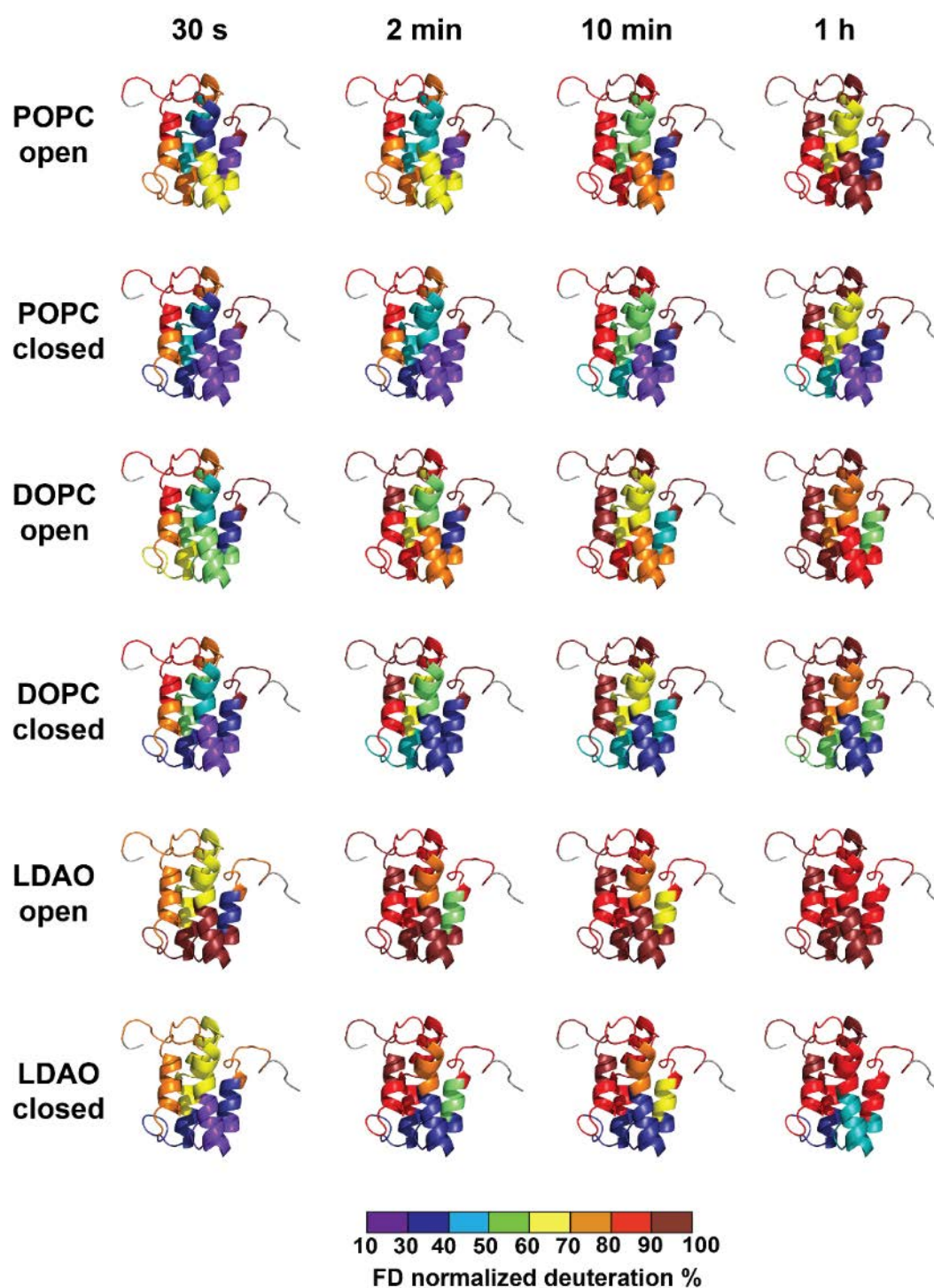
Figure S 26: Mistic deuteration in lipid vesicles (POPC, DOPC) and detergent (LDAO) at all labeling time points. The percentage reflects the FD normalized deuterium uptake averaged over several overlapping peptides in each region. For regions with bimodality deuteration of the "open" and "closed" conformation is displayed. Note that only protein regions shown in Figure 25 show statistically significant deuteration differences between the three conditions. Coloring was applied to visualize the increase in deuteration over time for certain areas of the protein and should only be compared for the "open" and "closed" states within one condition, but not across different conditions.

## 6.1.3 SEQUENCE COVERAGE MAPS
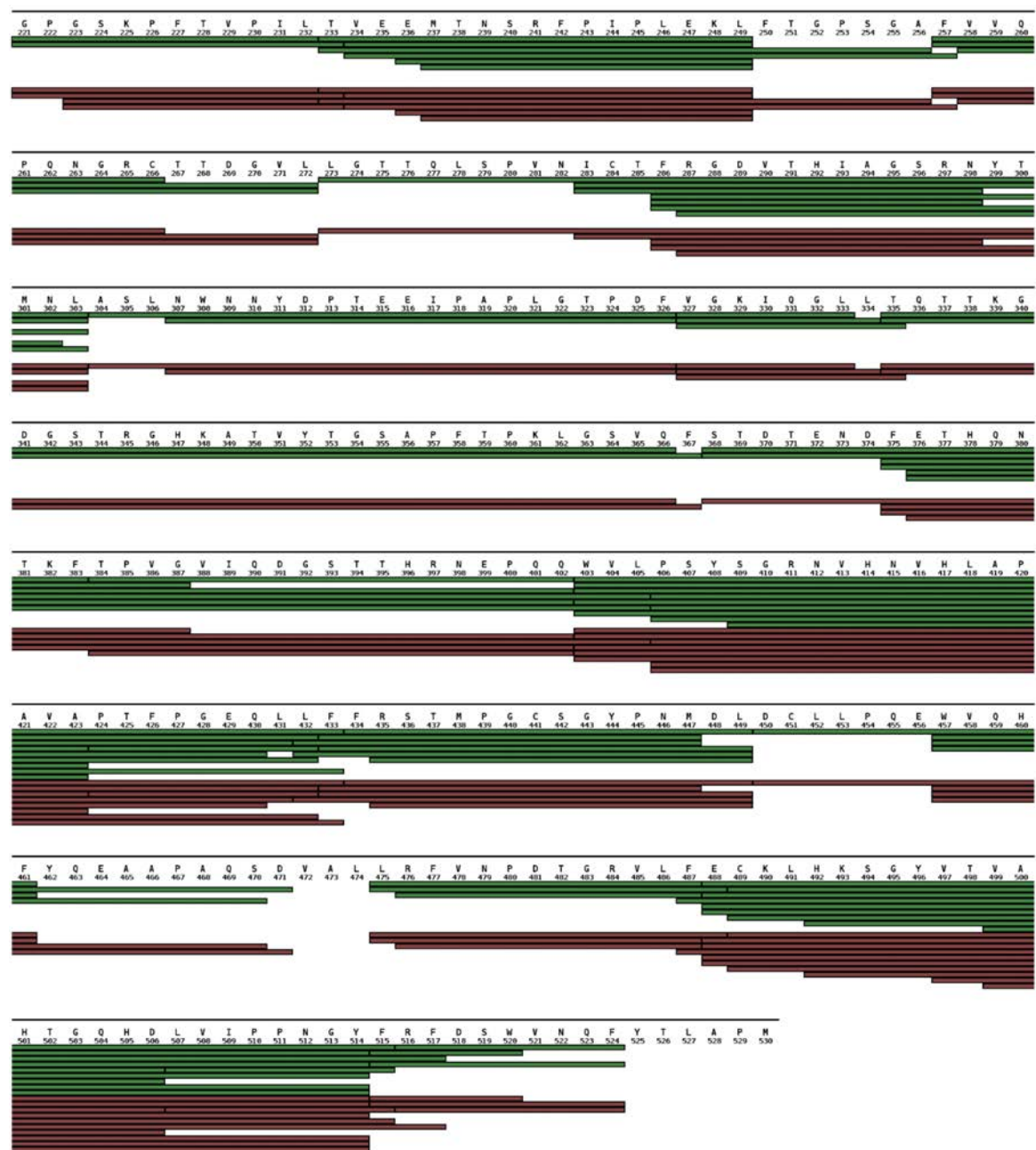


Figure S 27: GII.4 Saga P domain peptide coverage map for the wild type (green) and deamidated protein (red). Peptides for the two missing regions 472-474 and 525-530 were identified in the non-deuterated sample, but data quality in the deuterated samples was insufficient to be included in the HDX MS analysis. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.
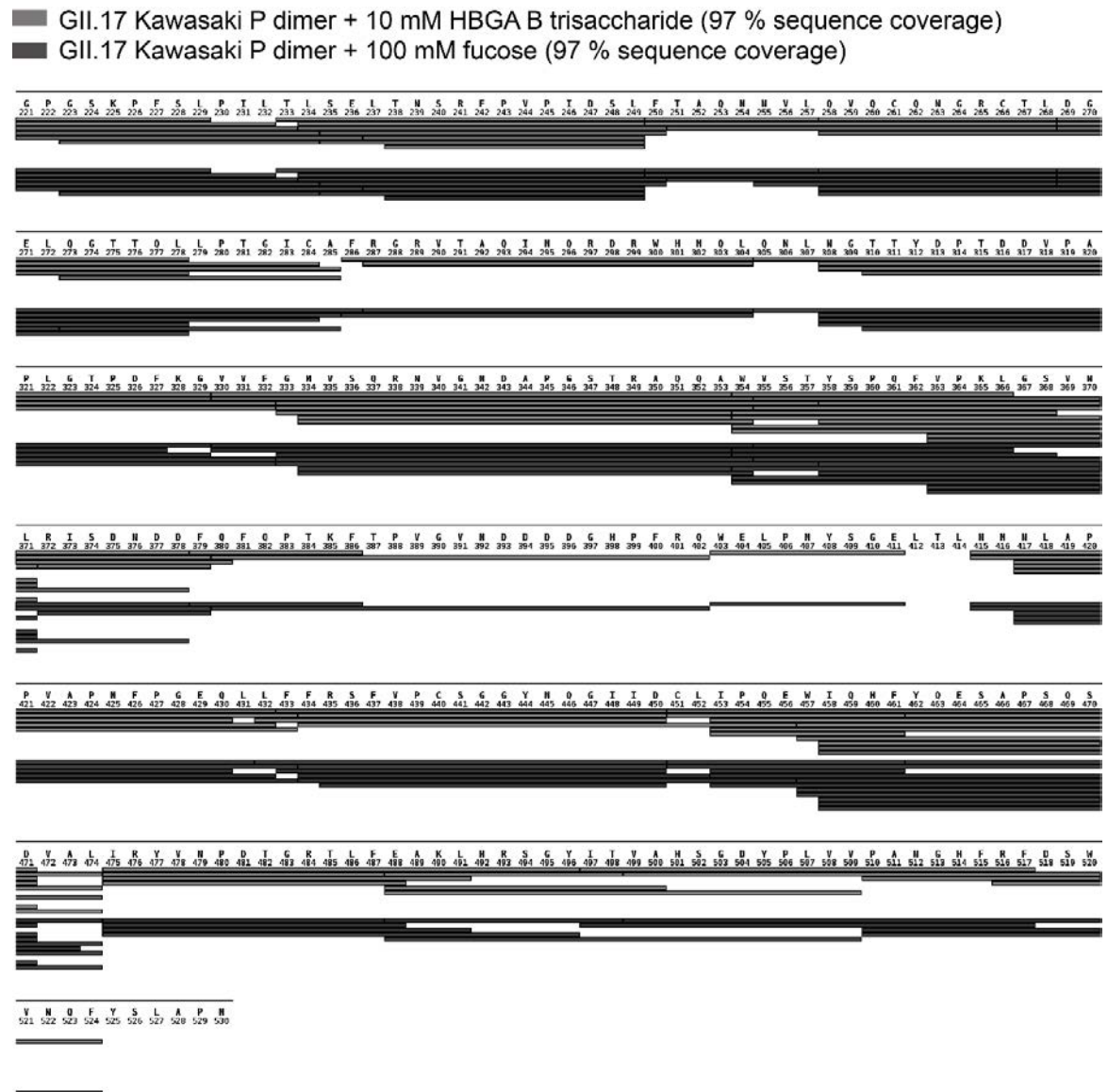
Figure S 28: GII.17 Kawasaki P domain peptide coverage map for HDX-MS experiments with HBGA B trisaccharide (grey) and fucose (dark grey).

Figure S 29: GII.10 Vietnam P domain peptide coverage map for HDX-MS experiments with HBGA B trisaccharide (grey) and fucose (dark grey).

▬ GII.4 MI001 (wt) P dimer + 10 mM HBGA B trisaccharide (96 % sequence coverage)
▬ GII.4 MI001 (wt) P dimer + 100 mM fucose (96 % sequence coverage)
▬ GII.4 MI001 (wt) P dimer + 100 mM fucose (87 % sequence coverage, single time points)
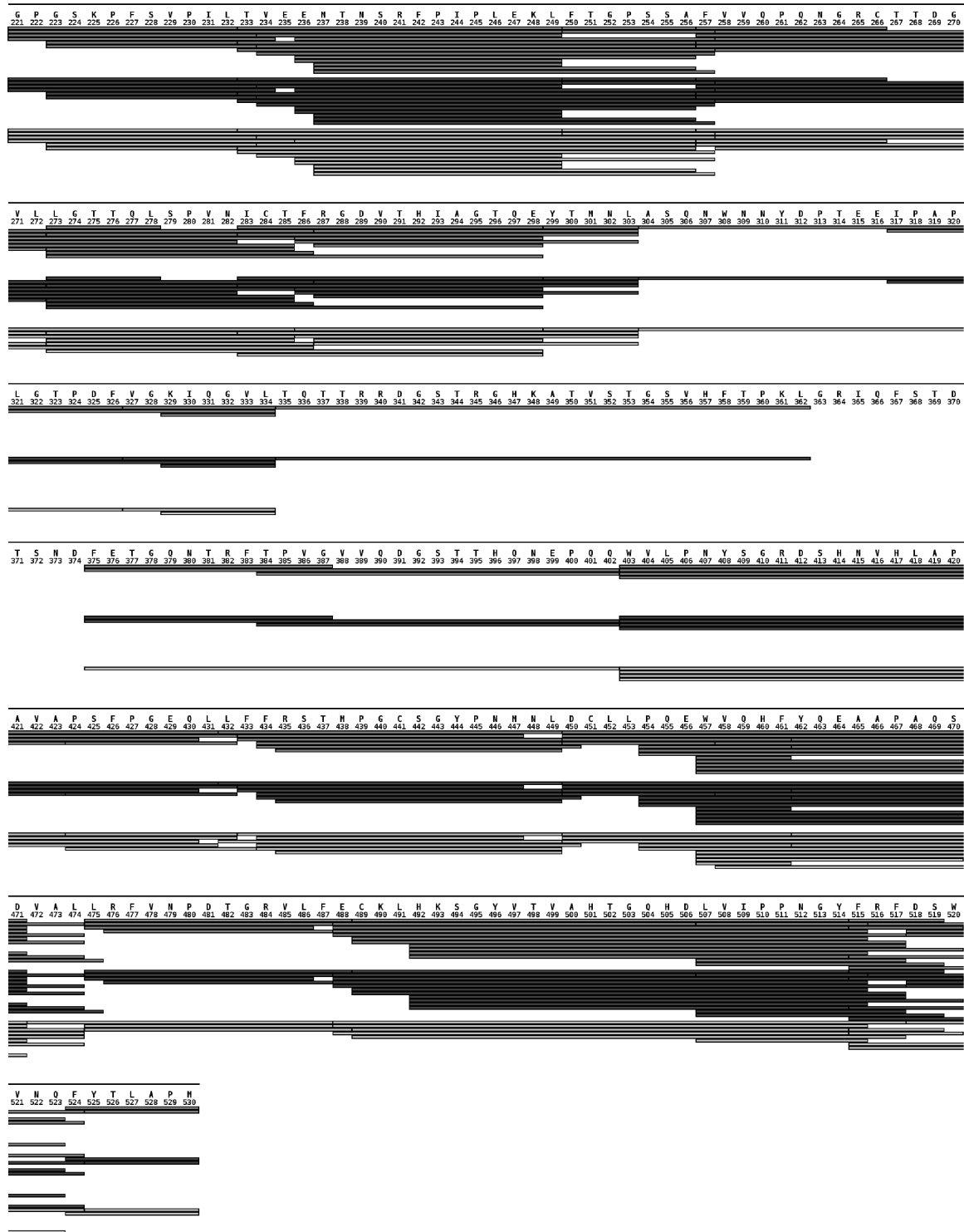
Figure S 30: GII.4 MI001 wild type P domain peptide coverage map for HDX-MS experiments with HBGA B trisaccharide (grey) and fucose triplicate measurement (dark grey) and single measurement (light grey).

■ GII.4 MI001 (partially deam) P dimer + 100 mM fucose (96 % sequence coverage)
■ GII.4 MI001 P dimer wt vs. deam (95 % sequence coverage)
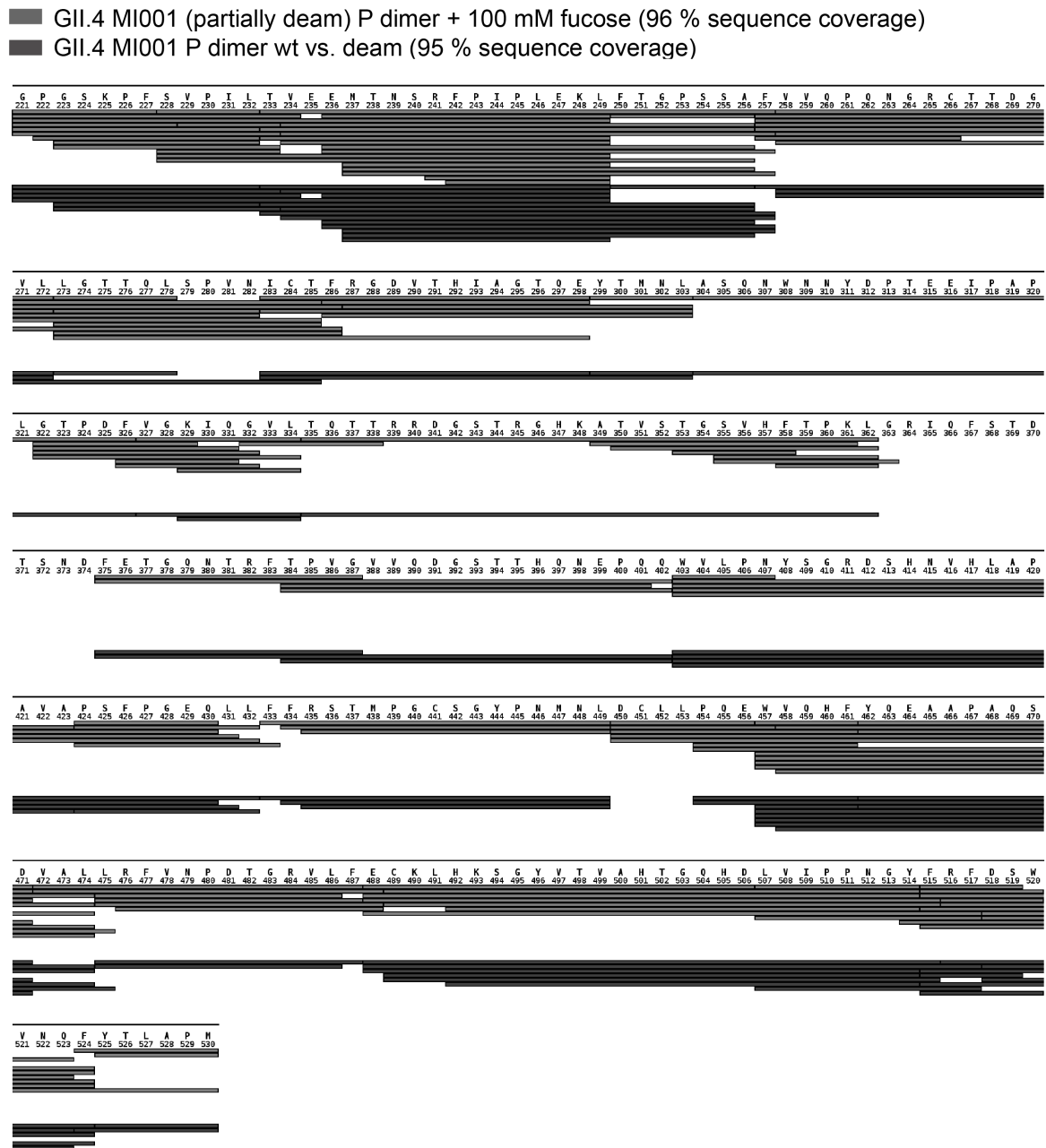


Figure S 31: GII.4 MI001 partially deamidated P domain peptide coverage map for HDX-MS experiments with fucose (grey) and wild type vs. partially deamidated without ligand (dark grey).
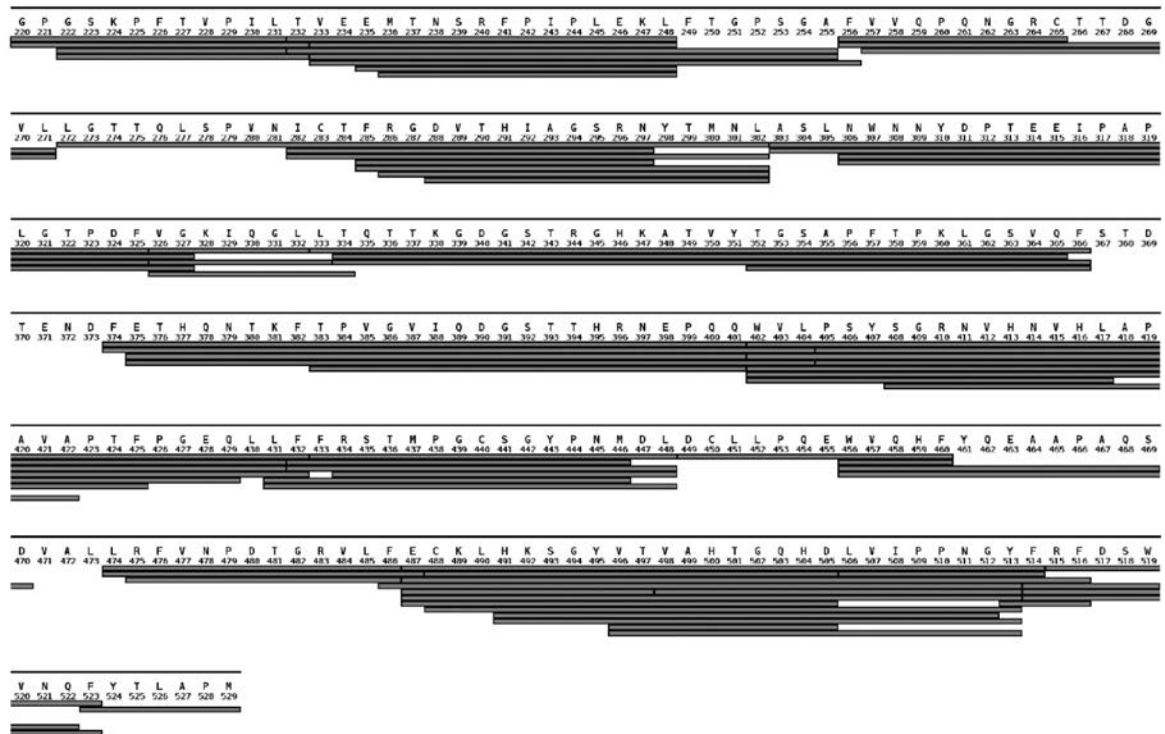
Figure S 32: GII.4 Saga partially deamidated P domain peptide coverage map for the HDX-MS experiment with fucose.
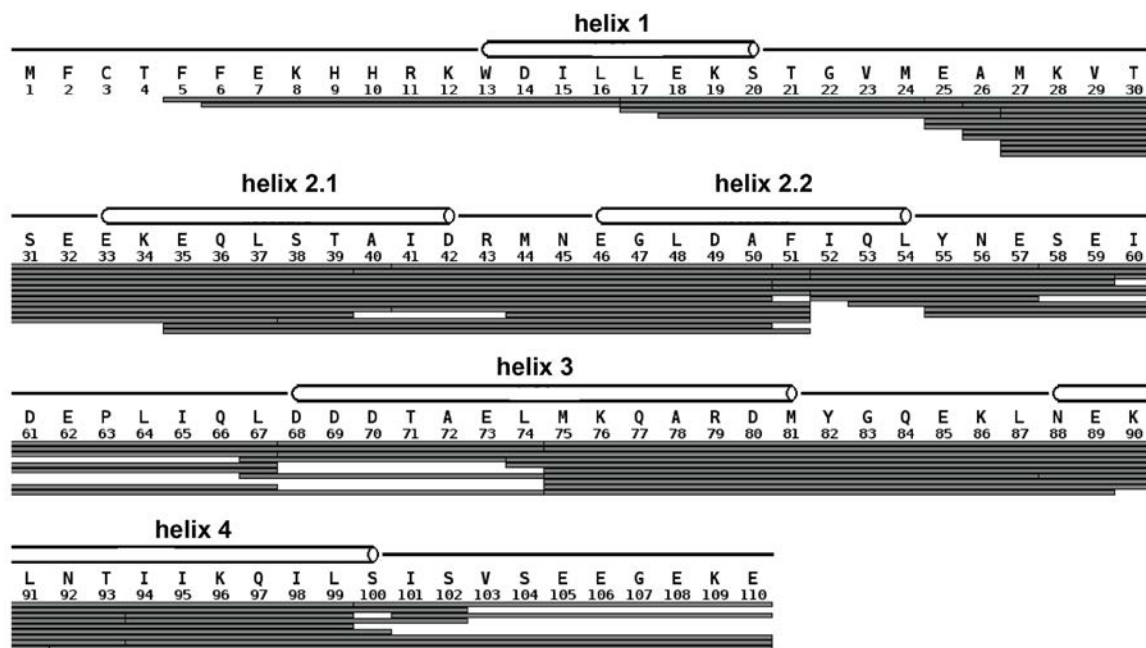
Figure S 33: Mistic peptide coverage map for the HDX-MS experiment in POPC, DOPC and LDAO (96 % coverage).

## 6.2 Supplemental tables

### 6.2.1 NOROVIRUS

Table S 1: Deamidated peptides identified by LC-MS/MS for different norovirus strains (compare Figure S 5 and Figure S 6). In addition to N373 deamidation, a second deamidation site was identified in GII.4 MI001 at either N446 or N448. The asparagine with the highest deamidation probability for the specific peptide is marked in bold. Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.

| Sample | Sequence (charge) | Peptide position | Theoretical *m/z* (wt) | Observed *m/z* | Deamidation site and probability | Fraction % |
|---|---|---|---|---|---|---|
| **GII.4 MI001 IEX-separated wild type fraction** | STDTSNDFETGQNTRFTPVG (2) | 368-387 | 1087.4852 | 1087.4844 | - | 100 |
| | FRSTMPGCSGYP**N**MNL (2) | 434-449 | 887.8891 | 887.8887 | - | 96 |
| | | | | 888.3805 | N446 (0.876) | 4 |
| | FFRSTMPGCSGYP**N**MNL (2) | 433-449 | 961.4233 | 961.4228 | - | 96 |
| | | | | 961.9153 | N446 (0.5) | 4 |
| **GII.4 MI001 IEX-separated deamidated fraction** | STDTS**N**DFETGQNTRFTPVG (2) | 368-387 | 1087.4852 | 1087.9762 | N373 (0.971) | 100 |
| | FRSTMPGCSGYPNM**N**L (2) | 434-449 | 887.8891 | 887.8887 | - | 86 |
| | | | | 888.3811 | N448 (0.546) | 14 |
| | FFRSTMPGCSGYPNM**N**L (2) | 433-449 | 961.4233 | 961.4228 | - | 83 |
| | | | | 961.9147 | N448 (0.58) | 17 |
| | RSTMPGCSGYP**N**MNL (2) | 435-449 | 814.3549 | 814.3545 | - | 86 |
| | | | | 814.8469 | N446 (0.688) | 14 |
| **GII.17 Kawasaki not incubated** | VNLRISDNDDF (2) | 369-379 | 654.3149 | 654.3139 | - | 100 |
| | LRISDNDDF (2) | 371-379 | 547.7593 | 547.7589 | - | 100 |
| | LRISDNDDFQ (2) | 371-380 | 611.7885 | 611.7877 | - | 100 |
| **GII.17 Kawasaki incubated** | VNLRISDNDDF (2) | 369-379 | 654.3149 | 654.3144 | - | 100 |
| | LRISDNDDF (2) | 371-379 | 547.7593 | 547.7593 | - | 100 |

| Sample | Sequence (charge) | Peptide position | Theoretical $m/z$ (wt) | Observed $m/z$ | Deamidation site and probability | Fraction % |
|---|---|---|---|---|---|---|
| **GII.10 Vietnam incubated** | STWETQDVSSGQPTKFTPVG (2) | 379-398 | 1076.5133 | 1076.5128 | - | 100 |
| | STWETQDVSSGQPTKFTPVGL (3) | 379-399 | 755.706 | 755.7054 | - | 100 |
| | STWETQDVSSGQPTKFTPVGLA (2) | 379-400 | 1168.5739 | 1168.5732 | - | 100 |
| | STWETQDVSSGQPTKFTPVGLAS (2) | 379-401 | 1212.0899 | 1212.0891 | - | 100 |

Table S 2: MNV GV P domain oligomer masses determined by native MS (N = 4, as shown in Figure 19). All values are given in Da.

| Virus strain | Oligomeric state | Theoretical mass | Experimental mass ± SD | FWHM ± SD |
|---|---|---|---|---|
| **CR10** | Monomer | 33416 | 33434 ± 38 | 12 ± 3 |
| | Dimer | 66832 | 66891 ± 62 | 13 ± 4 |
| | Trimer* | 100248 | 100362 ± 148 | 27 ± 12 |
| | Tetramer* | 133664 | 133895 ± 189 | 57 ± 12 |
| **MNV07** | Monomer | 33396 | 33427 ± 10 | 15 ± 8 |
| | Dimer | 66793 | 66854 ± 21 | 19 ± 13 |
| | Trimer* | 100189 | 100371 ±114 | 26 ± 19 |
| | Tetramer* | 133586 | 133805 ± 112 | 37 ± 27 |

*unspecific artefacts

## 6.2.2 ORBITRAP FUSION MS METHODS

Table S 3: MS method for HDX measurement

| Global Settings | Experiment Settings |
|---|---|
| Use Ion Source Settings from Tune = False | Experiment Name = MS |
| Method Duration (min)= 14 | Start Time (min) = 0 |
| Ion Source Type = H-ESI | End Time (min) = 14 |
| Spray Voltage: Positive Ion (V) = 3800 | Cycle Time (sec) = 3 |
| Spray Voltage: Negative Ion (V) = 2900 | **Master Scan** |
| Sheath Gas (Arb) = 6 | MSn Level = 1 |
| Aux Gas (Arb) = 0 | Use Wide Quad Isolation = True |
| Infusion Mode (LC)= False | Detector Type = Orbitrap |
| Sweep Gas (Arb) = 0 | Orbitrap Resolution = 120K |
| Ion Transfer Tube Temp (°C) = 275 | Mass Range = Normal |
| Vaporizer Temp (°C) = 20 | Scan Range (m/z) = 300-2000 |
| APPI Lamp = Not in use | Maximum Injection Time (ms) = 50 |
| FAIMS Mode = Not Installed | AGC Target = 200000 |
| Internal Mass Calibration= User Defined Lock Mass | Microscans = 4 |
| Application Mode = Peptide | RF Lens (%) = 60 |
| Default Charge State = 1 | Use ETD Internal Calibration = False |
| Advanced Peak Determination = False | DataType = Profile |
| Xcalibur AcquireX enabled for method modifications = False | Polarity = Positive |
| Internal Cal Positive (m/z) = 445.12 | Source Fragmentation = False |

Table S 4: MS/MS method for peptide and PTM identification

| Global Settings | Experiment Settings |
|---|---|
| Use Ion Source Settings from Tune = False | Experiment Name = MS |
| Method Duration (min)= 34 | Start Time (min) = 0 |
| Ion Source Type = H-ESI | End Time (min) = 34 |
| Spray Voltage: Positive Ion (V) = 3800 | Cycle Time (sec) = 3 |
| Spray Voltage: Negative Ion (V) = 2100 | **Master Scan** |
| Sheath Gas (Arb) = 6 | MSn Level = 1 |
| Aux Gas (Arb) = 0 | Use Wide Quad Isolation = True |
| Infusion Mode (LC)= False | Detector Type = Orbitrap |
| Sweep Gas (Arb) = 0 | Orbitrap Resolution = 240K |
| Ion Transfer Tube Temp (°C) = 275 | Mass Range = Normal |
| Vaporizer Temp (°C) = 20 | Scan Range (m/z) = 300-2000 |
| APPI Lamp = Not in use | Maximum Injection Time (ms) = 120 |
| FAIMS Mode = Not Installed | AGC Target = 200000 |
| Application Mode = Peptide | Microscans = 1 |
| Default Charge State = 1 | RF Lens (%) = 60 |

Table S 4: continued

| Global Settings | Master Scan |
|---|---|
| Advanced Peak Determination = False | Use ETD Internal Calibration = False |
| Xcalibur AcquireX enabled for method modifications = False | DataType = Profile |
| | Polarity = Positive |
| | Source Fragmentation = False |
| MS/MS settings | |
| **Filter MIPS** | **Scan ddMSnScan** |
| MIPS Mode = Peptide | MSn Level = 2 |
| Filter Charge State | Isolation Mode = Quadrupole |
| Include charge state(s) = 1-6 | Isolation Offset = Off |
| Include undetermined charge states = False | Isolation Window = 1.6 |
| Include charge states 25 and higher = False | Reported Mass = Original Mass |
| Filter Dynamic Exclusion | Multi-notch Isolation = False |
| Exclude after n times = 1 | Scan Range Mode = Auto Normal |
| Exclusion duration (s) = 5 | FirstMass = 120 |
| Mass Tolerance = ppm | Scan Priority= 1 |
| Mass tolerance low = 10 | ActivationType = HCD |
| Mass tolerance high = 10 | Collision Energy Mode = Fixed |
| Exclude isotopes = True | Collision Energy (%) = 30 |
| Perform dependent scan on single charge state per precursor only = False | Detector Type = IonTrap |
| **Data Dependent Properties** | Ion Trap Scan Rate = Rapid |
| Data Dependent Mode= Cycle Time | Maximum Injection Time (ms) = 60 |
| Filter IntensityThreshold | AGC Target = 10000 |
| Maximum Intensity = 1E+20 | Inject ions for all available parallelizable time = True |
| Minimum Intensity = 10000 | Microscans = 1 |
| Relative Intensity Threshold = 0 | Use ETD Internal Calibration = False |
| Intensity Filter Type = IntensityRange | DataType = Centroid |
| | Polarity = Positive |
| | Source Fragmentation = False |

## 6.2.3   HDX SUMMARY TABLES

Table S 5: HDX summary tables for measurements in chapter 3.1. Dataset numbering according to deposited data on PRIDE (chapter 5.7). Adapted from Mallagaray et al. (2019) under CC BY 4.0 license.

| | GII.4 Saga P dimer wild type | |
|---|---|---|
| **Dataset 1** | **unbound** | **10 mM B trisaccharide** |
| HDX reaction details | 20 mM Tris buffer, pH= 7.4, 150 mM NaCl, 25°C | 20 mM Tris buffer, pH= 7.4, 150 mM NaCl, 25°C |
| HDX time course (min) | 1, 10, 60, 480 | 1, 10, 60, 480 |
| HDX control samples | - | - |
| Back-exchange (mean) | mean 37% | mean 37% |
| # of Peptides | 68 | 68 |
| Sequence coverage | 97% | 97% |
| Average peptide length / Redundancy | 18.8 / 4.12 | 18.8 / 4.12 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability | 0.083 (average standard deviation) | 0.111 (average standard deviation) |
| Significant differences in HDX (delta HDX > X D) | t-test with α=0.01 was used to compare significant differences between both states at a given time point | |

Table S 5: continued

| | GII.4 Saga P dimer fully deamidated | |
|---|---|---|
| **Dataset 2** | **unbound** | **10 mM B trisaccharide** |
| HDX reaction details | 20 mM Tris buffer, pH= 7.4, 150 mM NaCl, 25°C | |
| HDX time course (min) | 1, 10, 60, 480 | |
| HDX control samples | - | |
| Back-exchange (mean) | mean 37% | |
| # of Peptides | 65 | |
| Sequence coverage | 97% | |
| Average peptide length / Redundancy | 18.6 / 3.87 | |
| Replicates (biological or technical) | 3 (technical) | |
| Repeatability | 0.126 (average standard deviation) | 0.06 (average standard deviation) |
| Significant differences in HDX (delta HDX > X D) | t-test with α=0.01 was used to compare significant differences between both states at a given time point | |

| | GII.4 Saga P dimer | |
|---|---|---|
| **Dataset 3** | **wild type (unbound)** | **deamidated (unbound)** |
| HDX reaction details | 20 mM Tris buffer, pH= 7.4, 150 mM NaCl, 25°C | 20 mM Tris buffer, pH= 7.4, 150 mM NaCl, 25°C |
| HDX time course (min) | 1, 10, 60, 480 | 1, 10, 60, 480 |
| HDX control samples | - | - |
| Back-exchange (mean) | mean 37% | mean 37% |
| # of Peptides | 65 | 65 |
| Sequence coverage | 97% | 97% |
| Average peptide length / Redundancy | 18.6 / 3.87 | 18.6 / 3.87 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability | 0.081 (average standard deviation) | 0.126 (average standard deviation) |
| Significant differences in HDX (delta HDX > X D) | combination of t-test with α=0.01 and delta HDX > 0.64 D (99% percentile calculated according to Arora et al. ) | |

Table S 5: continued

| | GII.4 Saga P dimer wild type | |
|---|---|---|
| **Dataset 4** | **unbound** | **10 mM B trisaccharide** |
| HDX reaction details | 20 mM Tris buffer, pH= 7.4, 150 mM NaCl, 25°C | 20 mM Tris buffer, pH= 7.4, 150 mM NaCl, 25°C |
| HDX time course (min) | 0.25, 1, 10, 60, 480 | 0.25, 1, 10, 60, 480 |
| HDX control samples | - | - |
| Back-exchange (mean) | mean 37% | mean 37% |
| # of Peptides | 72 | 72 |
| Sequence coverage | 99% | 99% |
| Average peptide length / Redundancy | 18.5 / 4.29 | 18.5 / 4.29 |
| Replicates (biological or technical) | 1 (technical) | 1 (technical) |
| Repeatability | - | - |
| Significant differences in HDX (delta HDX > X D) | - | - |
| | **100 mM methyl-fucose** | **100 mM galactose** |
| HDX reaction details | 20 mM Tris buffer, pH= 7.4, 150 mM NaCl, 25°C | 20 mM Tris buffer, pH= 7.4, 150 mM NaCl, 25°C |
| HDX time course (min) | 0.25, 1, 10, 60, 480 | 0.25, 1, 10, 60, 480 |
| HDX control samples | - | - |
| Back-exchange (mean / IQR) | mean 37% | mean 37% |
| # of Peptides | 72 | 72 |
| Sequence coverage | 99% | 99% |
| Average peptide length / Redundancy | 18.5 / 4.29 | 18.5 / 4.29 |
| Replicates (biological or technical) | 1 (technical) | 1 (technical) |
| Repeatability | - | - |
| Significant differences in HDX (delta HDX > X D) | - | - |

Table S 6: HDX summary tables for measurements in chapter 3.2. Dataset numbering according to deposited data on PRIDE (chapter 5.7).

| | GII.10 Vietnam P dimer | |
|---|---|---|
| **Data Set T1** | **unbound** | **100 mM fucose** |
| HDX reaction details | 20 mM Tris, 150 mM NaCl, pH 7, 25 °C | |
| HDX time course (min) | 1, 10, 60, 480 | |
| HDX control samples | FD model peptides in T9 | |
| Back-exchange (mean) | NA | |
| # of Peptides | 117 | 117 |
| Sequence coverage | 91% | 91% |
| Average peptide length / Redundancy | 15.9 / 5.8 | 15.9 / 5.8 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability (average standard deviation) | 0.1 | 0.073 |
| Significant differences in HDX (delta HDX > X D) | T test ($p<0.05$) and delta D > 0.25 (2x pooled average SD), delta D > 0.55 (99% percentile) for 8h time point single measurement, manual validation in case of bimodality | |

| | GII.10 Vietnam P dimer | |
|---|---|---|
| **Data Set T2** | **unbound** | **10 mM HBGA B trisaccharide** |
| HDX reaction details | 20 mM Tris, 150 mM NaCl, pH 7, 25 °C | |
| HDX time course (min) | 1, 10, 60, 480 | |
| HDX control samples | Fully deuterated protein control, labeled for 24 h, buffer 20 mM Tris, 6 M urea, pH 7 + FD model peptides in T9 | |
| Back-exchange (mean) | 50 ± 15 % (FD protein) | |
| # of Peptides | 126 | 126 |
| Sequence coverage | 89% | 89% |
| Average peptide length / Redundancy | 16.4 / 6.49 | 16.4 / 6.49 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability (average standard deviation) | 0.12 | 0.178 |
| Significant differences in HDX (delta HDX > X D) | T test ($p<0.05$) and delta D > 0.43 (2x pooled average SD), manual validation in case of bimodality | |

Table S 6: continued

| | GII.17 Kawasaki P dimer | |
|---|---|---|
| **Data Set T3** | **unbound** | **100 mM fucose** |
| HDX reaction details | 20 mM Tris, 150 mM NaCl, pH 7, 25 °C | |
| HDX time course (min) | 1, 10, 60, 480 | |
| HDX control samples | Fully deuterated protein control, labeled for 72 h, buffer 20 mM Tris, 6 M urea, pH 7 + FD model peptides in T9 | |
| Back-exchange (mean) | 64 ± 11 % (FD protein) | |
| # of Peptides | 89 | 89 |
| Sequence coverage | 97% | 97% |
| Average peptide length / Redundancy | 16.3 / 4.69 | 16.3 / 4.69 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability (average standard deviation) | 0.057 | 0.083 |
| Significant differences in HDX (delta HDX > X D) | T test (p<0.05) and delta D > 0.20 (2x pooled average SD), manual validation in case of bimodality | |

| | GII.17 Kawasaki P dimer | |
|---|---|---|
| **Data Set T4** | **unbound** | **10 mM HBGA B trisaccharide** |
| HDX reaction details | 20 mM Tris, 150 mM NaCl, pH 7, 25 °C | |
| HDX time course (min) | 1, 10, 60, 480 | |
| HDX control samples | Fully deuterated protein control, labeled for 72 h, buffer 20 mM Tris, 6 M urea, pH 7 + FD model peptides in T9 | |
| Back-exchange (mean) | 58 ± 14 % (FD protein) | |
| # of Peptides | 84 | 84 |
| Sequence coverage | 97% | 97% |
| Average peptide length / Redundancy | 16.1 / 4.36 | 16.1 / 4.36 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability (average standard deviation) | 0.115 | 0.157 |
| Significant differences in HDX (delta HDX > X D) | T test (p<0.05) and delta D > 0.39 (2x pooled average SD), manual validation in case of bimodality | |

Table S 6: continued

| Data Set T5 | GII.4 MI001 P dimer wild type | |
| --- | --- | --- |
| | **unbound** | **10 mM HBGA B trisaccharide** |
| HDX reaction details | 20 mM Tris, 150 mM NaCl, pH 7, 25 °C | |
| HDX time course (min) | 1, 10, 60, 480 | |
| HDX control samples | Fully deuterated protein control, labeled for 72 h, buffer 20 mM Tris, 6 M urea, pH 7 + FD model peptides in T9 | |
| Back-exchange (mean) | 47 ± 12 % (FD protein) | |
| # of Peptides | 100 | 100 |
| Sequence coverage | 96% | 96% |
| Average peptide length / Redundancy | 16.9 / 5.43 | 16.9 / 5.43 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability (average standard deviation) | 0.08 | 0.073 |
| Significant differences in HDX (delta HDX > X D) | T test (p<0.05) and delta D > 0.22 (2x pooled average SD), manual validation in case of bimodality | |
| | **unbound** | **100 mM fucose** |
| HDX reaction details | 20 mM Tris, 150 mM NaCl, pH 7, 25 °C | |
| HDX time course (min) | 1, 10, 60, 480 | |
| HDX control samples | Fully deuterated protein control, labeled for 72 h, buffer 20 mM Tris, 6 M urea, pH 7 + FD model peptides in T9 | |
| Back-exchange (mean) | 47 ± 12 % (FD protein) | |
| # of Peptides | 100 | 100 |
| Sequence coverage | 96% | 96% |
| Average peptide length / Redundancy | 16.9 / 5.43 | 16.9 / 5.43 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability (average standard deviation) | 0.08 | 0.09 |
| Significant differences in HDX (delta HDX > X D) | T test (p<0.05) and delta D > 0.23 (2x pooled average SD), manual validation in case of bimodality | |

Table S 6: continued

| Data Set T6 | GII.4 MI001 P dimer wild type | |
|---|---|---|
| | **unbound** | **100 mM fucose** |
| HDX reaction details | 20 mM Tris, 150 mM NaCl, pH 7, 25 °C | |
| HDX time course  (min) | 0.25, 1, 10, 60, 480 | |
| HDX control samples | Fully deuterated protein control, labeled for 24 h, buffer 20 mM Tris, 6 M urea, pH 7 + FD model peptides in T9 | |
| Back-exchange (mean) | 40 ± 8 % (FD protein) | |
| # of Peptides | 81 | 81 |
| Sequence coverage | 87% | 87% |
| Average peptide length / Redundancy | 16.6 / 4.34 | 16.6 / 4.34 |
| Replicates (biological or technical) | 1 (technical) | 1 (technical) |
| Repeatability   (average standard deviation) | 0.1 for FD (2 technical replicates) | - |
| Significant   differences in HDX (delta HDX > X D) | delta D > 0.5 (99% percentile of dataset T5 with similar SD), manual validation in case of bimodality | |

| Data Set T7 | GII.4 MI001 P dimer partially deamidated | |
|---|---|---|
| | **unbound** | **100 mM fucose** |
| HDX reaction details | 20 mM Tris, 150 mM NaCl, pH 7, 25 °C | |
| HDX time course  (min) | 1, 10, 60, 480 | |
| HDX control samples | Fully deuterated protein control, labeled for 72 h, buffer 20 mM Tris, 6 M urea, pH 7 + FD model peptides in T9 | |
| Back-exchange (mean) | 41 ± 13 % (FD protein) | |
| # of Peptides | 123 | 123 |
| Sequence coverage | 96% | 96% |
| Average peptide length / Redundancy | 15.0 / 5.94 | 15.0 / 5.94 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability   (average standard deviation) | 0.074 | 0.079 |
| Significant   differences in HDX (delta HDX > X D) | T test ($p<0.05$) and delta D > 0.21 (2x pooled average SD), manual validation in case of bimodality | |

Table S 6: continued

| | GII.4 MI001 P dimer wild type vs. partially deamidated | |
|---|---|---|
| **Data Set T8** | **unbound (wt) (T5)** | **unbound (deam) (T7)** |
| HDX reaction details | 20 mM Tris, 150 mM NaCl, pH 7, 25 °C | |
| HDX time course (min) | 1, 10, 60, 480 | |
| HDX control samples | fully deuterated protein control (FD T5 and T7) + FD model peptides in T9 | |
| Back-exchange (mean) | see T5 and T7 for FD protein | |
| # of Peptides | 69 | 69 |
| Sequence coverage | 95% | 95% |
| Average peptide length / Redundancy | 16.4 / 3.65 | 16.4 / 3.65 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability (average standard deviation) | 0.081 | 0.072 |
| Significant differences in HDX (delta HDX > X D) | delta D normalized with FD controls ratio, T test ($p<0.05$) and delta D > 0.42 (99% percentile) | |

| | fully deuterated model peptides for quality control | |
|---|---|---|
| **Data Set T9** | **Angiotensin I** | **Bradykinin** |
| HDX reaction details | 20 mM deuterated Tris, 6 M GndHCl, pH 7, 25 °C, labeled for 24 h | |
| HDX time course (min) | - | |
| HDX control samples | fully deuterated peptide mix serves as general back exchange control for the standard bottom-up HDX MS workflow | |
| Back-exchange (mean) | 23 ± 1% (FD peptide) | 32 ± 2 % (FD peptide) |
| # of Peptides | 1 | 1 |
| Sequence coverage | 100% | 100% |
| Average peptide length / Redundancy | 10 | 9 |
| Replicates (biological or technical) | 2 (technical) | 2 (technical) |
| Repeatability (average standard deviation) | 0.1 | 0.08 |
| Significant differences in HDX (delta HDX > X D) | - | |

Table S 6: continued

| | GII.4 Saga P dimer partially deamidated | |
|---|---|---|
| **Data Set T10** | **unbound** | **100 mM fucose** |
| HDX reaction details | 20 mM Tris, 150 mM NaCl, pH 7, 25 °C | |
| HDX time course (min) | 0.25. 1, 10 | |
| HDX control samples | Fully deuterated protein control, labeled for 24 h, buffer 20 mM Tris, 6 M urea, pH 7 + FD model peptides in T9 | |
| Back-exchange (mean) | 37 ± 9 % (FD protein) | |
| # of Peptides | 78 | 78 |
| Sequence coverage | 97% | 97% |
| Average peptide length / Redundancy | 18.2 / 4.59 | 18.2 / 4.59 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability (average standard deviation) | 0.084 | 0.109 |
| Significant differences in HDX (delta HDX > X D) | T test (p<0.05) and delta D > 0.27 (2x pooled average SD) | |

Table S 7: HDX summary tables for measurements in chapter 3.4.

| | fully deuterated model peptides for quality control | |
|---|---|---|
| **Dataset 1** | **Angiotensin I** | **Bradykinin** |
| HDX reaction details | 20 mM deuterated Tris, 6 M GndHCl, pH 7, 25 °C, labeled for 24 h | |
| HDX time course (min) | - | |
| HDX control samples | fully deuterated peptide mix serves as general back exchange control for the HDX MS workflow adapted for membrane proteins | |
| Back-exchange (mean) | 59 ± 29 % (FD peptide) | 66 ± 11 % (FD peptide) |
| # of Peptides | 1 | 1 |
| Sequence coverage | 100% | 100% |
| Average peptide length / Redundancy | 10 | 9 |
| Replicates (biological or technical) | 2 (technical) | 2 (technical) |
| Repeatability (average standard deviation) | 0.1 | 0.08 |
| Significant differences in HDX (delta HDX > X D) | - | |

Table S 7: continued

| | Mistic | |
|---|---|---|
| **Dataset 2** | **DOPC** | **POPC** |
| HDX reaction details | 50 mM Tris, pH 7.4, 50 mM NaCl, 25 °C | |
| HDX time course  (min) | 0.5, 2, 10, 60 | |
| HDX control samples | Fully deuterated protein control in 48 μM LDAO, labeled for 24 h, buffer 50 mM Tris, 50 mM NaCl, pH 7.4+ FD model peptides in T11 | |
| Back-exchange (mean) | 54 ± 7 % (FD protein) | |
| # of Peptides | 54 | 54 |
| Sequence coverage | 96% | 96% |
| Average peptide length / Redundancy | 15.2 / 7.5 | 15.2 / 7.5 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability  (average standard deviation) | 0.18 | 0.27 |
| Significant differences in HDX (delta HDX > X D) | T test (p<0.05) and delta D > 0.63 (2x pooled average SD), manual validation in case of bimodality | |
| | **LDAO** | **DOPC** |
| HDX reaction details | 50 mM Tris, pH 7.4, 50 mM NaCl, 25 °C, additionally 3 mM LDAO for LDAO sample | |
| HDX time course  (min) | 0.5, 2, 10, 60 | |
| HDX control samples | Fully deuterated protein control in 48 μM LDAO, labeled for 24 h, buffer 50 mM Tris, 50 mM NaCl, pH 7.4+ FD model peptides in T11 | |
| Back-exchange (mean) | 54 ±  7 % (FD protein) | |
| # of Peptides | 54 | 54 |
| Sequence coverage | 96% | 96% |
| Average peptide length / Redundancy | 15.2 / 7.5 | 15.2 / 7.5 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability  (average standard deviation) | 0.21 | 0.18 |
| Significant differences in HDX (delta HDX > X D) | T test (p<0.05) and delta D > 0.55 (2x pooled average SD), manual validation in case of bimodality | |

Table S 7: continued

| Dataset 2 continued | Mistic | |
|---|---|---|
| | **LDAO** | **POPC** |
| HDX reaction details | 50 mM Tris, pH 7.4, 50 mM NaCl, 25 °C, additionally 3 mM LDAO for LDAO sample | |
| HDX time course (min) | 0.5, 2, 10, 60 | |
| HDX control samples | Fully deuterated protein control in 48 µM LDAO, labeled for 24 h, buffer 50 mM Tris, 50 mM NaCl, pH 7.4+ FD model peptides in T11 | |
| Back-exchange (mean) | 54 ± 7 % (FD protein) | |
| # of Peptides | 54 | 54 |
| Sequence coverage | 96% | 96% |
| Average peptide length / Redundancy | 15.2 / 7.5 | 15.2 / 7.5 |
| Replicates (biological or technical) | 3 (technical) | 3 (technical) |
| Repeatability (average standard deviation) | 0.21 | 0.27 |
| Significant differences in HDX (delta HDX > X D) | T test (p<0.05) and delta D > 0.67 (2x pooled average SD), manual validation in case of bimodality | |

## 6.3  Deuterium uptake plots

**Detailed figure descriptions for deuterium uptake plots:**

Figure S 34: Deuterium uptake plots for the wild type P dimers (black) with 10 mM HBGA B Trisaccharide (blue). All time points (1 min, 10 min, 1 h, 8 h) were performed in triplicate. Error bars indicate the standard deviation of each triplicate analysis.

Figure S 35: Deuterium uptake plots for the deamidated P dimers (black) with 10 mM HBGA B trisaccharide (blue). All time points (1 min, 10 min, 1 h, 8 h) were performed in triplicate. Error bars indicate the standard deviation of each triplicate analysis. No time point showed significant ($p < 0.01$) deuteration differences.

Figure S 36: Deuterium uptake plot comparison for the wild type (black) and deamidated P dimers (orange) without any ligand. All time points (1 min, 10 min, 1 h, 8 h) were performed in triplicate. Error bars indicate the standard deviation of each triplicate analysis. Smaller deuteration differences can also be detected in the N-terminal domain loop (residues 221-272). However, these are not considered significant according to the used significance definition.

Figure S 37: Deuterium uptake plots for the wild type P dimers (black) with 10 mM HBGA B trisaccharide (blue), 100 mM methyl α-L-fucopyranoside (red) and 100 mM galactose (green, negative control). All time points (15 s, 1 min, 10 min, 1 h, 8 h) represent single measurements, but deuterium levels are well in line with the triplicate measurements presented in Figure S 34.

Figure S 38: Deuterium uptake plots for Vietnam P dimers (black) with 100 mM fucose (blue) (dataset T1). Time points 1 min, 10 min, and 1 h were performed in triplicate, the 8 h time point with fucose represents a single measurement. Error bars indicate the standard deviation of each triplicate analysis.

Figure S 39: Deuterium uptake plots for Vietnam P dimers (black) with 10 mM HBGA B trisaccharide (blue) (dataset T2). All time points (1 min, 10 min, 1 h and 8 h) were performed in triplicate, the fully deuterated control (FD, red) represents two measurements. Error bars indicate the standard deviation.

Figure S 40: Deuterium uptake plots for Kawasaki P dimers (black) with 100 mM fucose (blue) (dataset T3). All time points (1 min, 10 min, 1 h and 8 h) were

performed in triplicate, the fully deuterated control (FD, red) represents two measurements. Error bars indicate the standard deviation.

Figure S 41: Deuterium uptake plots for Kawasaki P dimers (black) with 10 mM HBGA B trisaccharide (blue) (dataset T4). All time points (1 min, 10 min, 1 h and 8 h) were performed in triplicate, the fully deuterated control (FD, red) represents two measurements. Error bars indicate the standard deviation.

Figure S 42: Deuterium uptake plots for wild type MI001 P dimers (black) with 10 mM HBGA B trisaccharide (blue) and 100 mM fucose (light blue) (dataset T5). All time points (1 min, 10 min, 1 h and 8 h) were performed in triplicate, the fully deuterated control (FD, red) represents two measurements. Error bars indicate the standard deviation.

Figure S 43: Deuterium uptake plots for wild type MI001 P dimers (black) with 100 mM fucose (blue) (dataset T6). All time points (1 min, 10 min, 1 h and 8 h) were performed in single measurements, the fully deuterated control (FD, red) represents two measurements. Error bars indicate the standard deviation.

Figure S 44: Deuterium uptake plots for partially deamidated MI001 P dimers (black) with 100 mM fucose (blue) (dataset T7). All time points (1 min, 10 min, 1 h and 8 h) were performed in triplicate, the fully deuterated control (FD, red) represents two measurements. Error bars indicate the standard deviation.

Figure S 45: Deuterium uptake plot comparison for wild type (black) and partially deamidated MI001 P dimers (orange) without any ligand (dataset T8). All time points (1 min, 10 min, 1 h, 8 h) were performed in triplicate. The ratio of the FD controls from wild type and deamidated measurements was used for normalization to account for different back exchange levels between the two datasets. Error bars indicate the standard deviation.

Figure S 46: Deuterium uptake plots for partially deamidated Saga P dimers (black) with 100 mM fucose (blue) (dataset T10). All time points (15 s, 1 min and 10 min) were performed in triplicate, the fully deuterated control (FD, red) represents two measurements. Error bars indicate the standard deviation.

Figure S 47: Deuterium uptake plots for Mistic in POPC (blue), DOPC (black) and LDAO (orange). All time points (30 s, 2 min, 10 min, 1 h) were performed in triplicate. Error bars indicate the standard deviation.
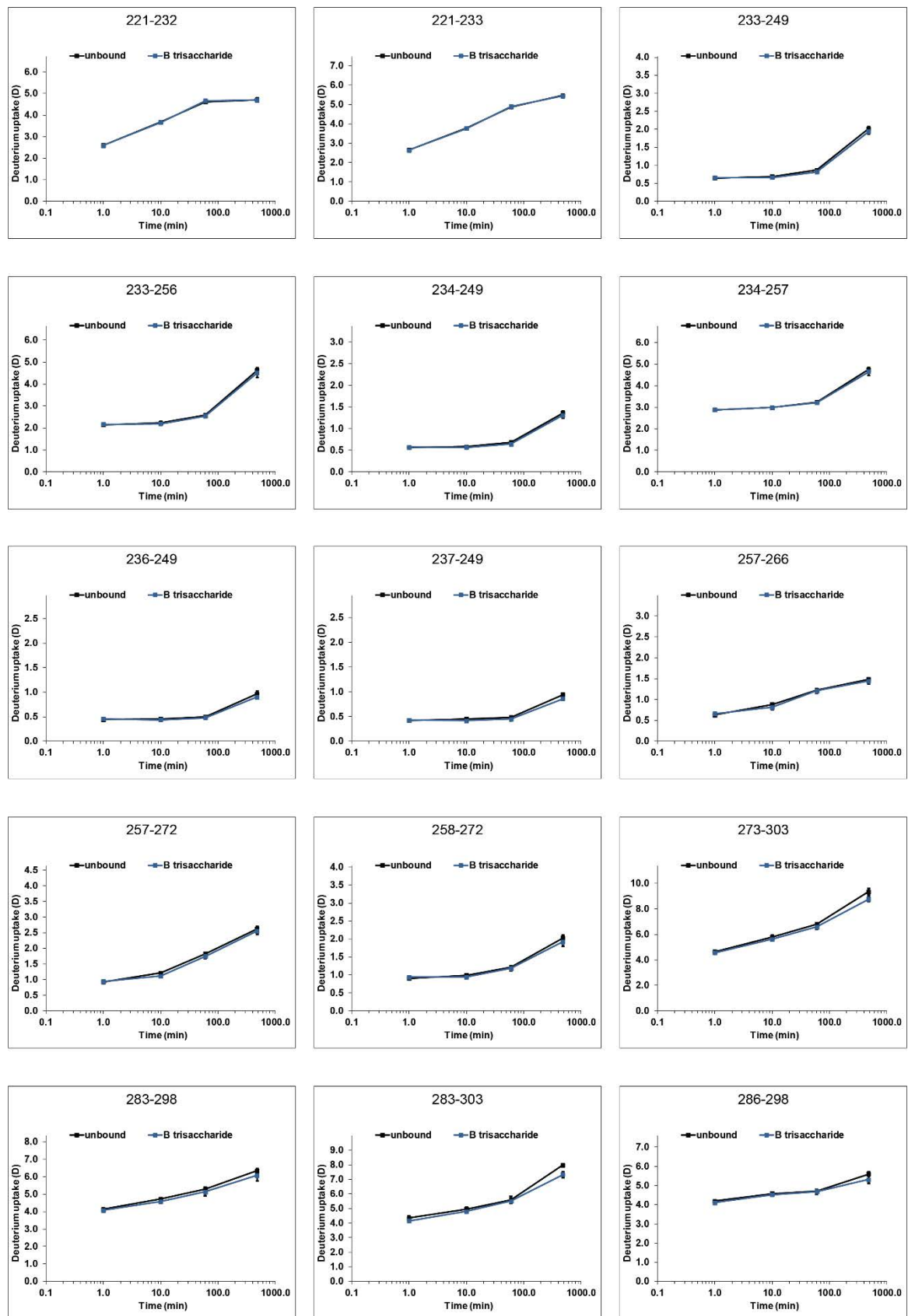
Figure S 34: GII.4 Saga wild type P dimer with 10 mM B trisaccharide
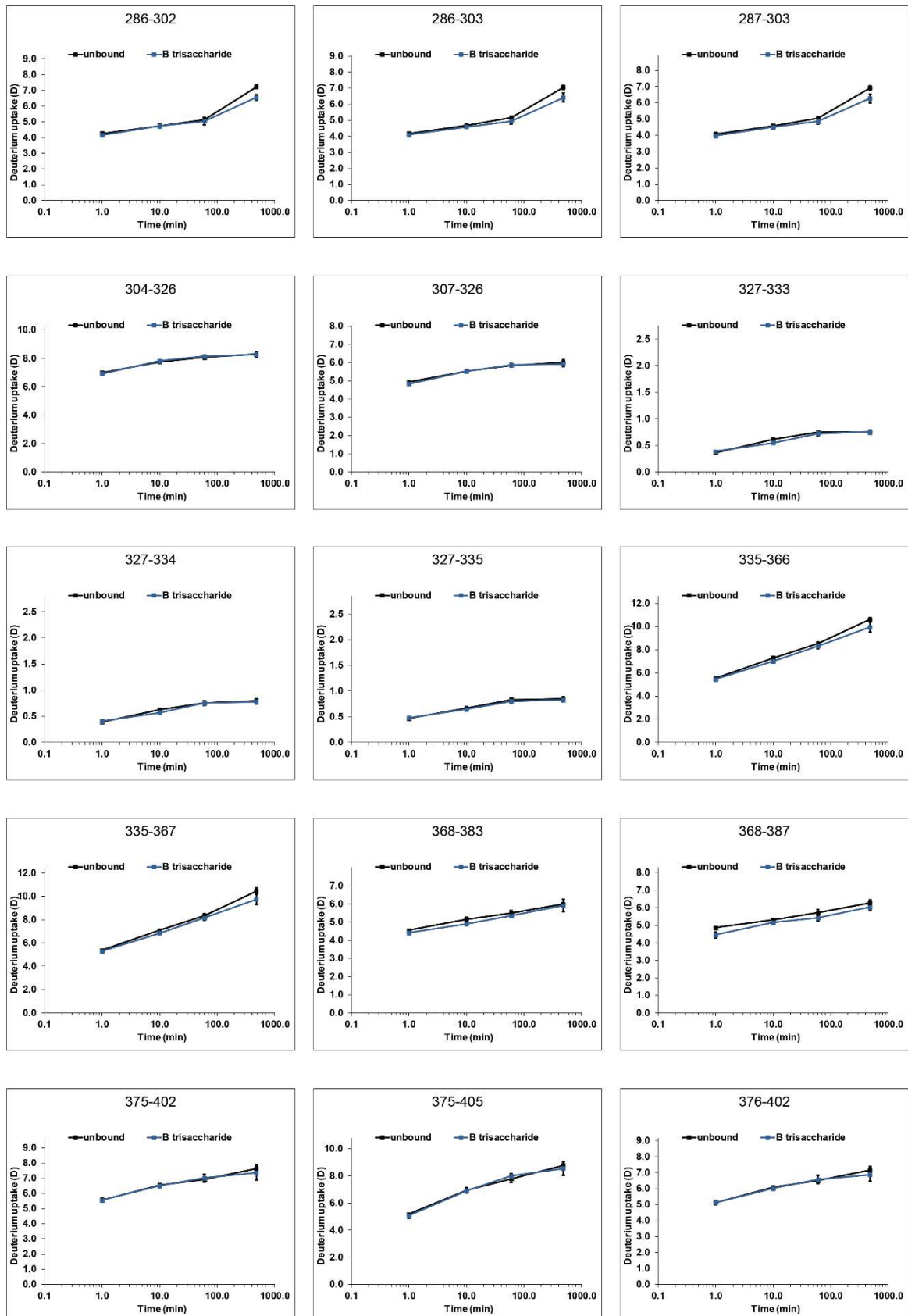
Figure S 34: continued
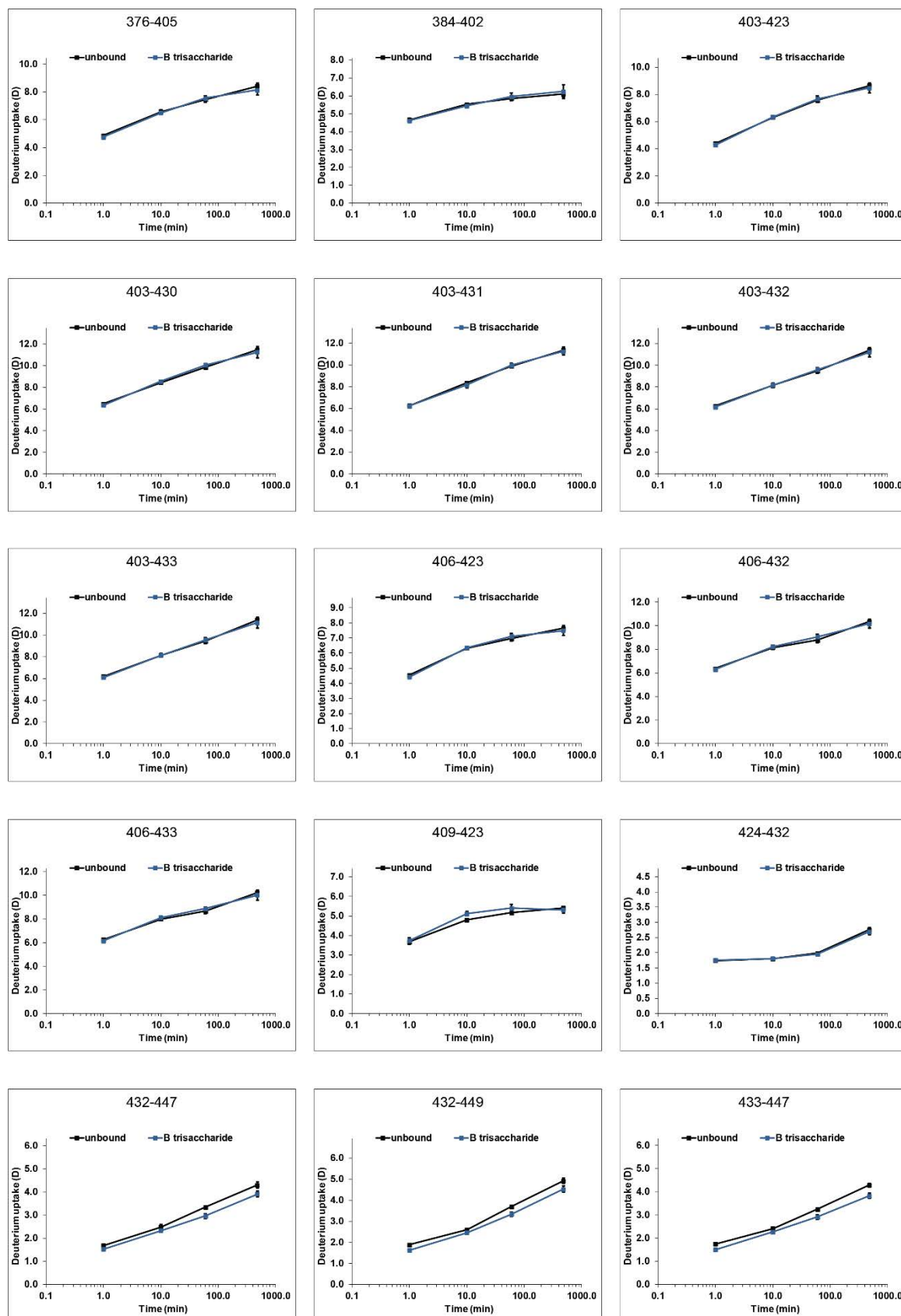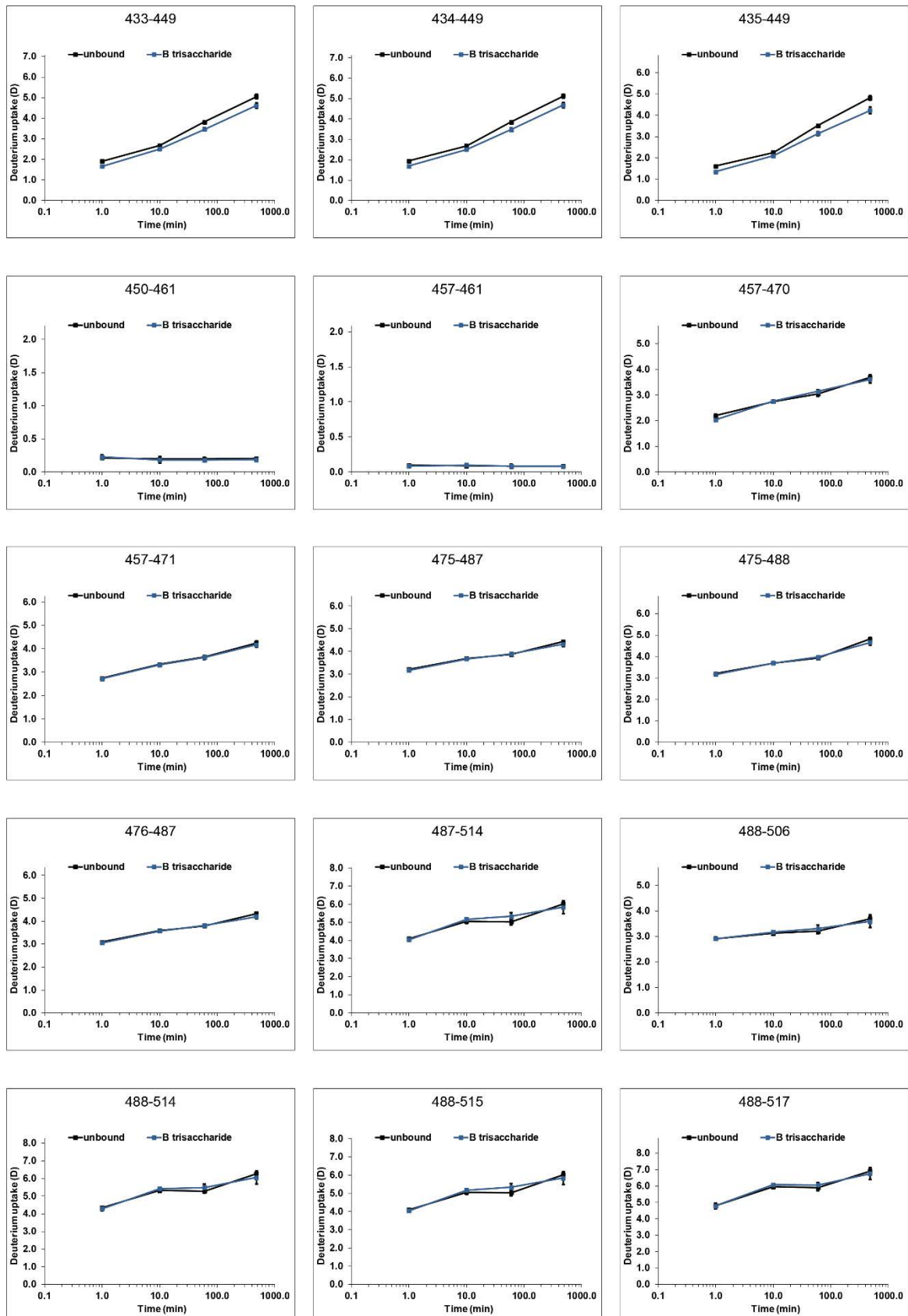
Figure S 34: continued

Figure S 34: continued
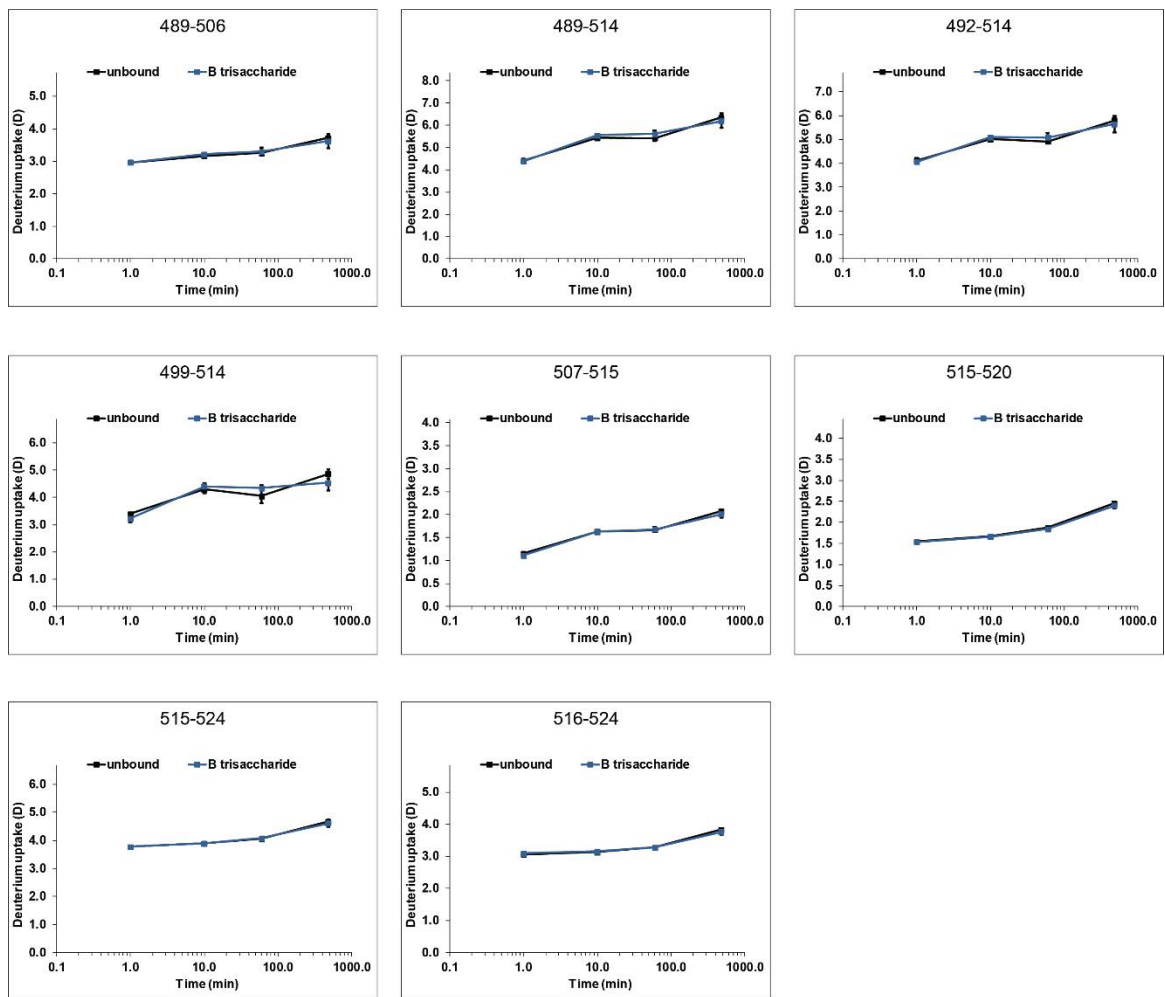
Figure S 34: continued
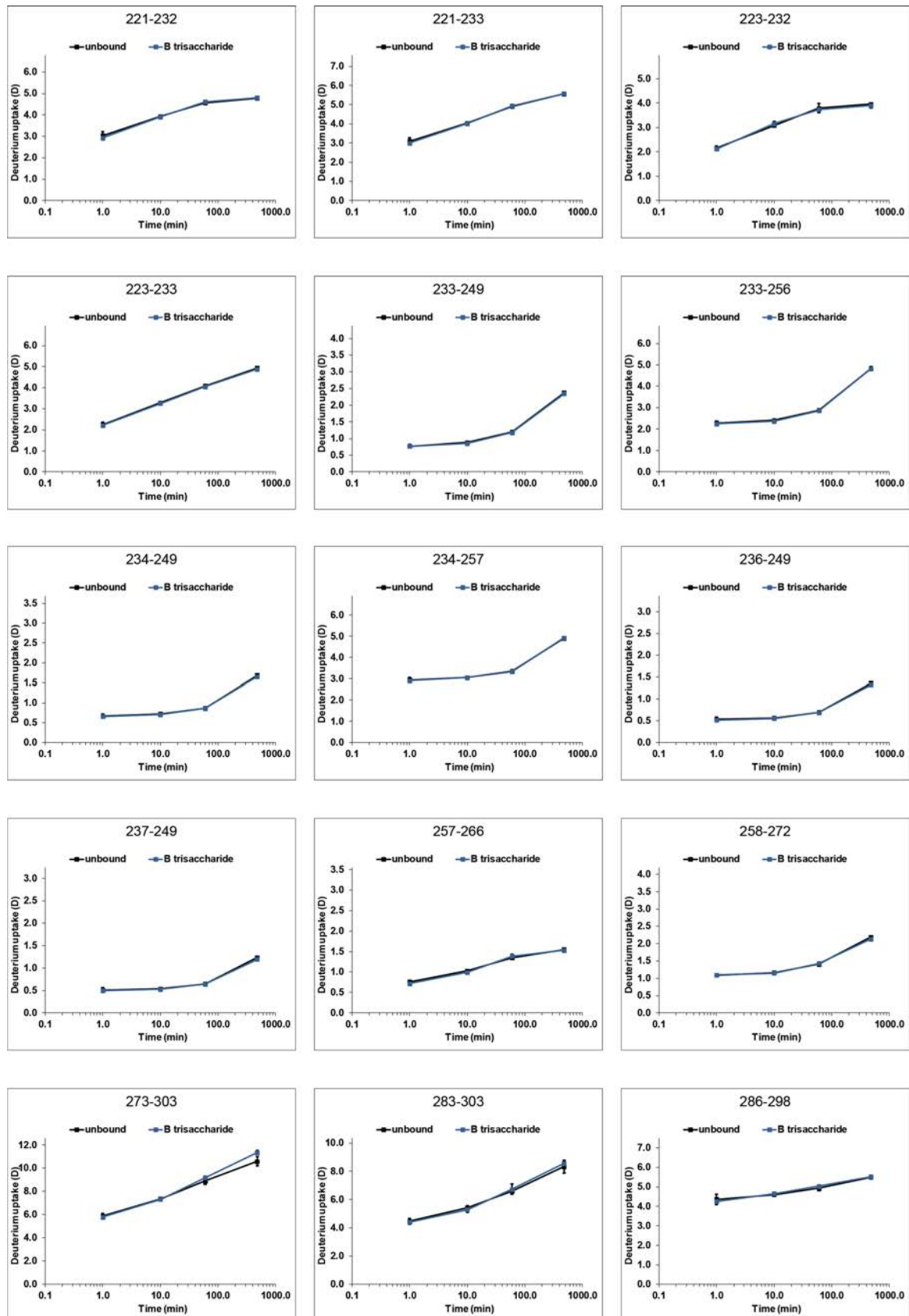
Figure S 35: GII.4 Saga fully deamidated P dimer with 10 mM B trisaccharide
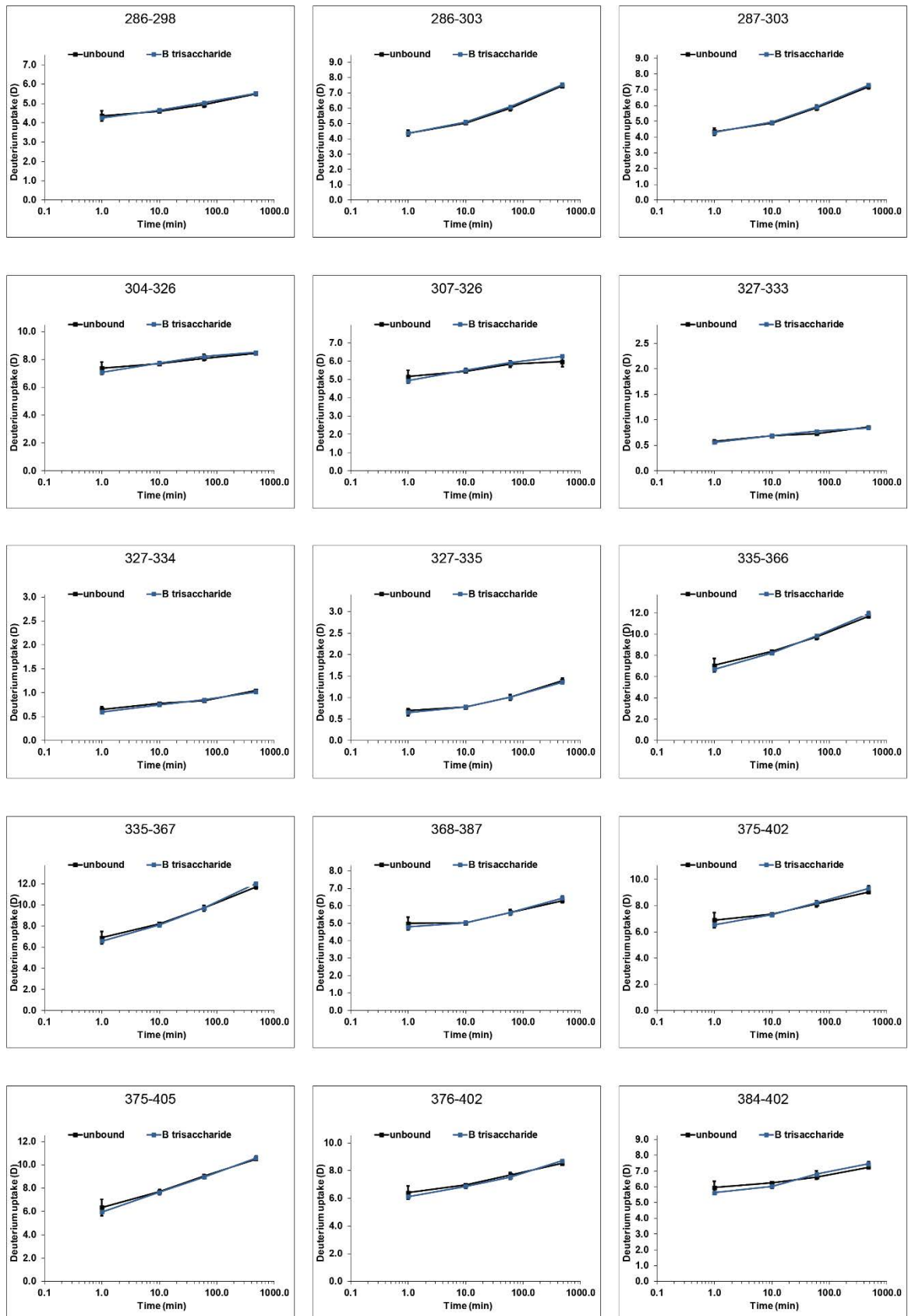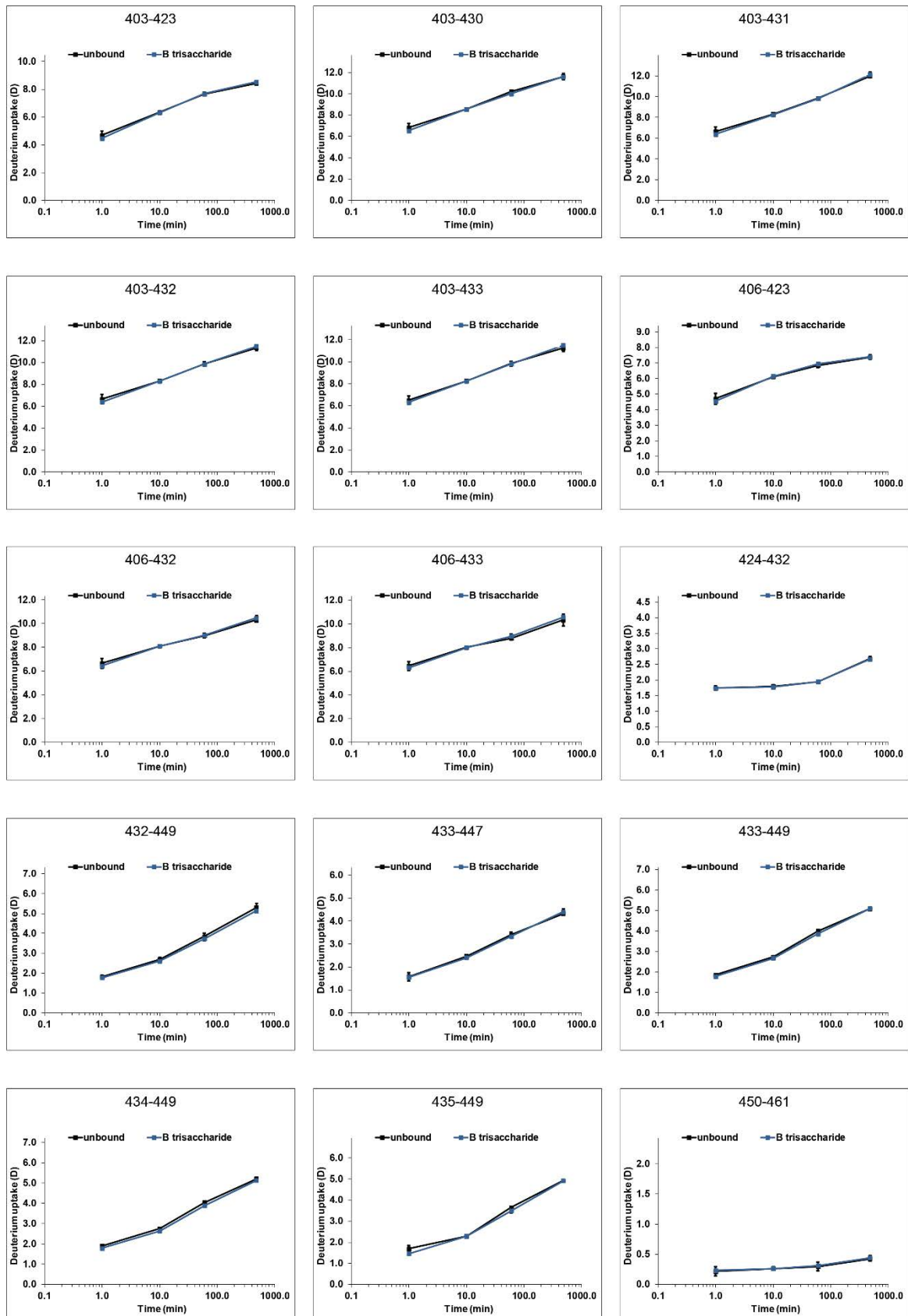
Figure S 35: continued
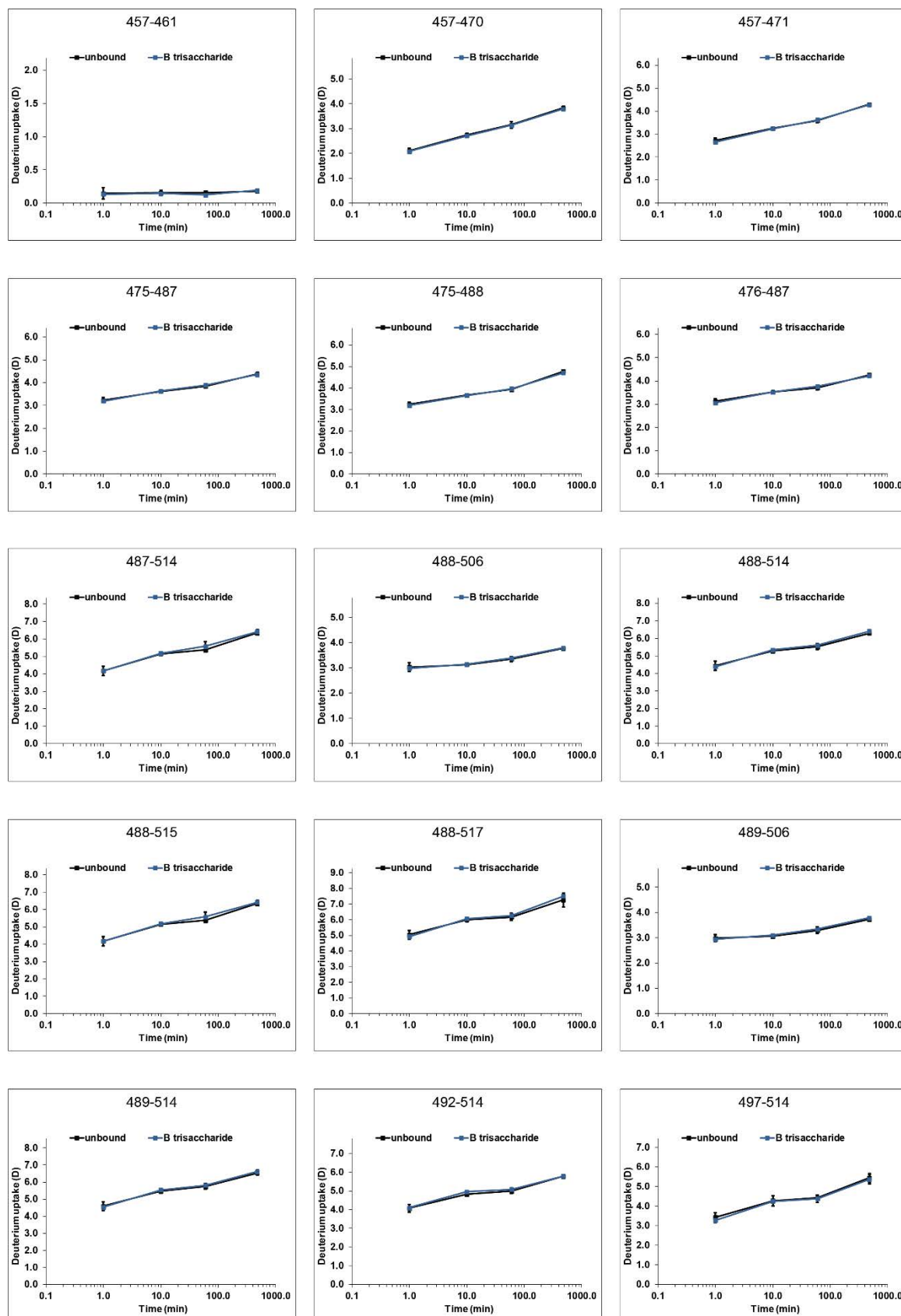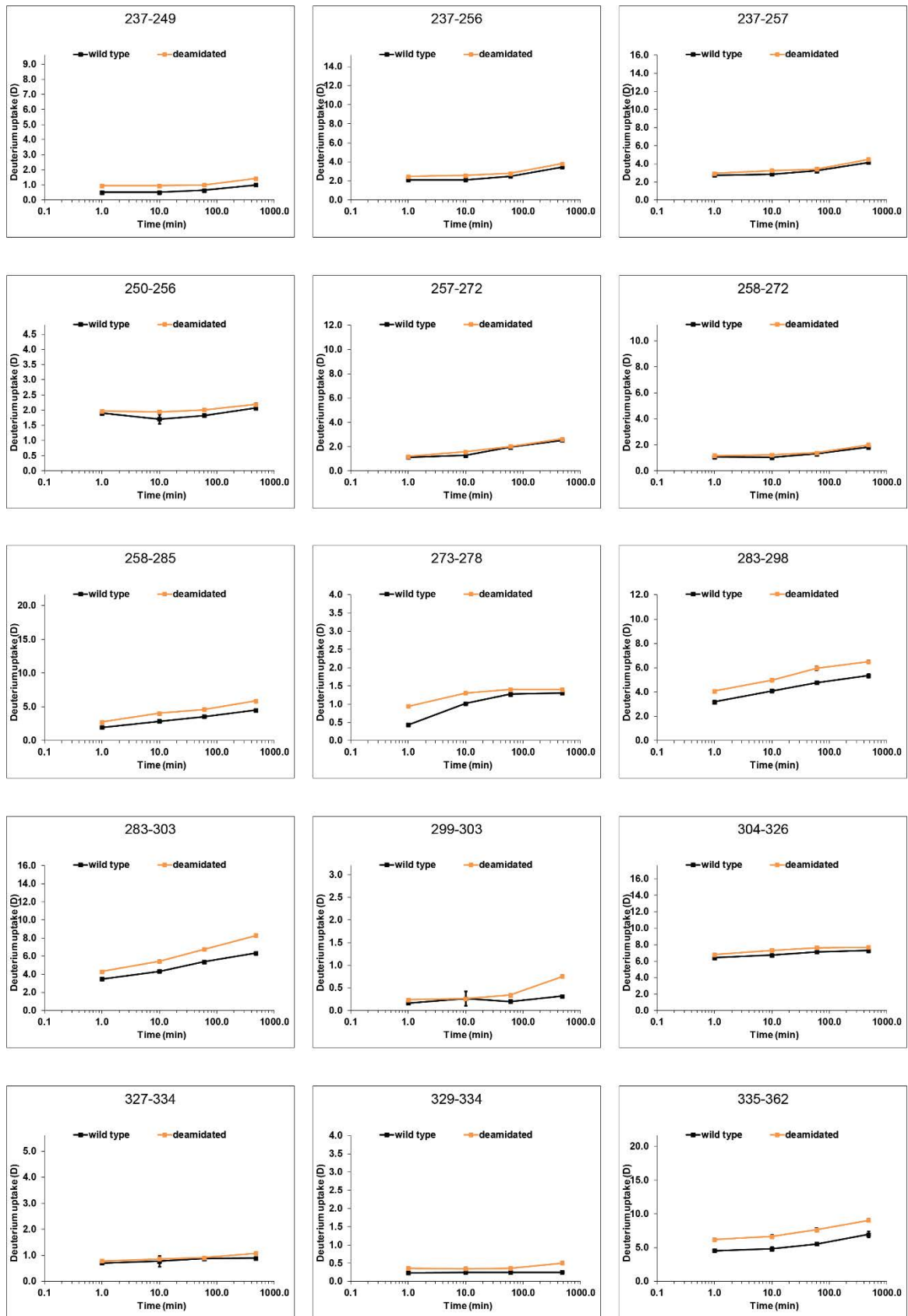
Figure S 35: continued
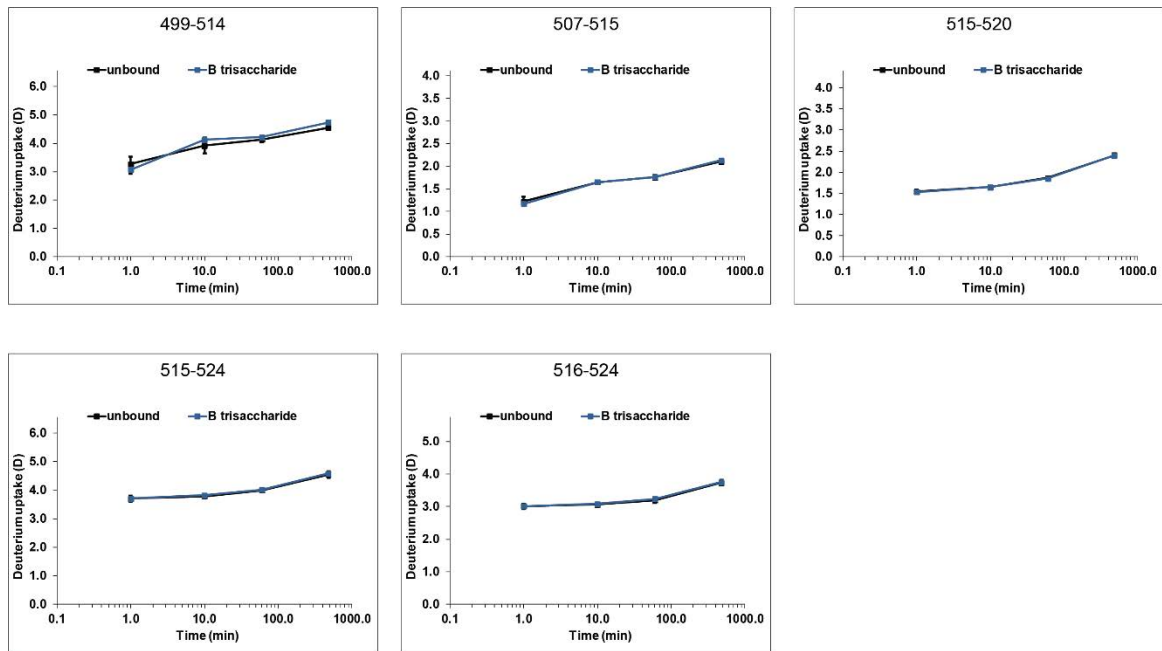
Figure S 35: continued

Figure S 35: continued
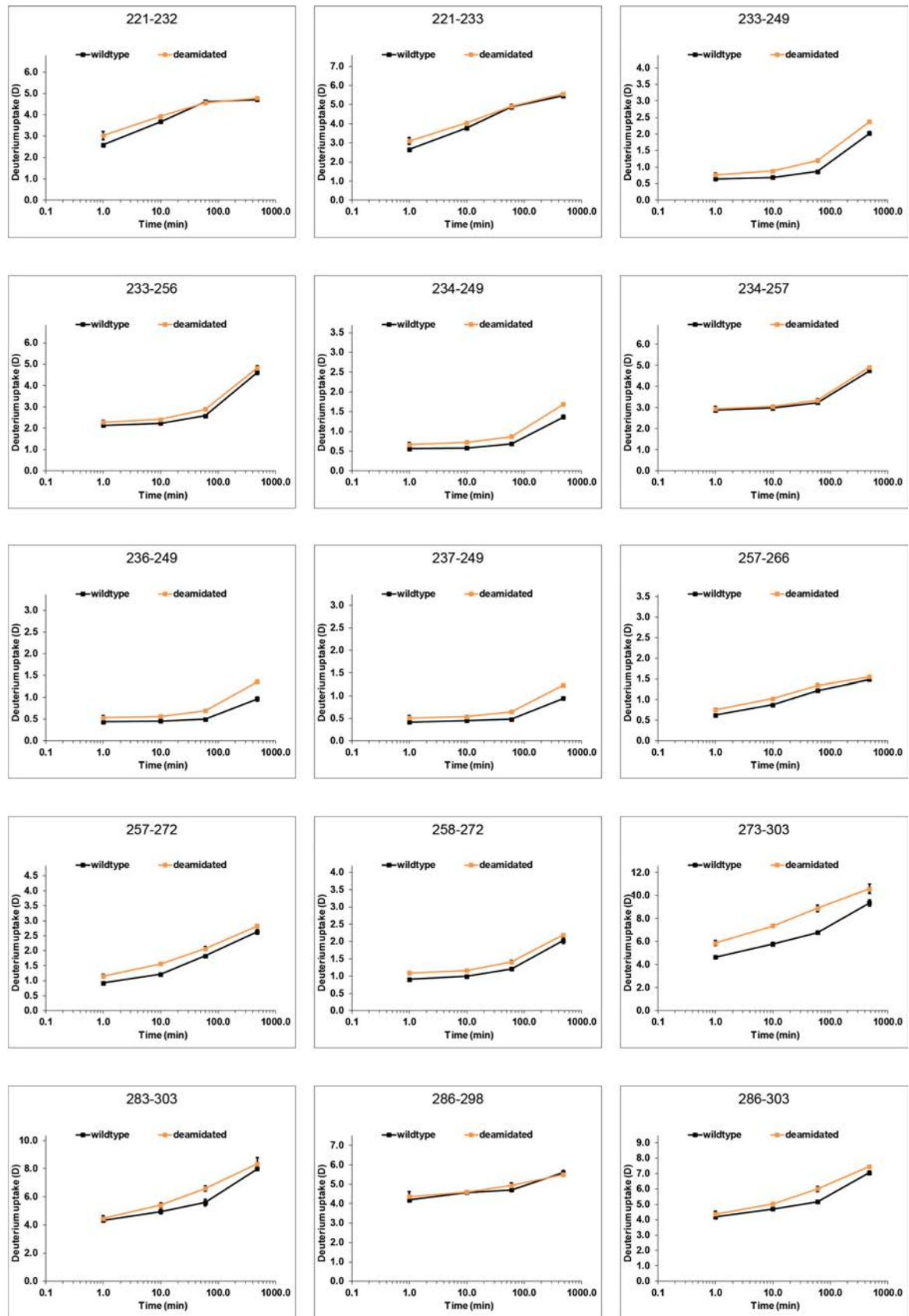
Figure S 36: GII.4 Saga wild type versus fully deamidated P dimer
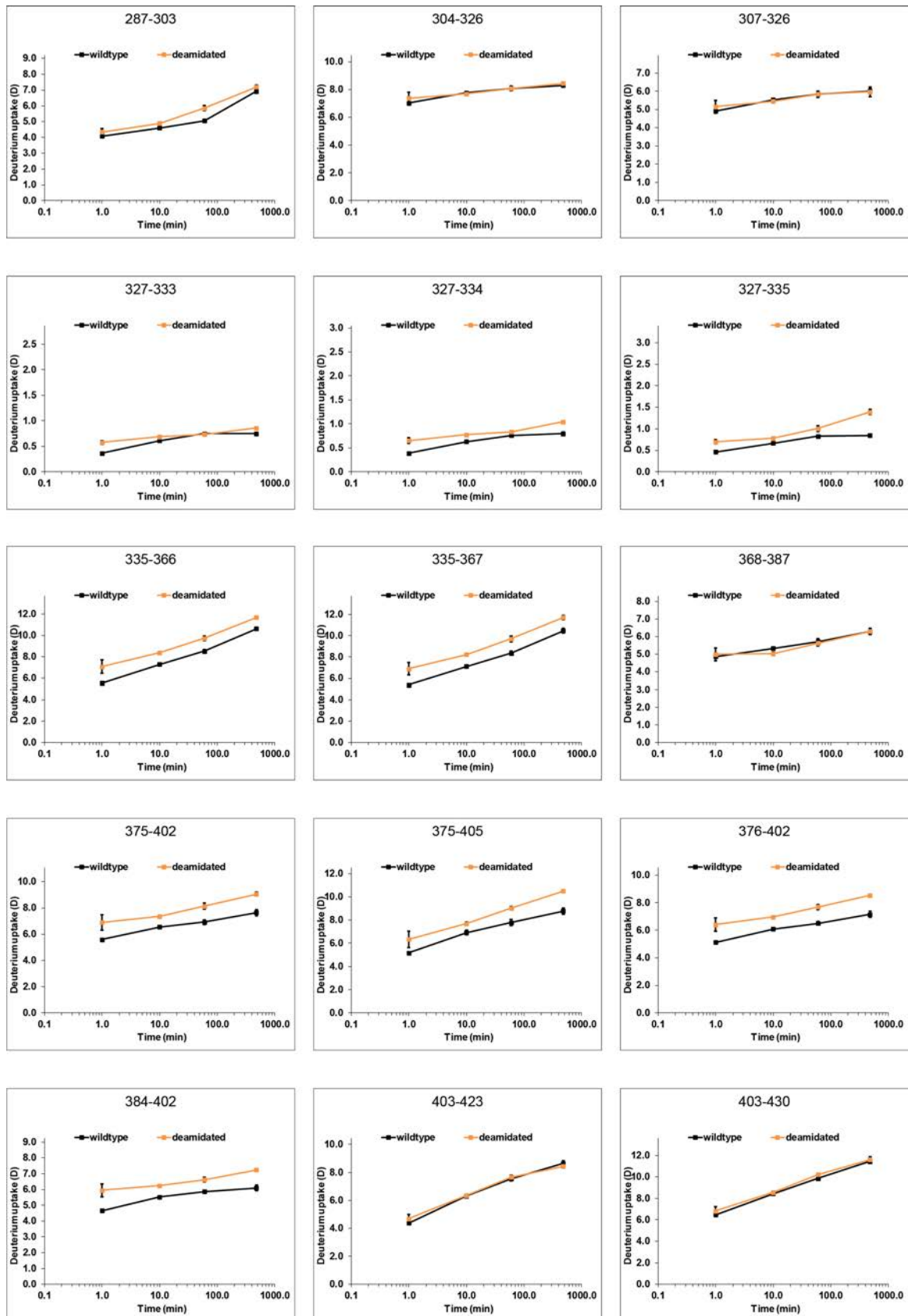
Figure S 36: continued
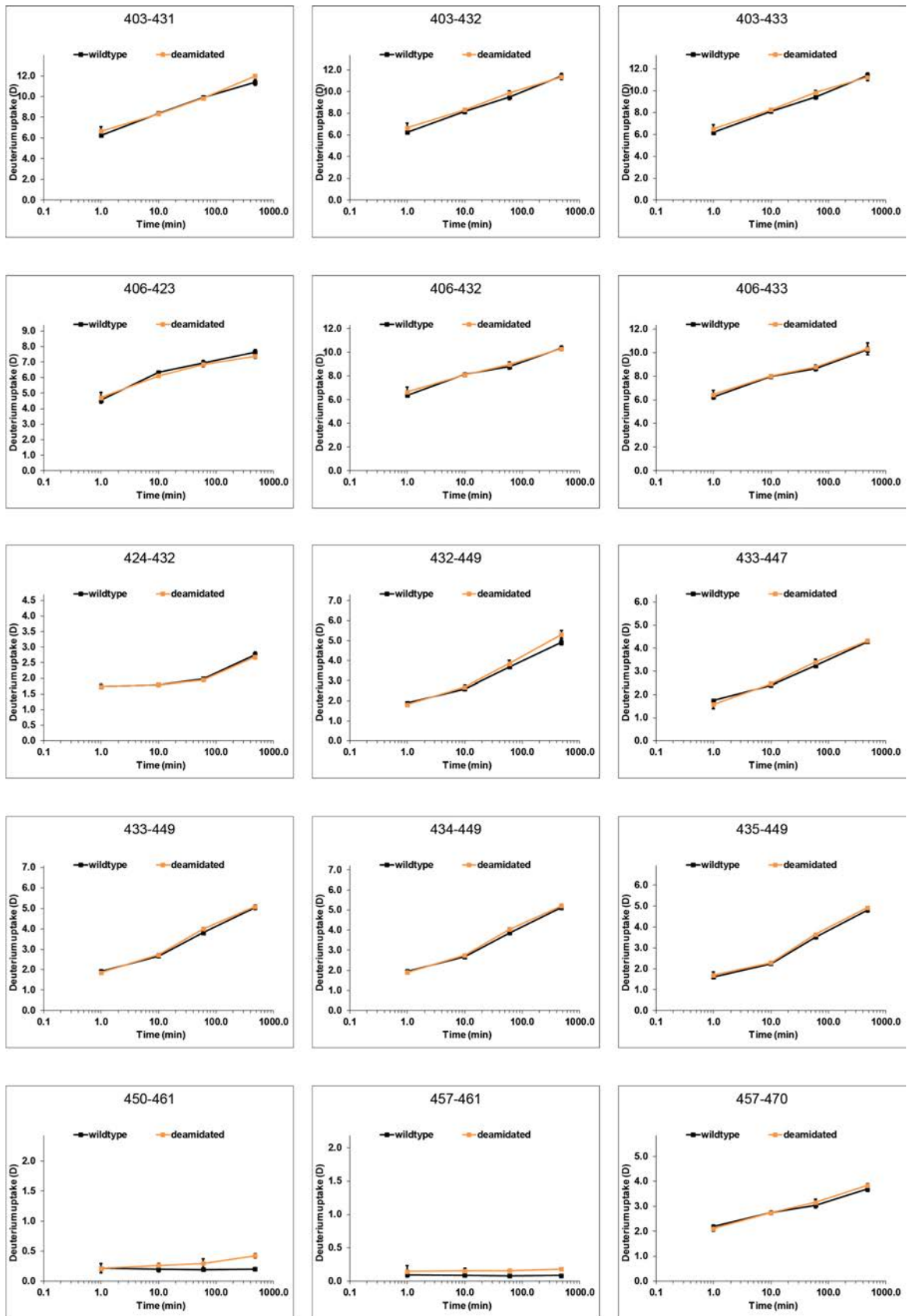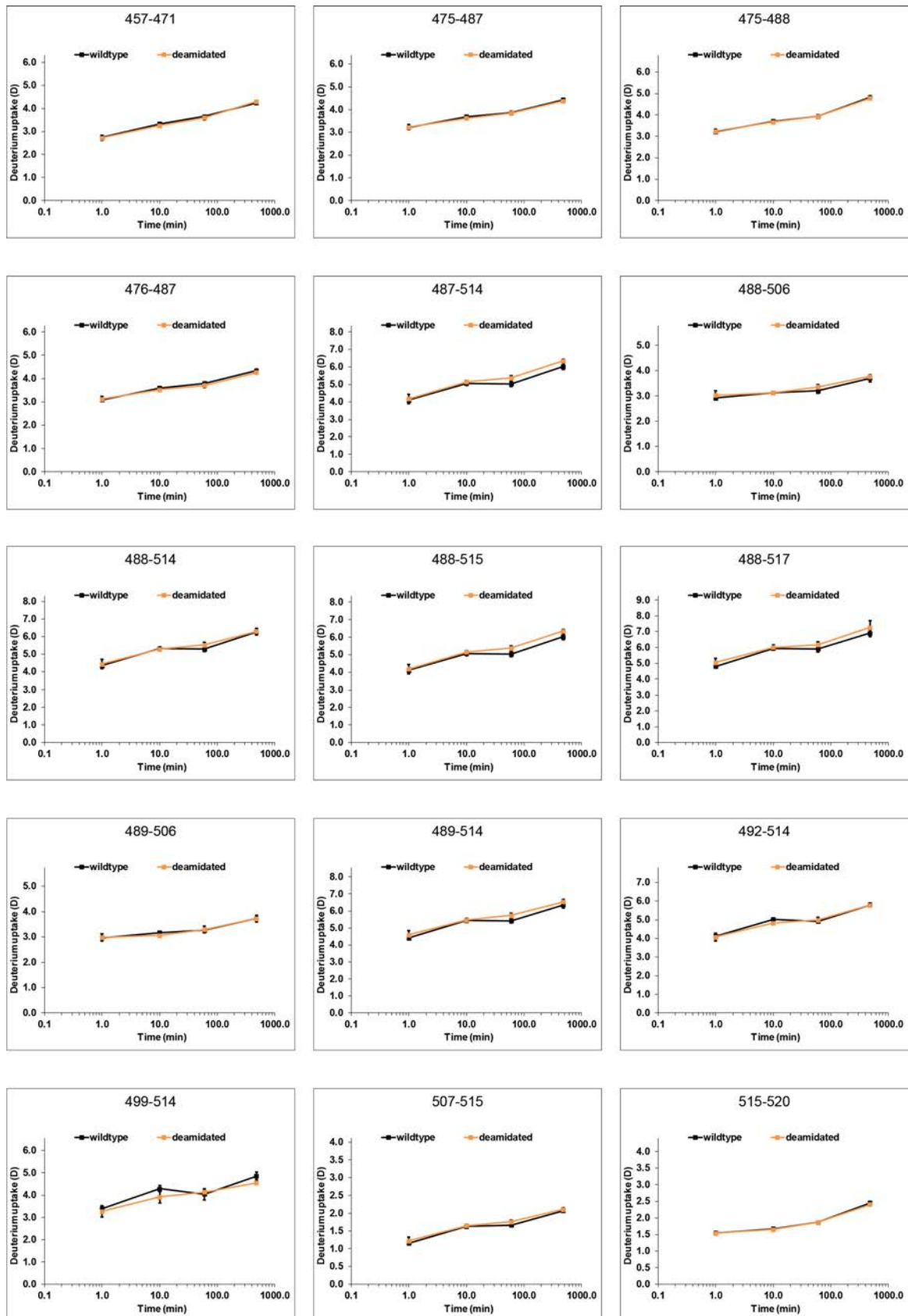
Figure S 36: continued

Figure S 36: continued
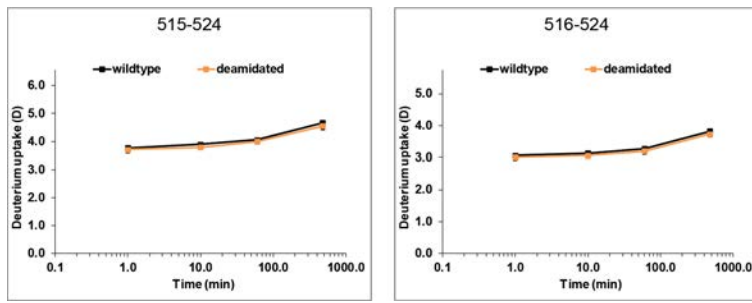
Figure S 36: continued
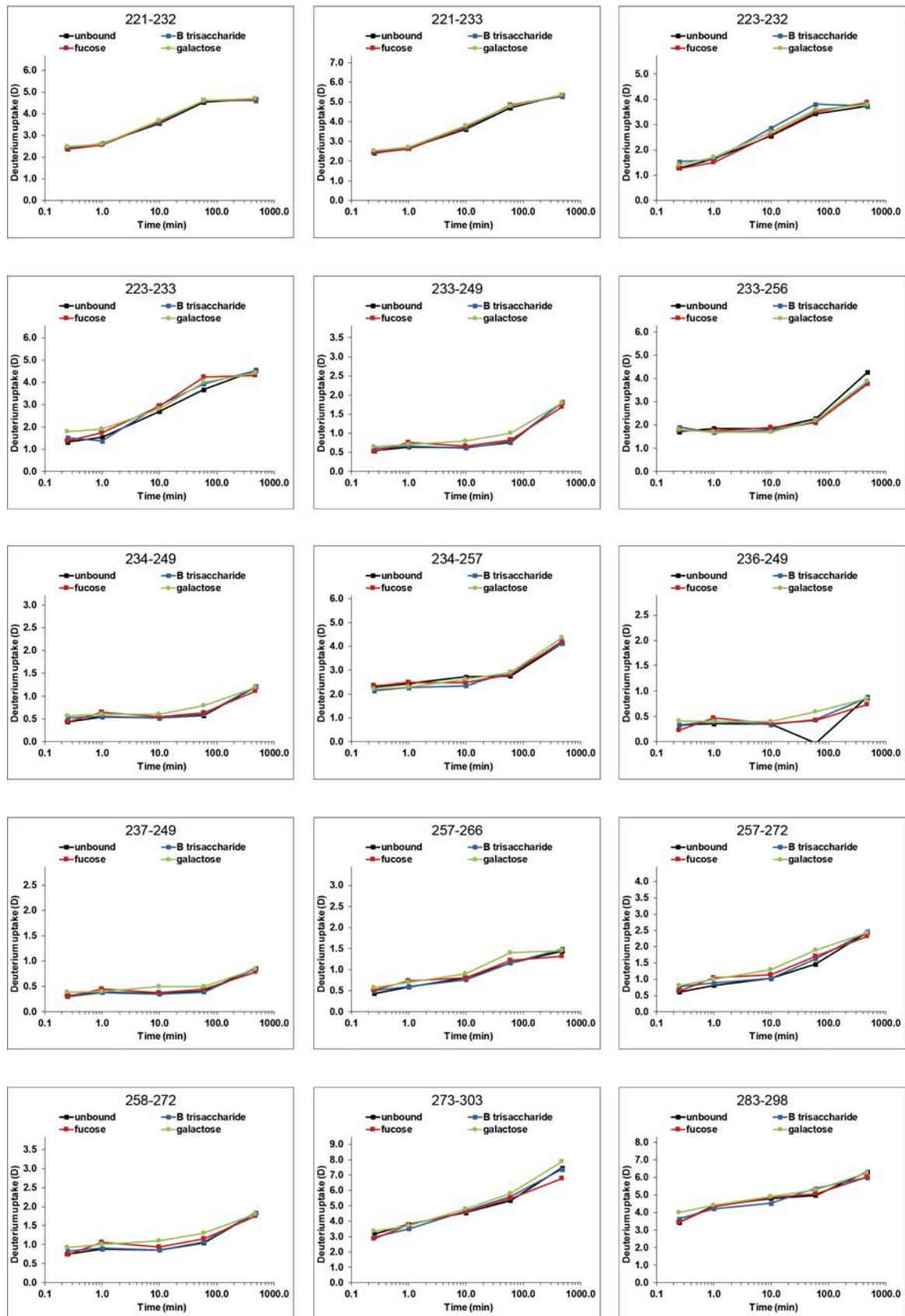
Figure S 37: GII.4 Saga wild type P domain with with three glycan ligands:10 mM B trisaccharide, 100 mM fucose and 100 mM galactose (negative control).
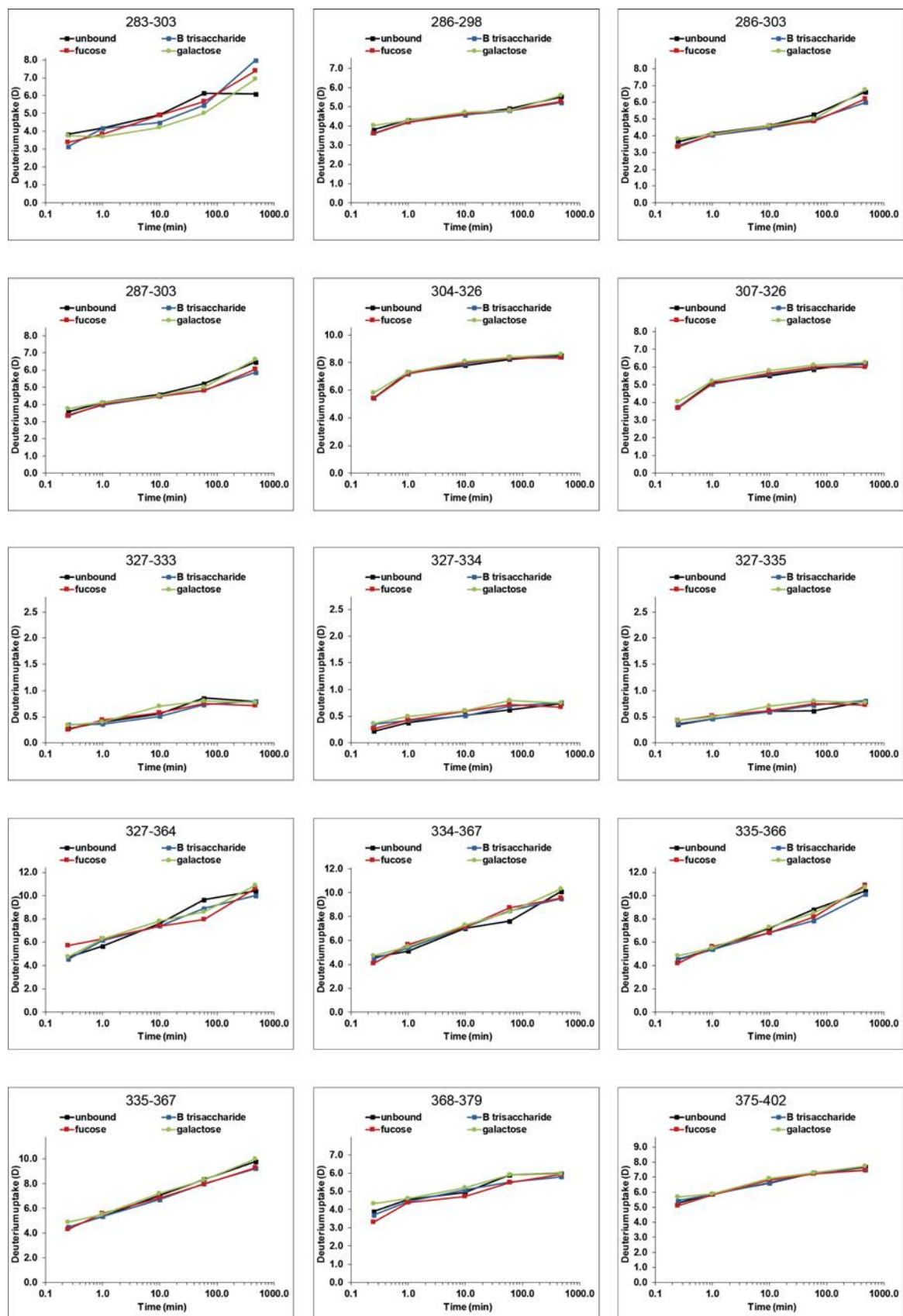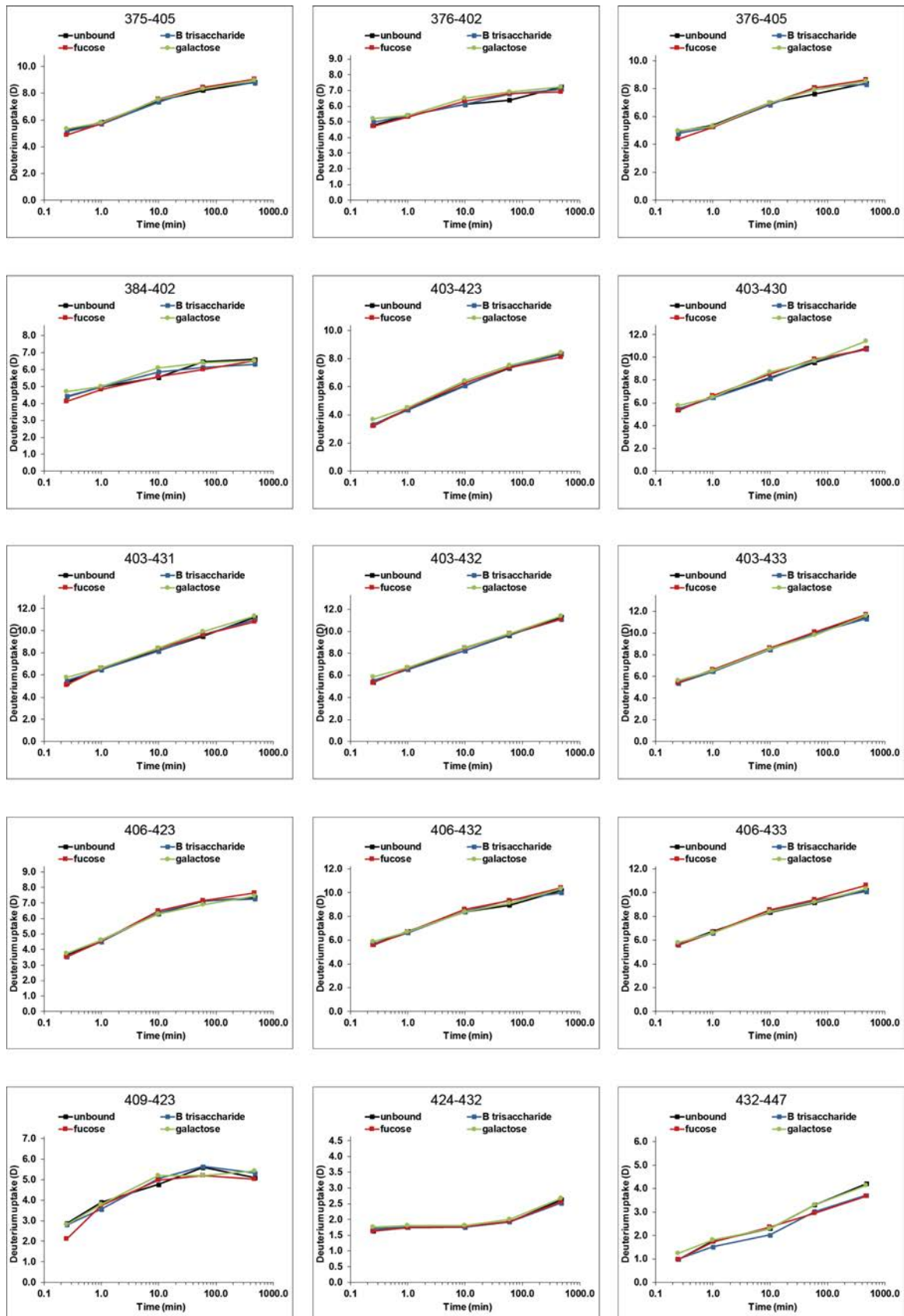
Figure S 37: continued
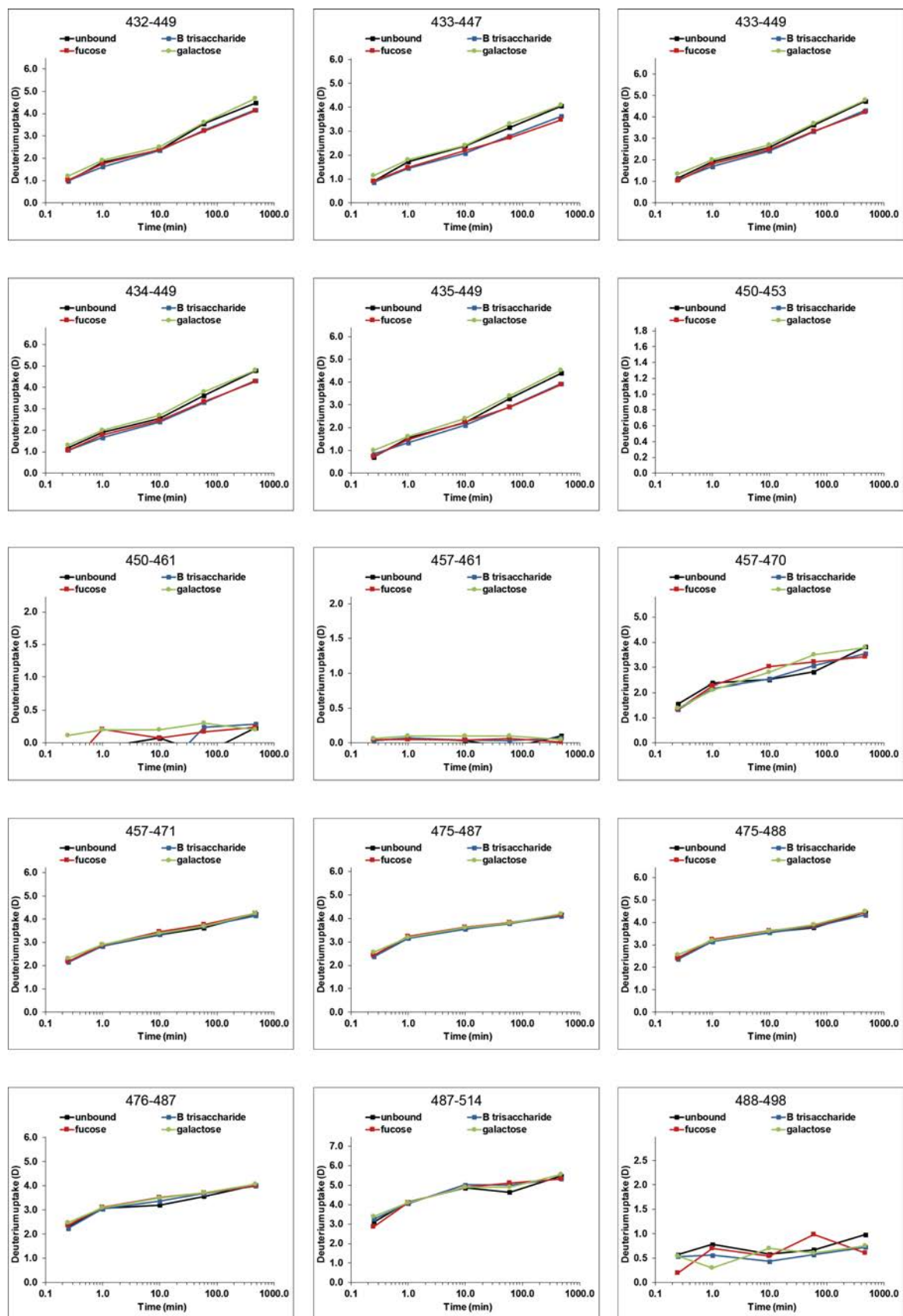
Figure S 37: continued
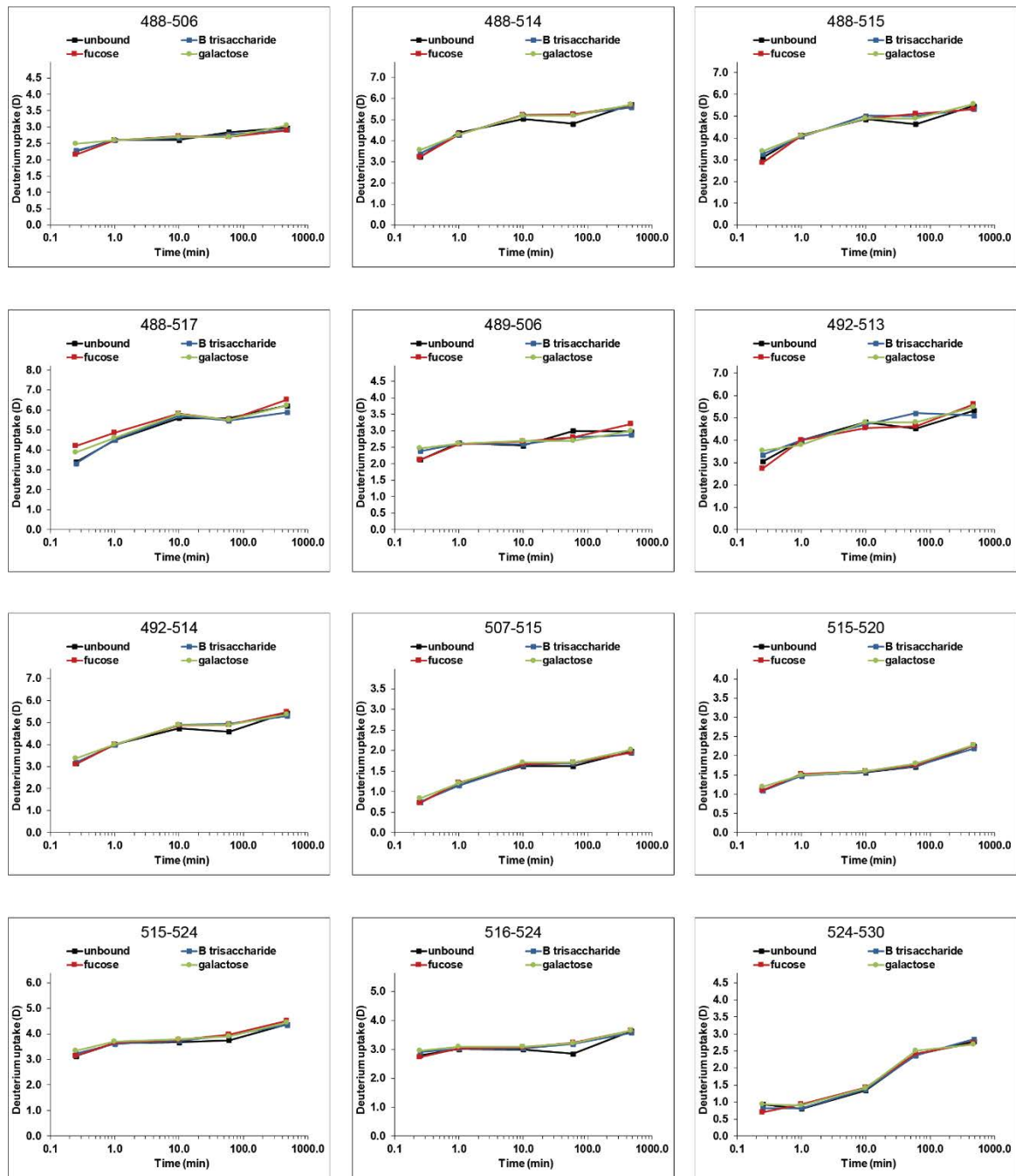
Figure S 37: continued

Figure S 37: continued
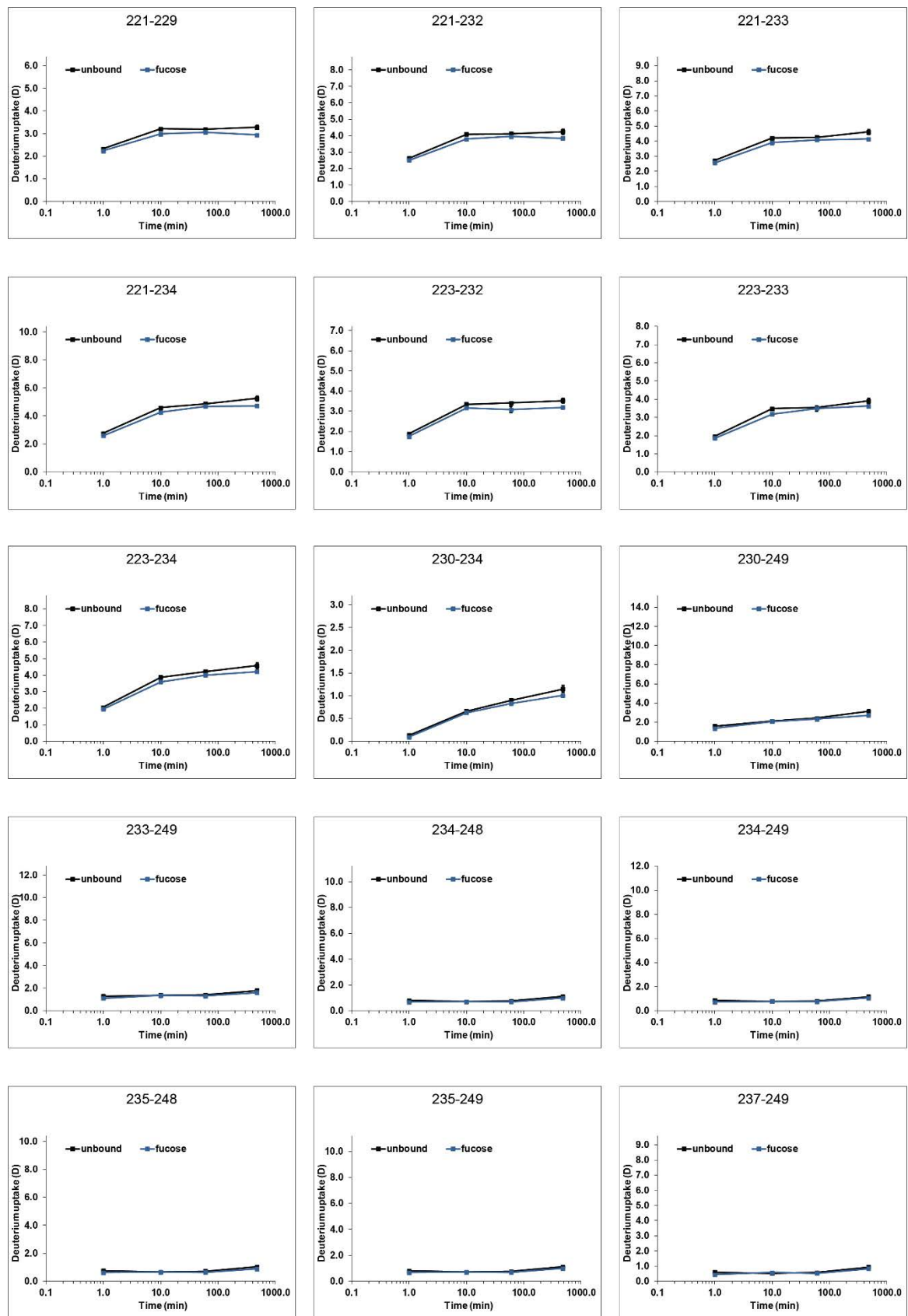
Figure S 38: GII.10 Vietnam P dimer with 100 mM fucose
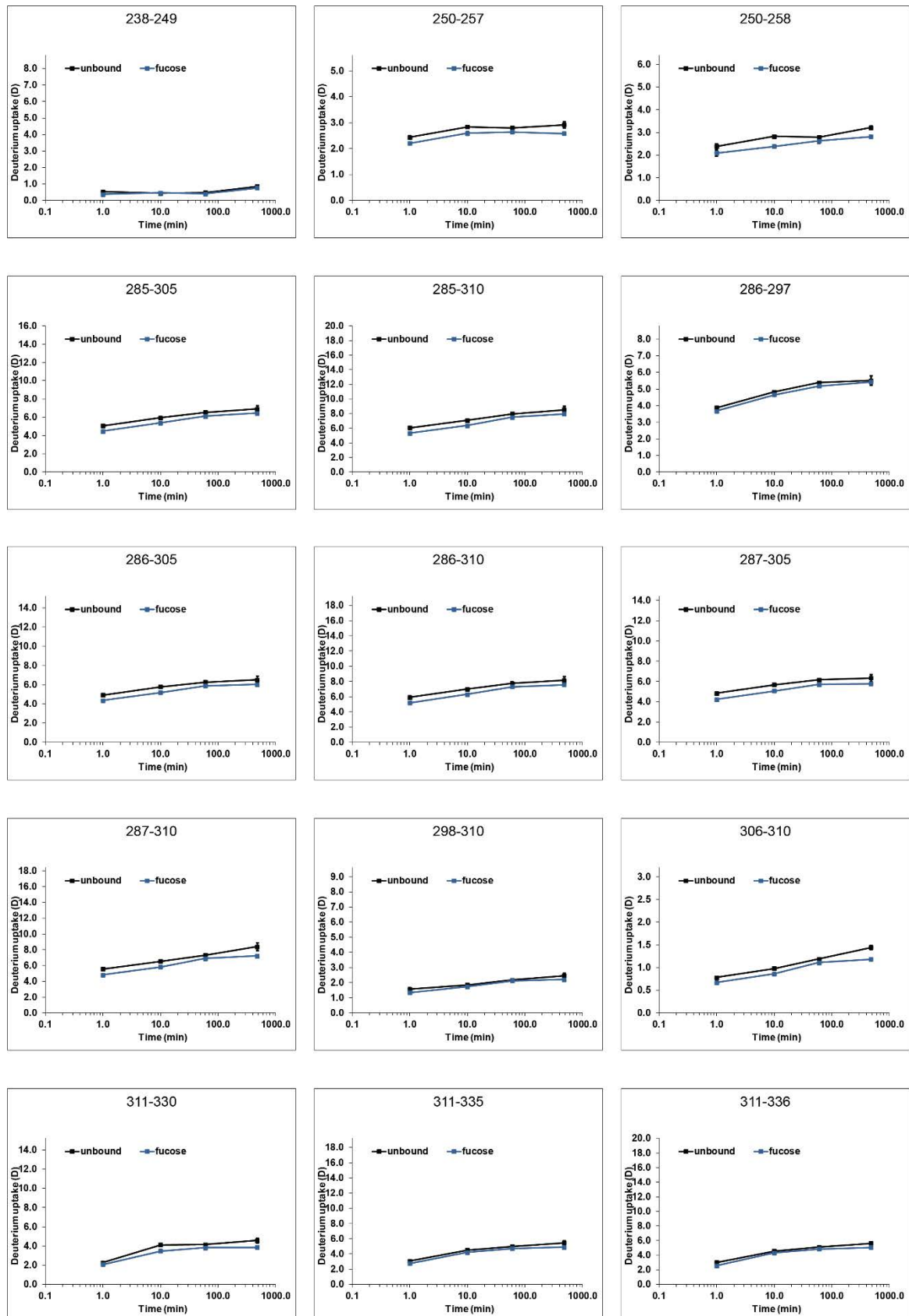
Figure S 38: continued
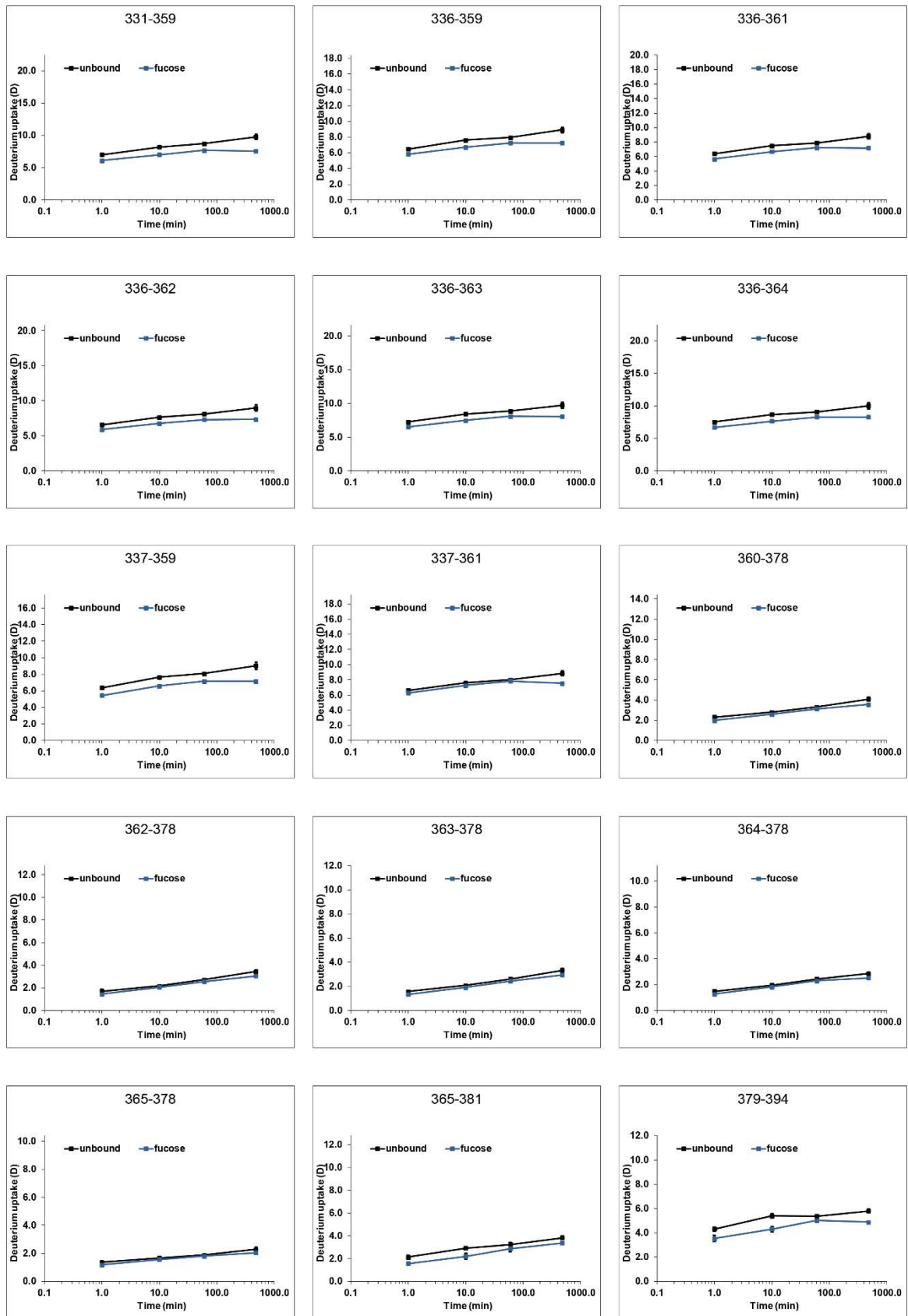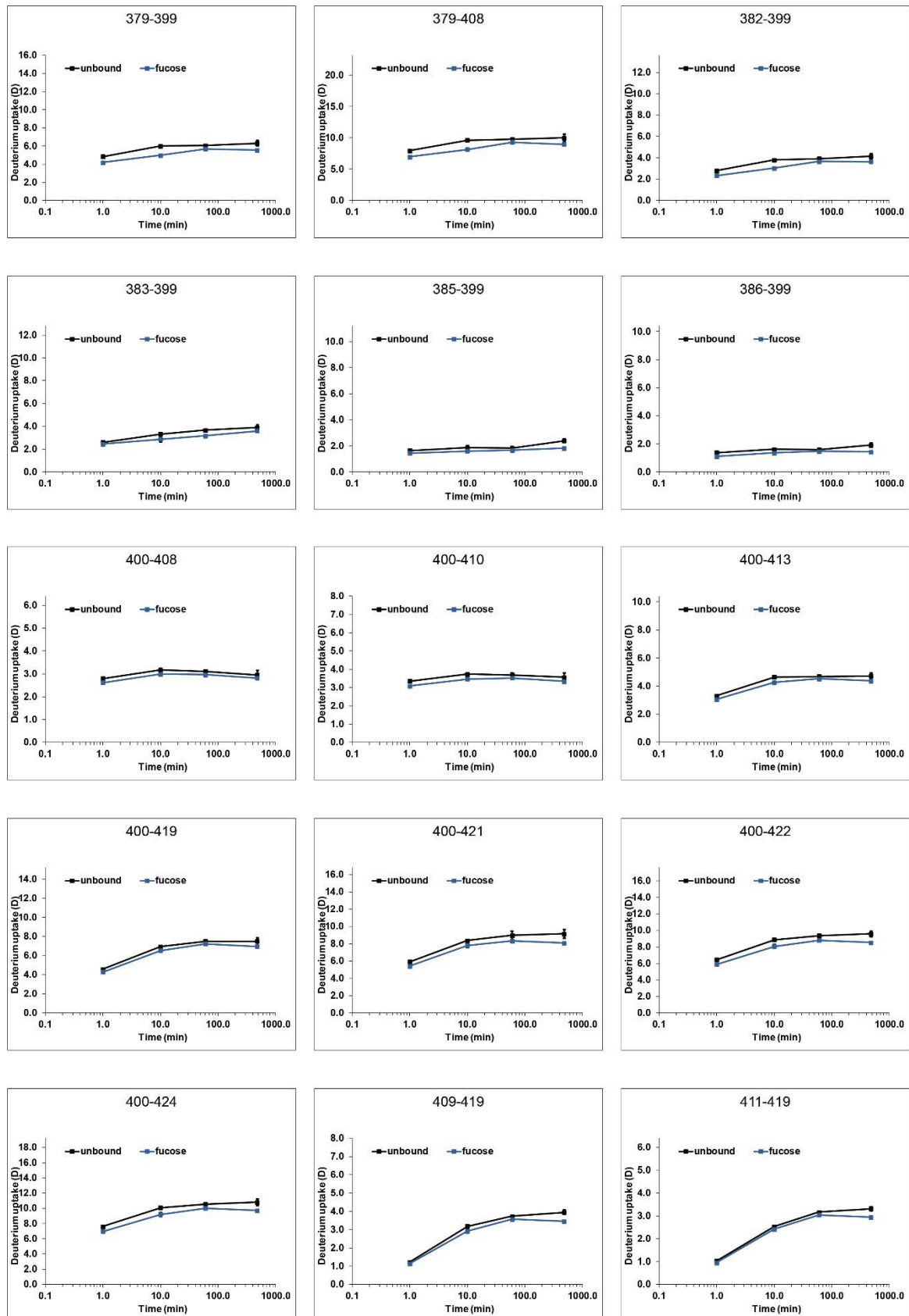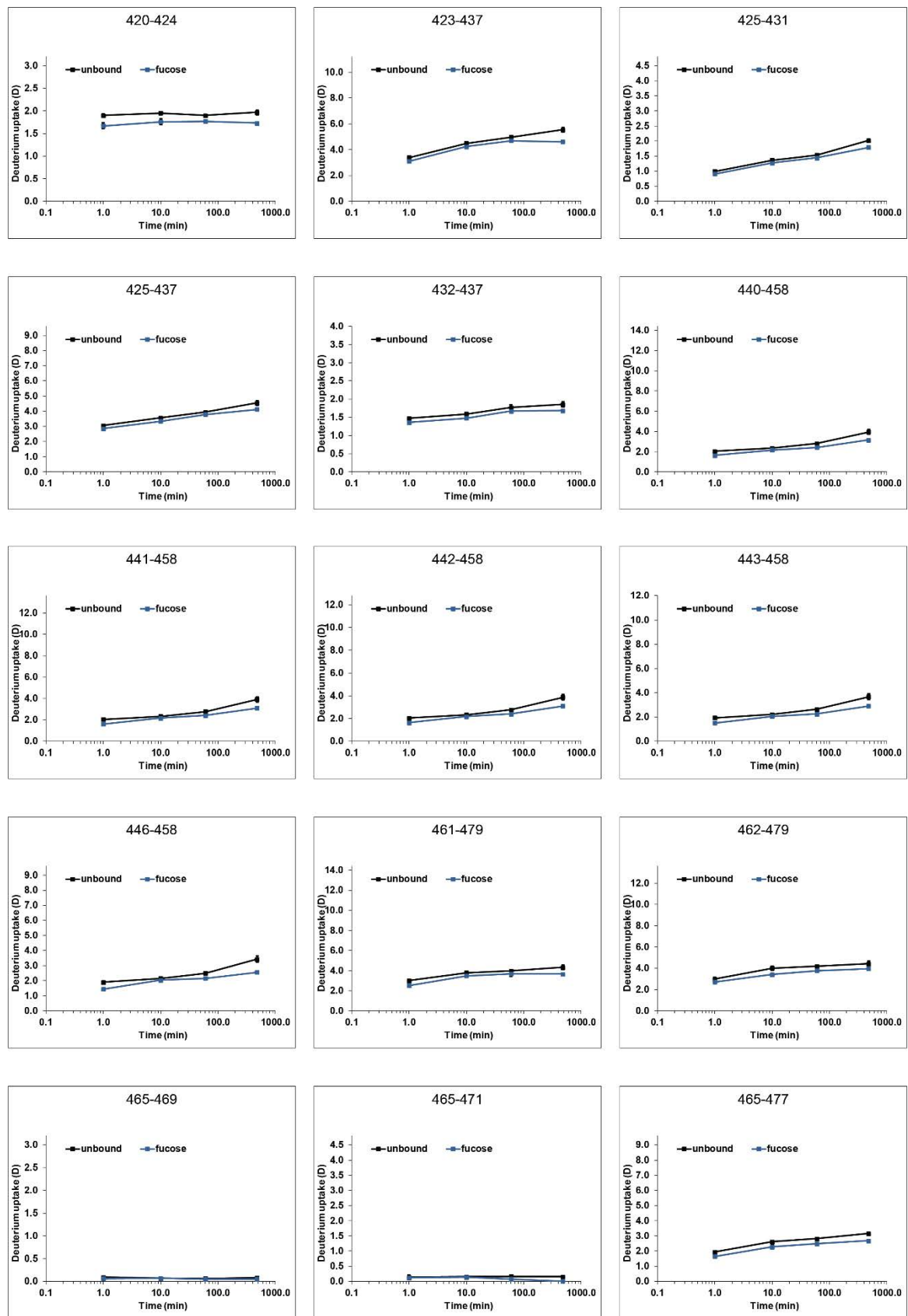
Figure S 38: continued
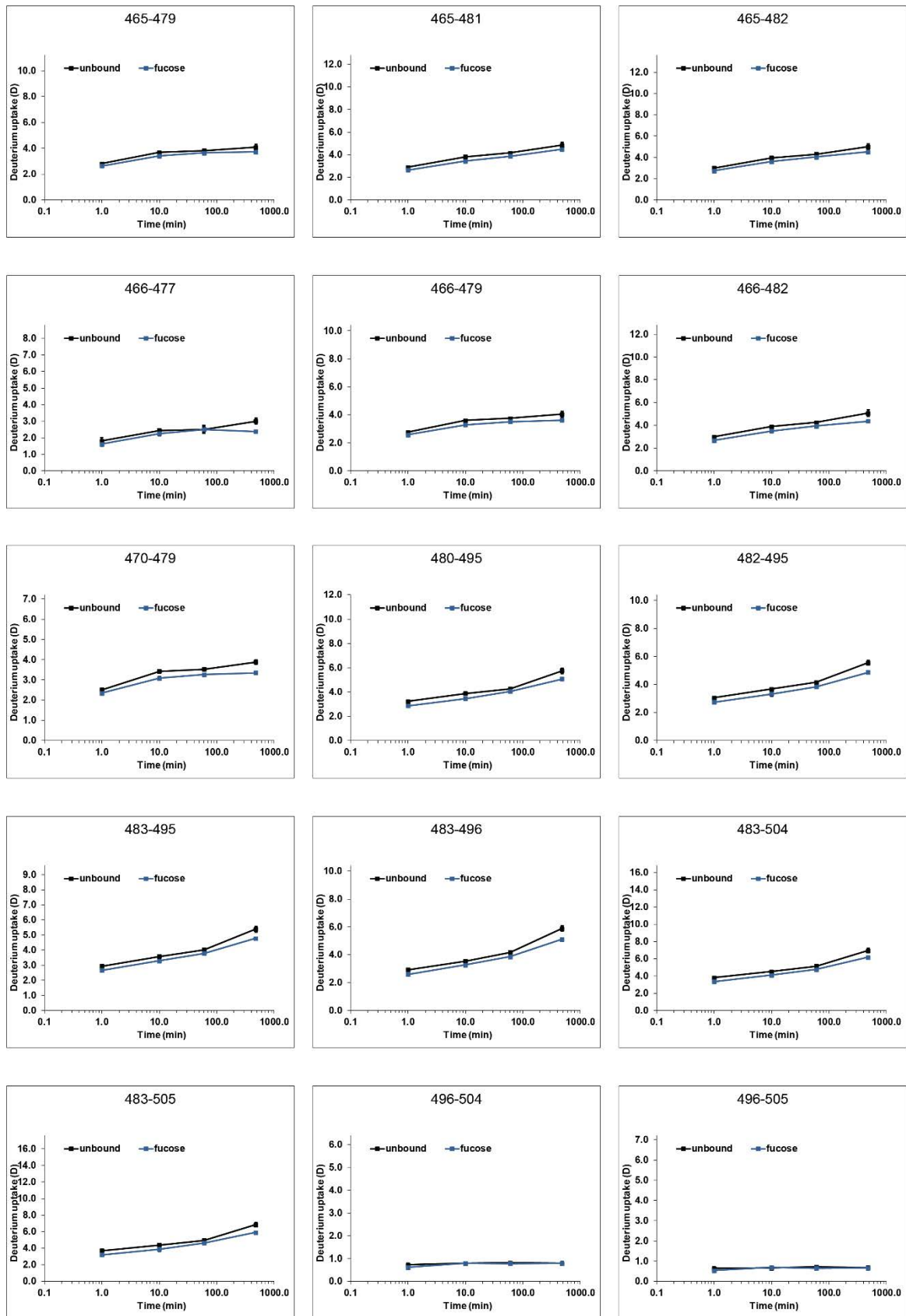
Figure S 38: continued

Figure S 38: continued
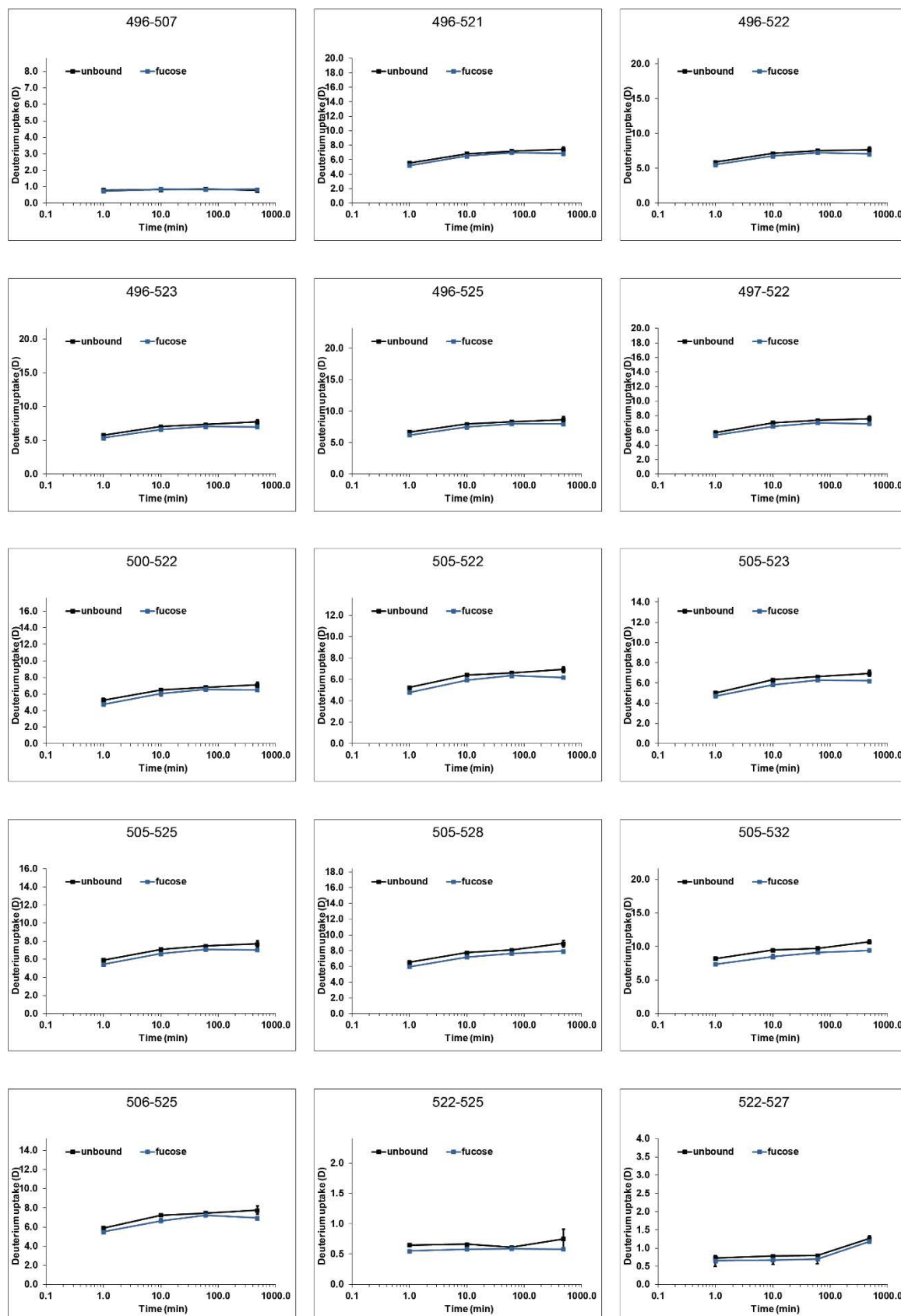
Figure S 38: continued
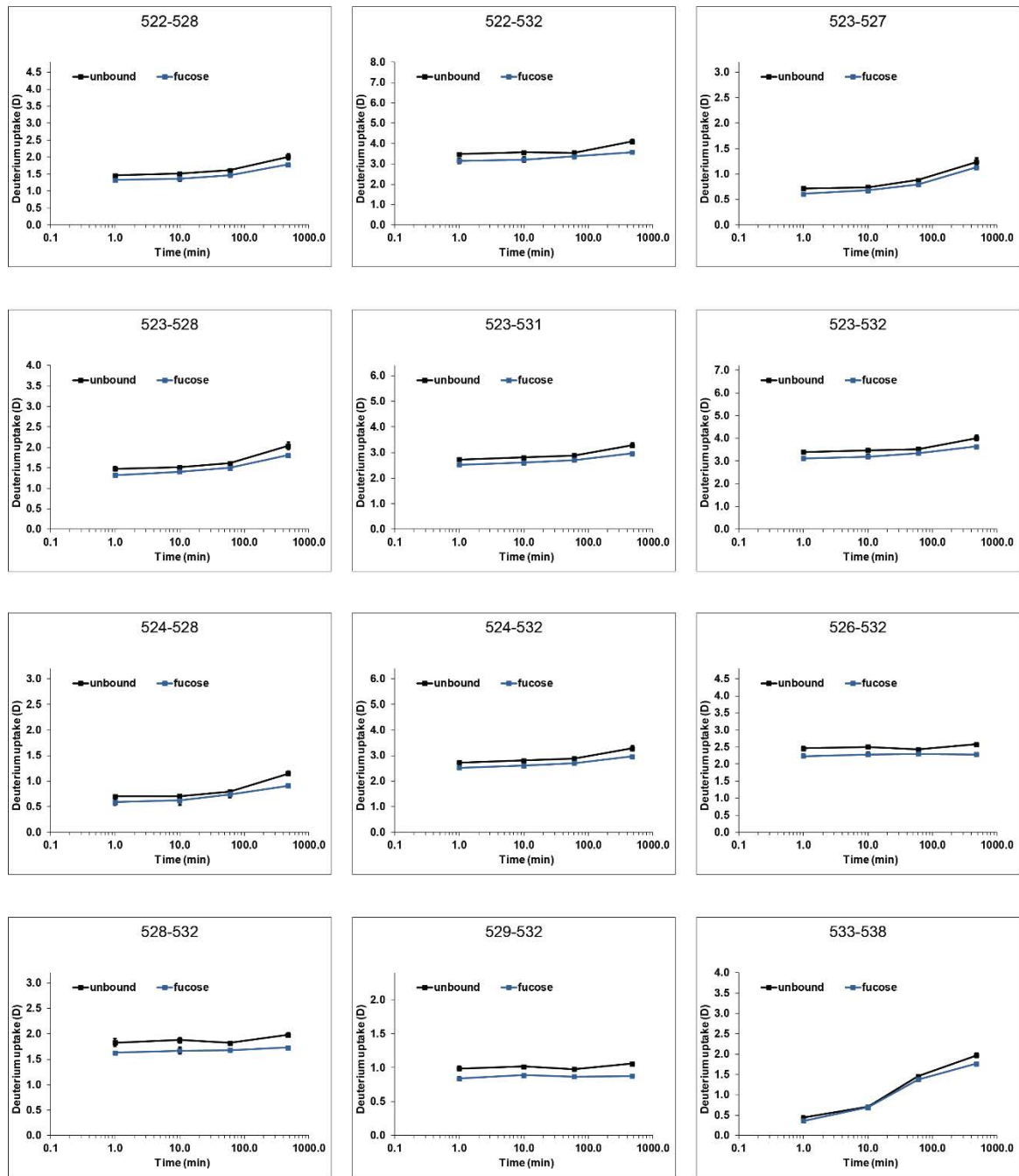
Figure S 38: continued

Figure S 38: continued
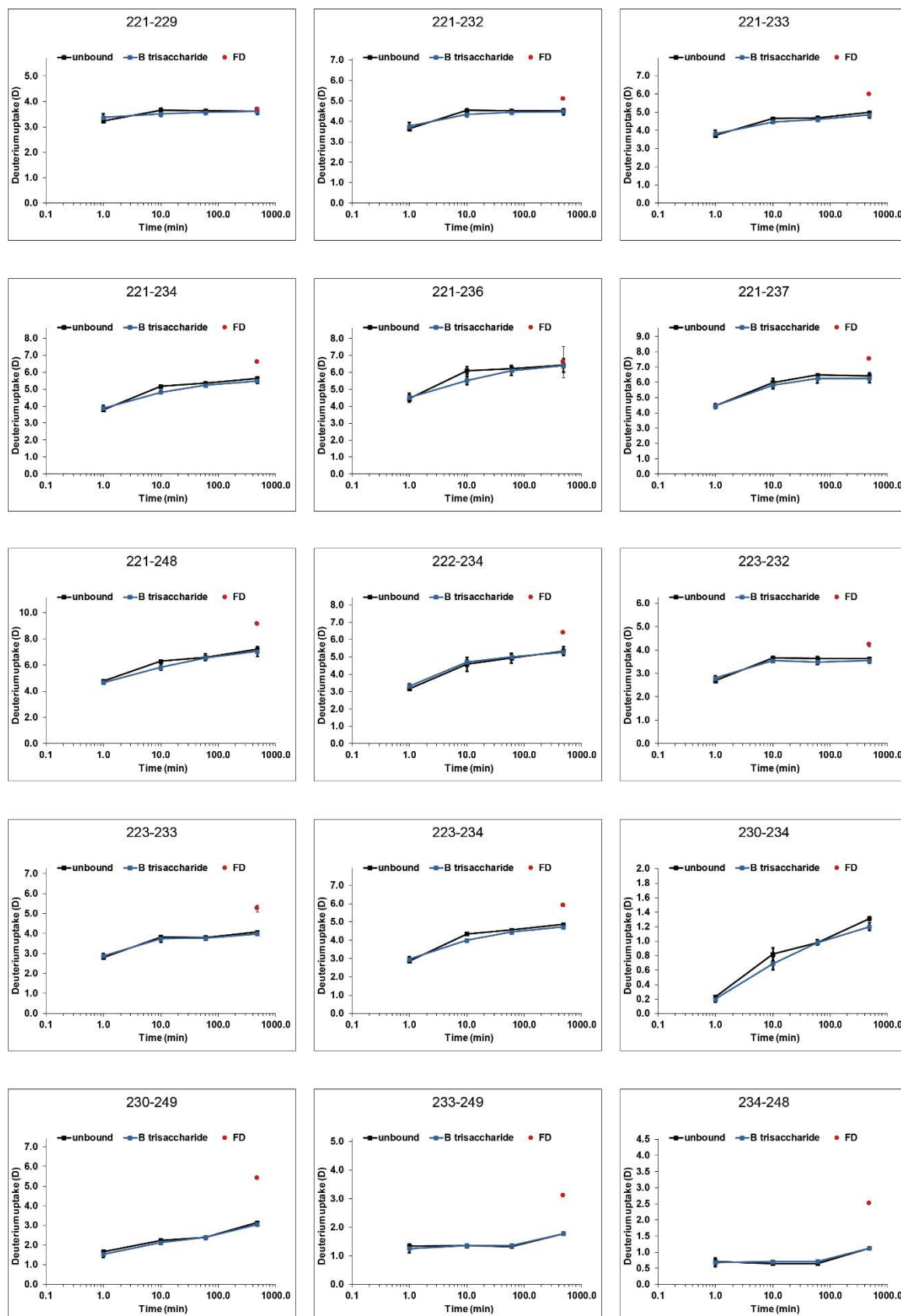
Figure S 39: GII.10 Vietnam P dimer with 10 mM B trisaccharide
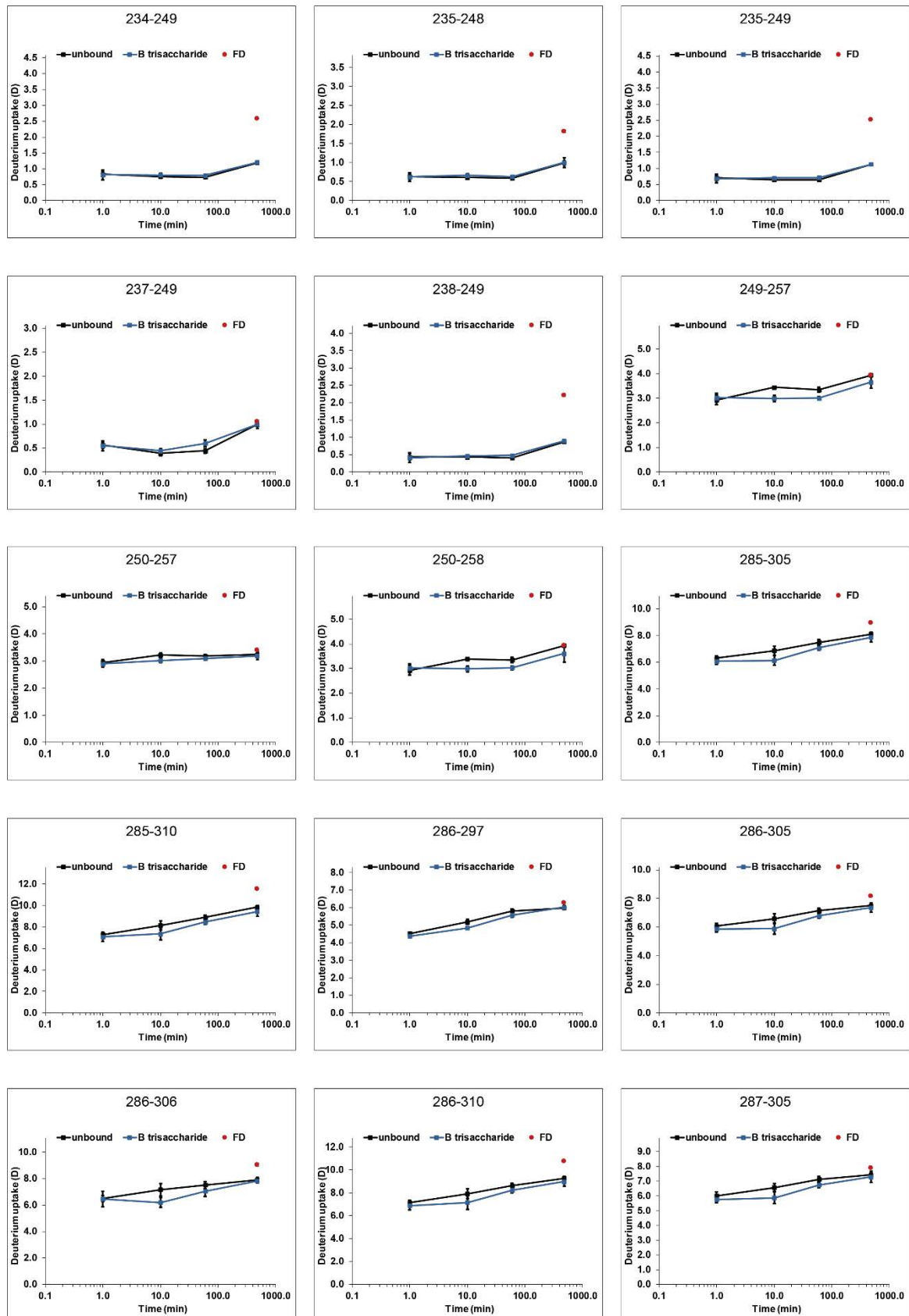
Figure S 39: continued
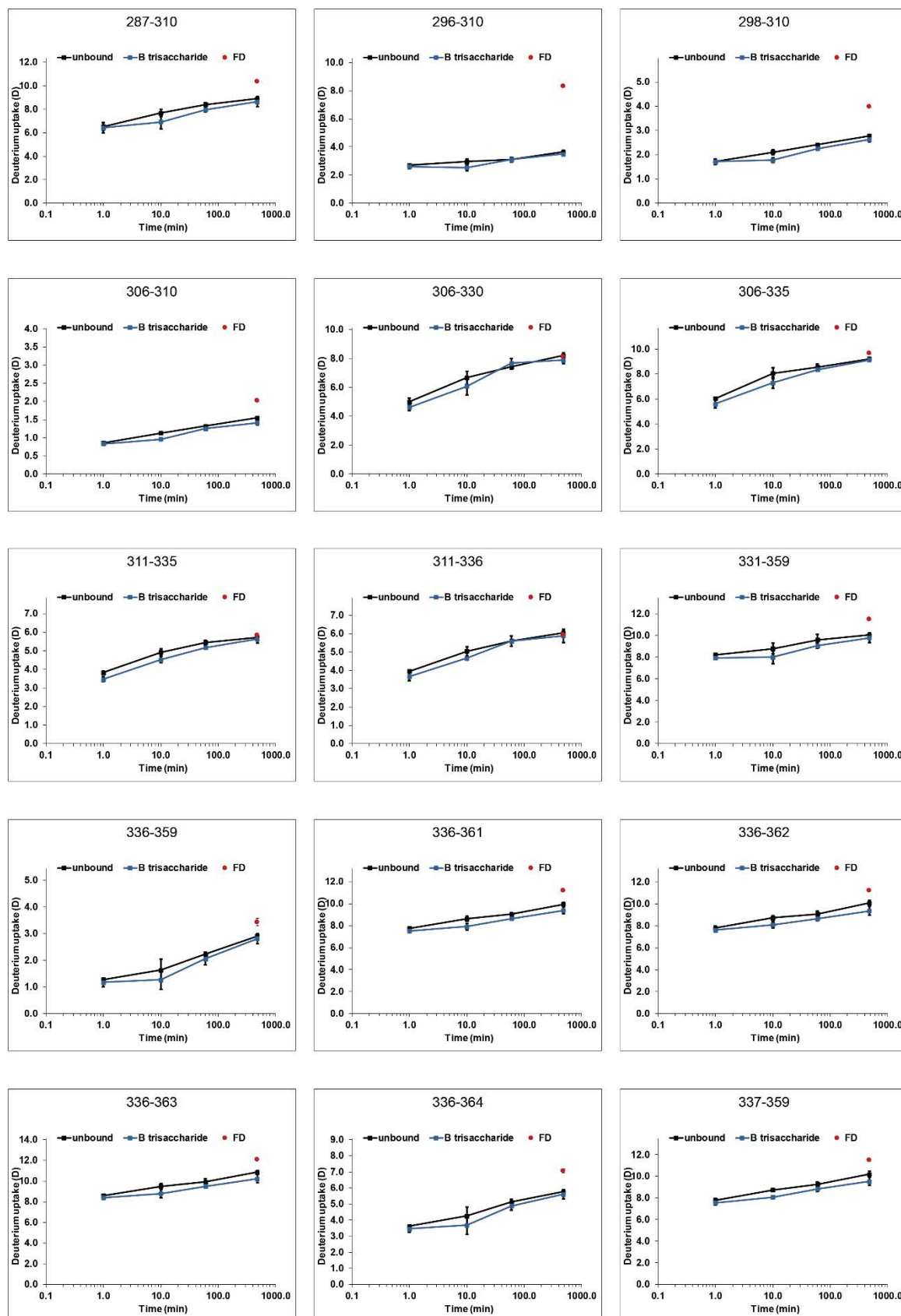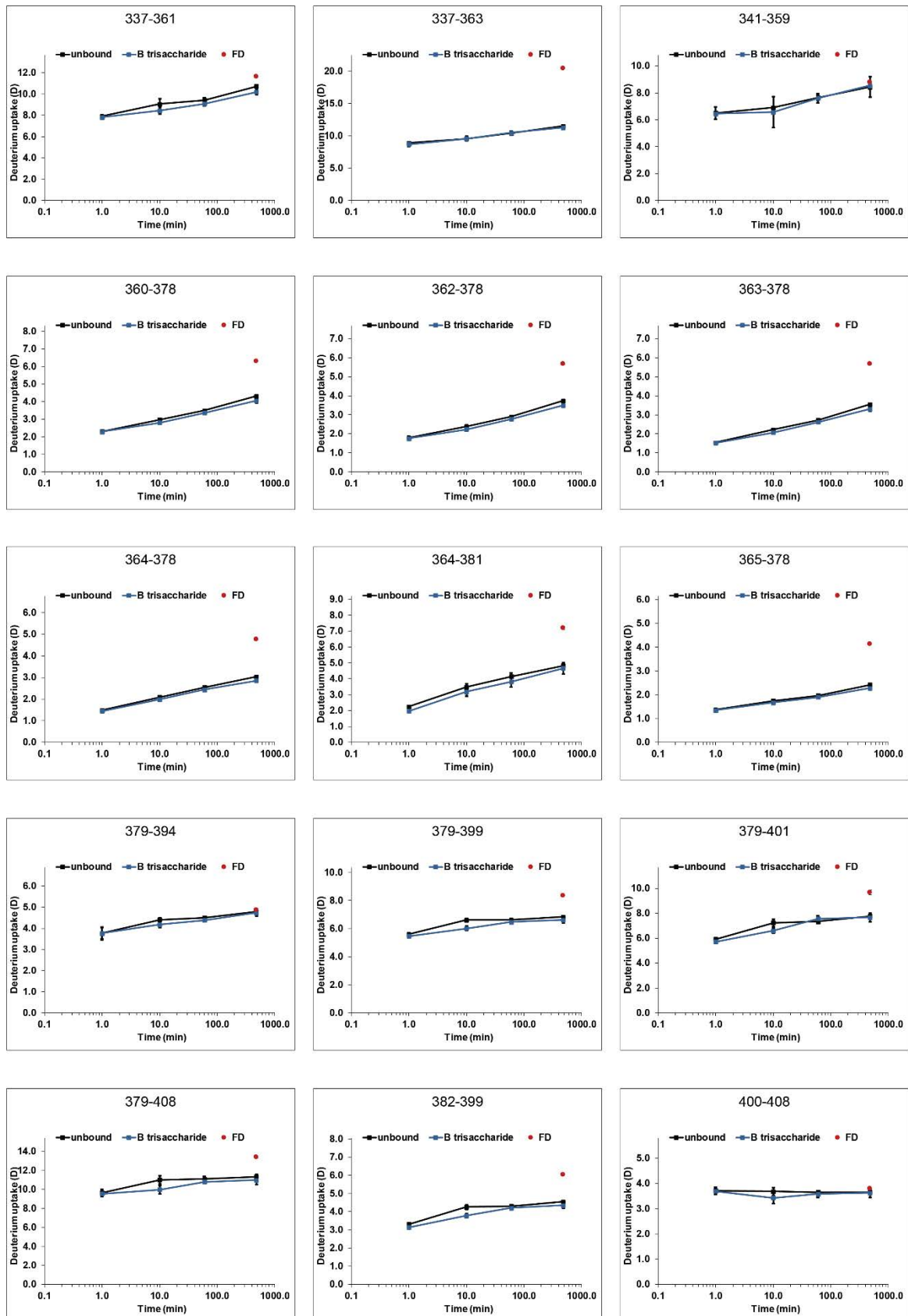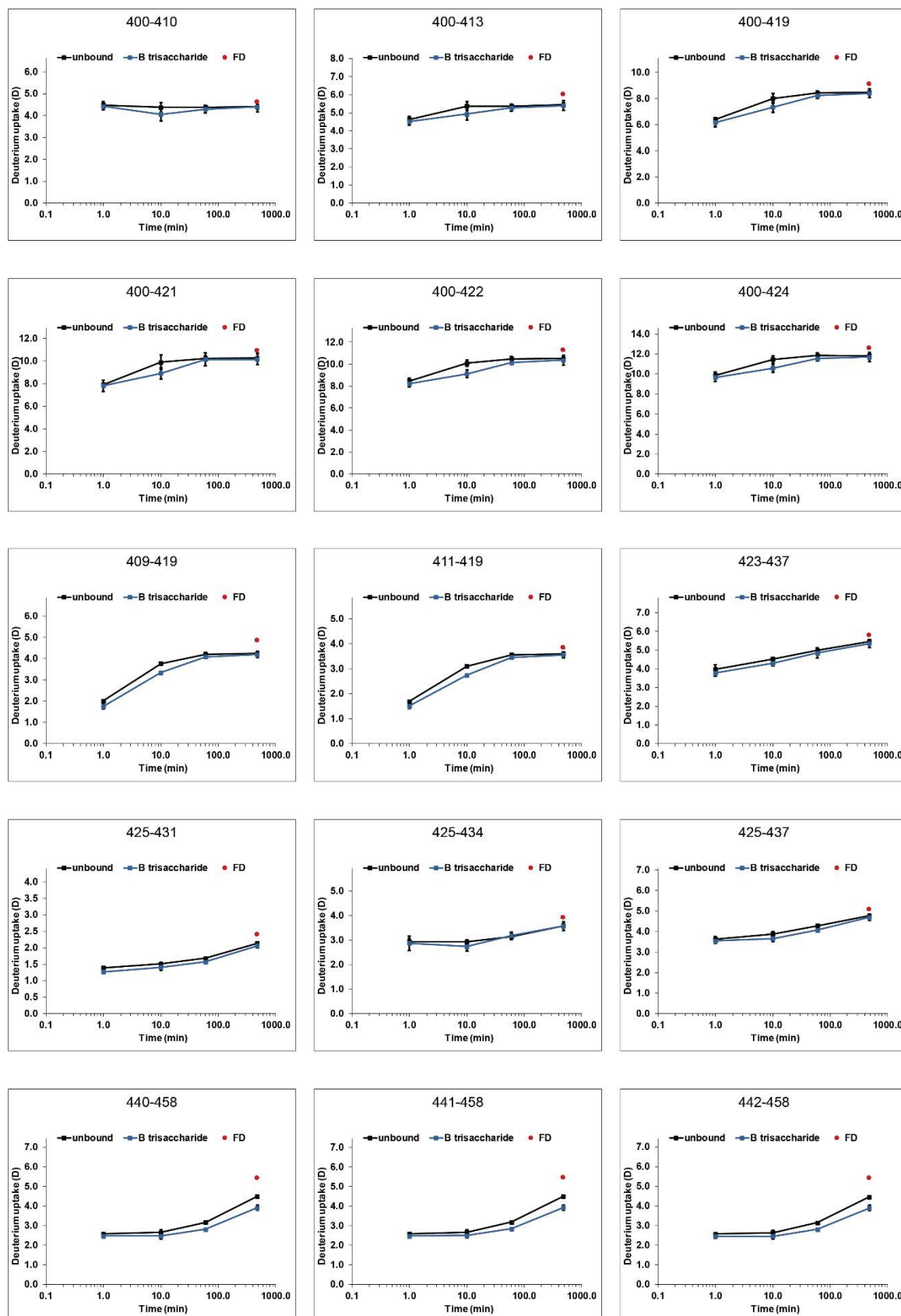
Figure S 39: continued
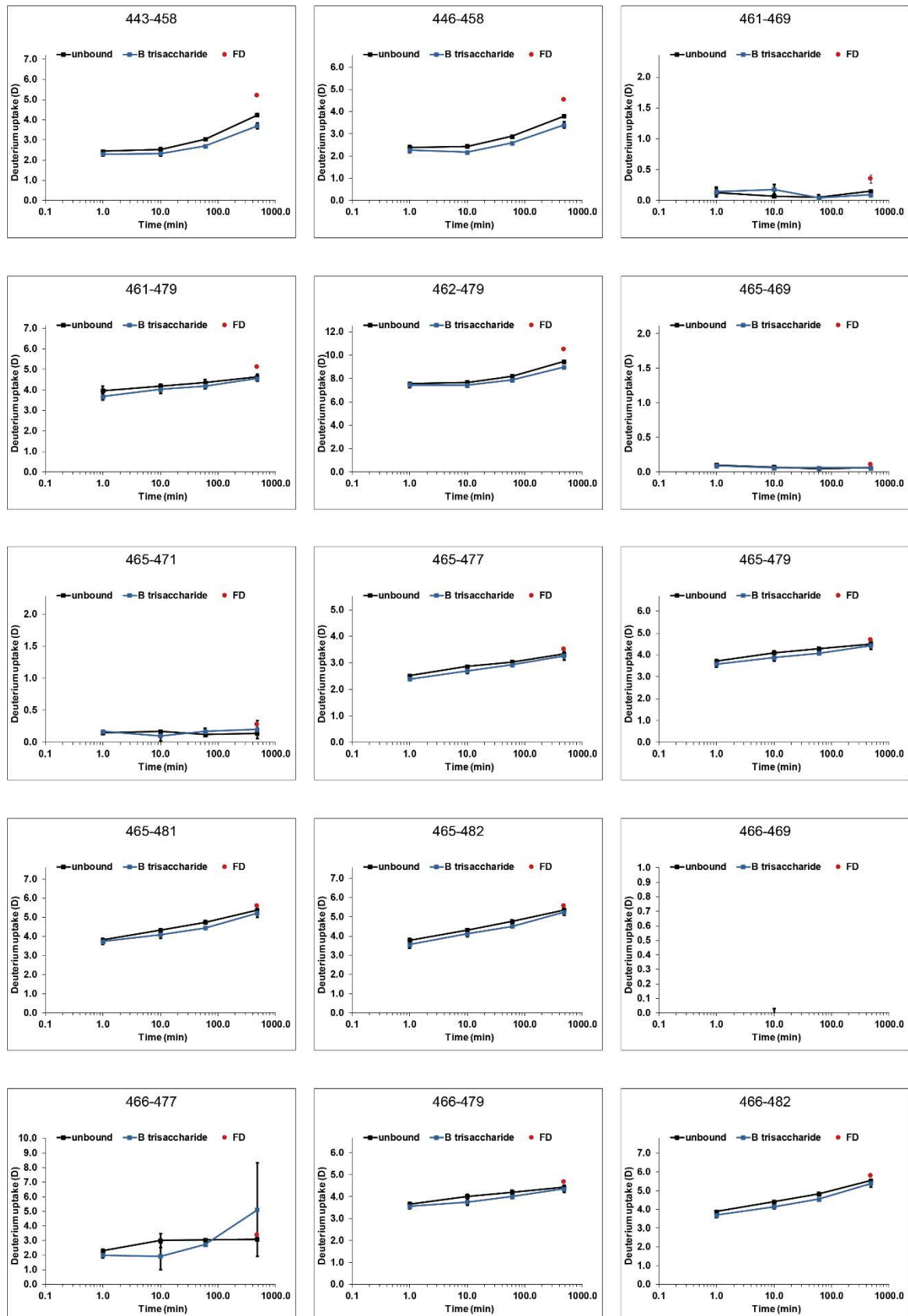
Figure S 39: continued

Figure S 39: continued
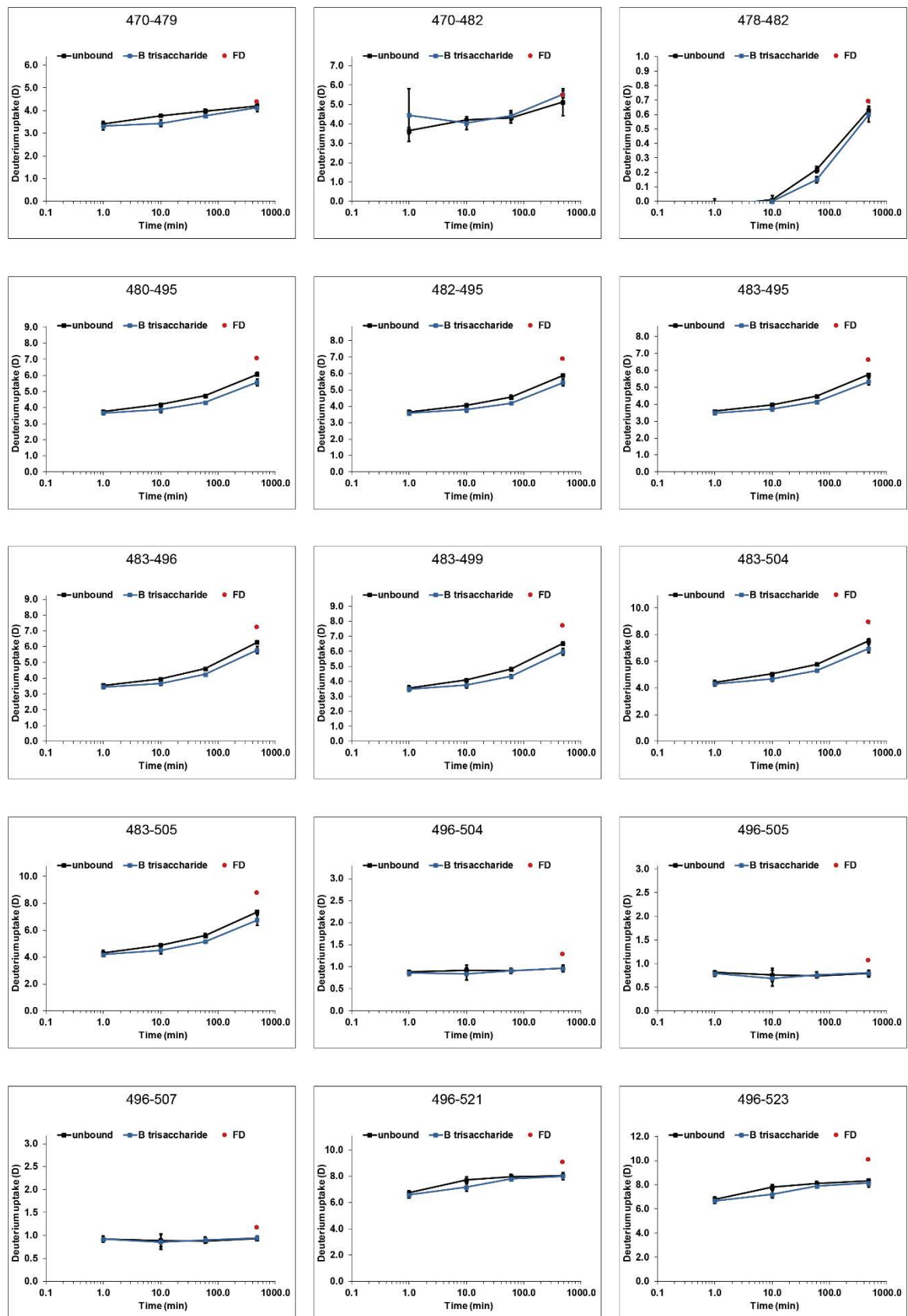
Figure S 39: continued
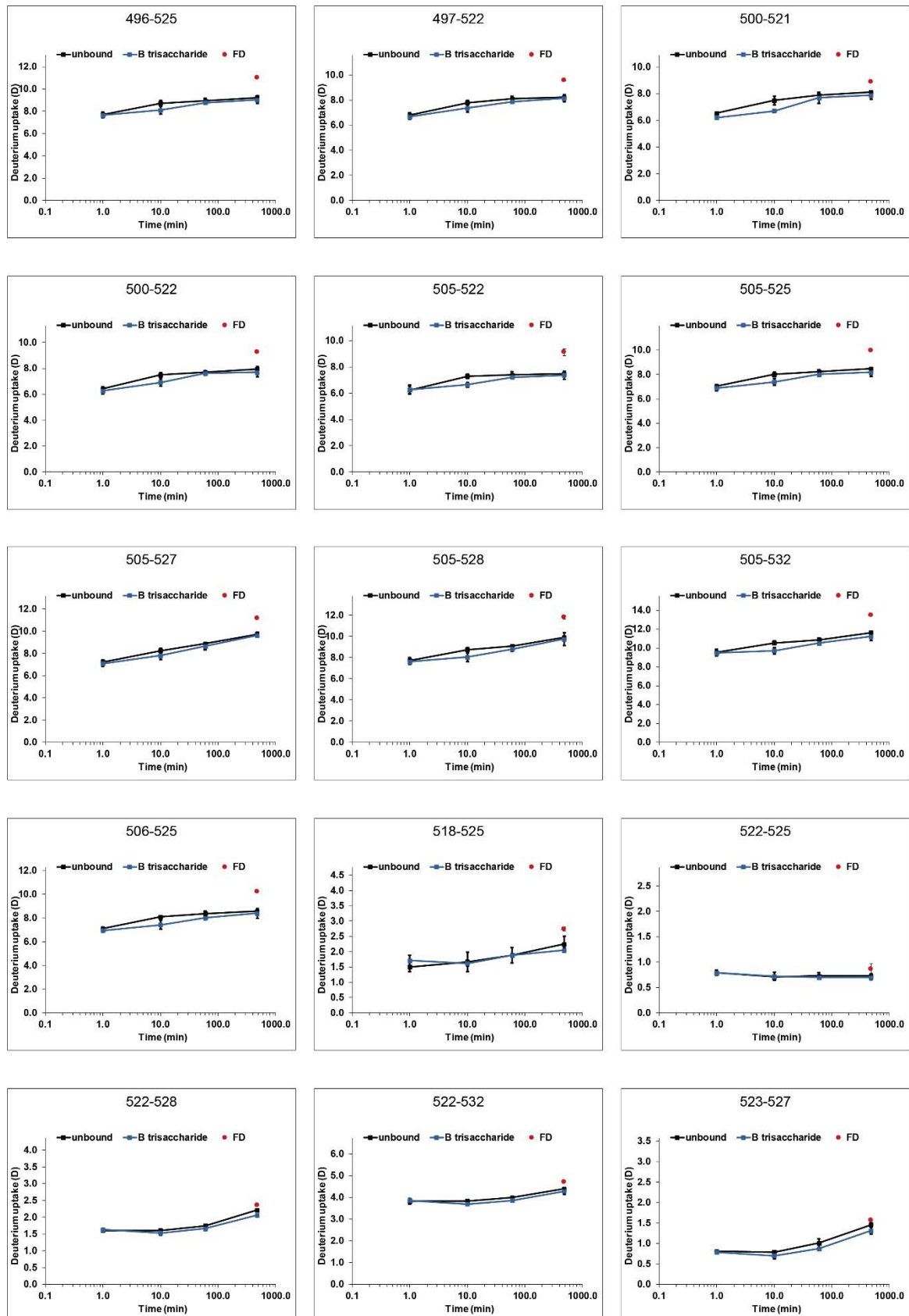
Figure S 39: continued

Figure S 39: continued
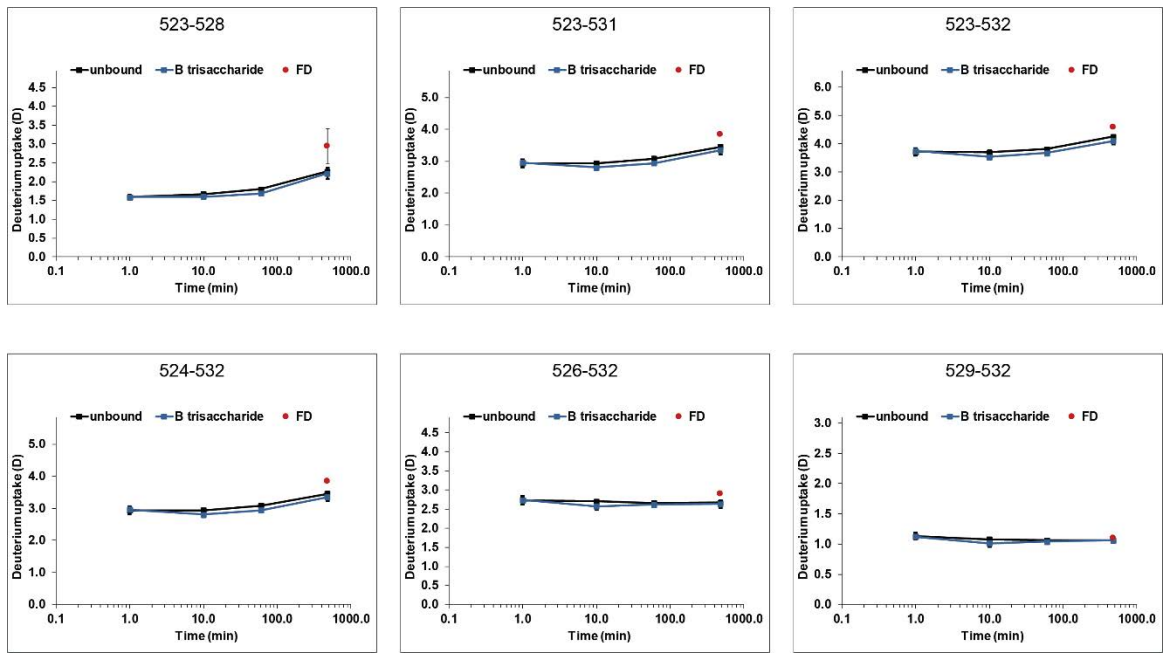
Figure S 39: continued
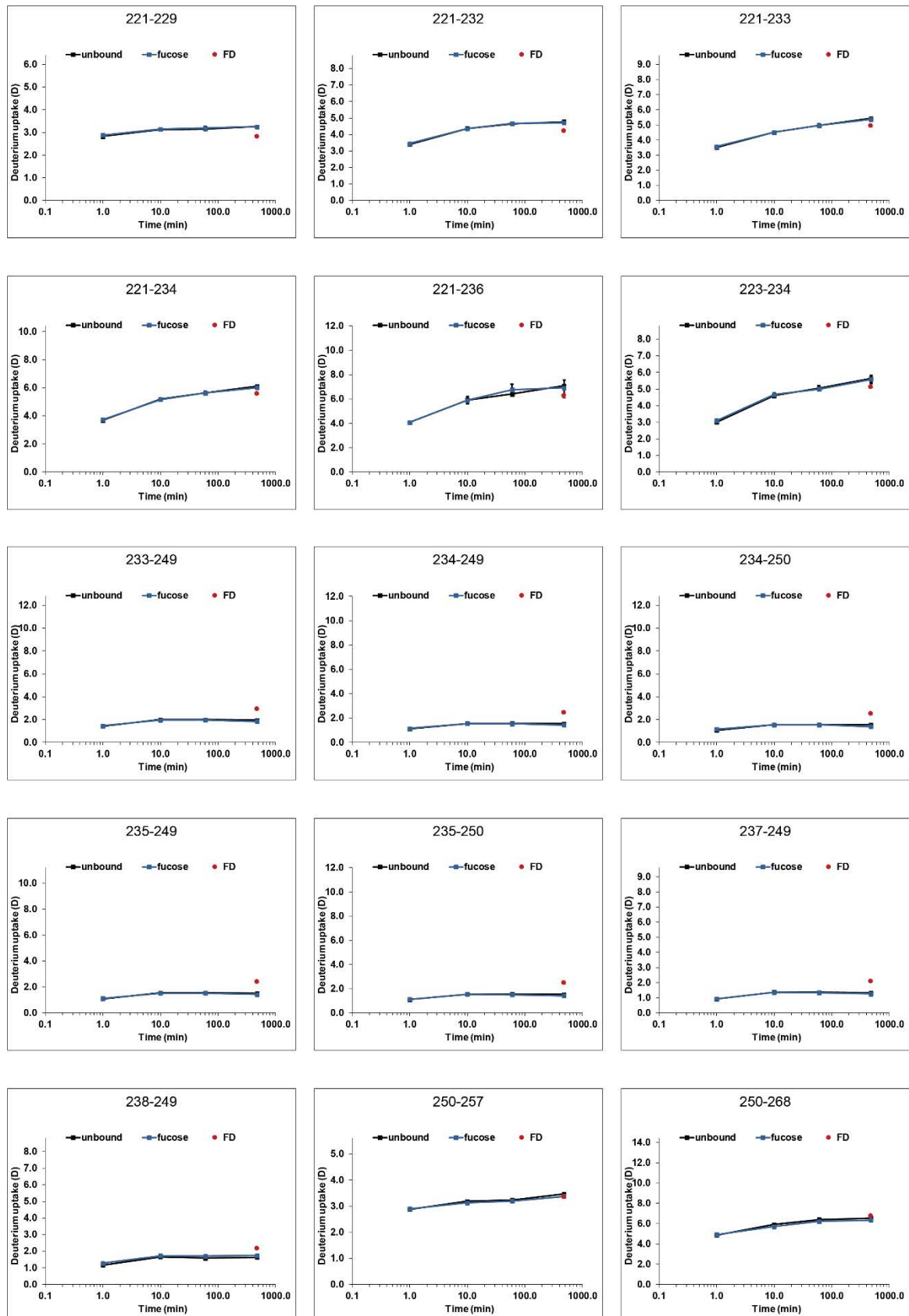
Figure S 40: GII.17 Kawasaki P dimer with 100 mM fucose
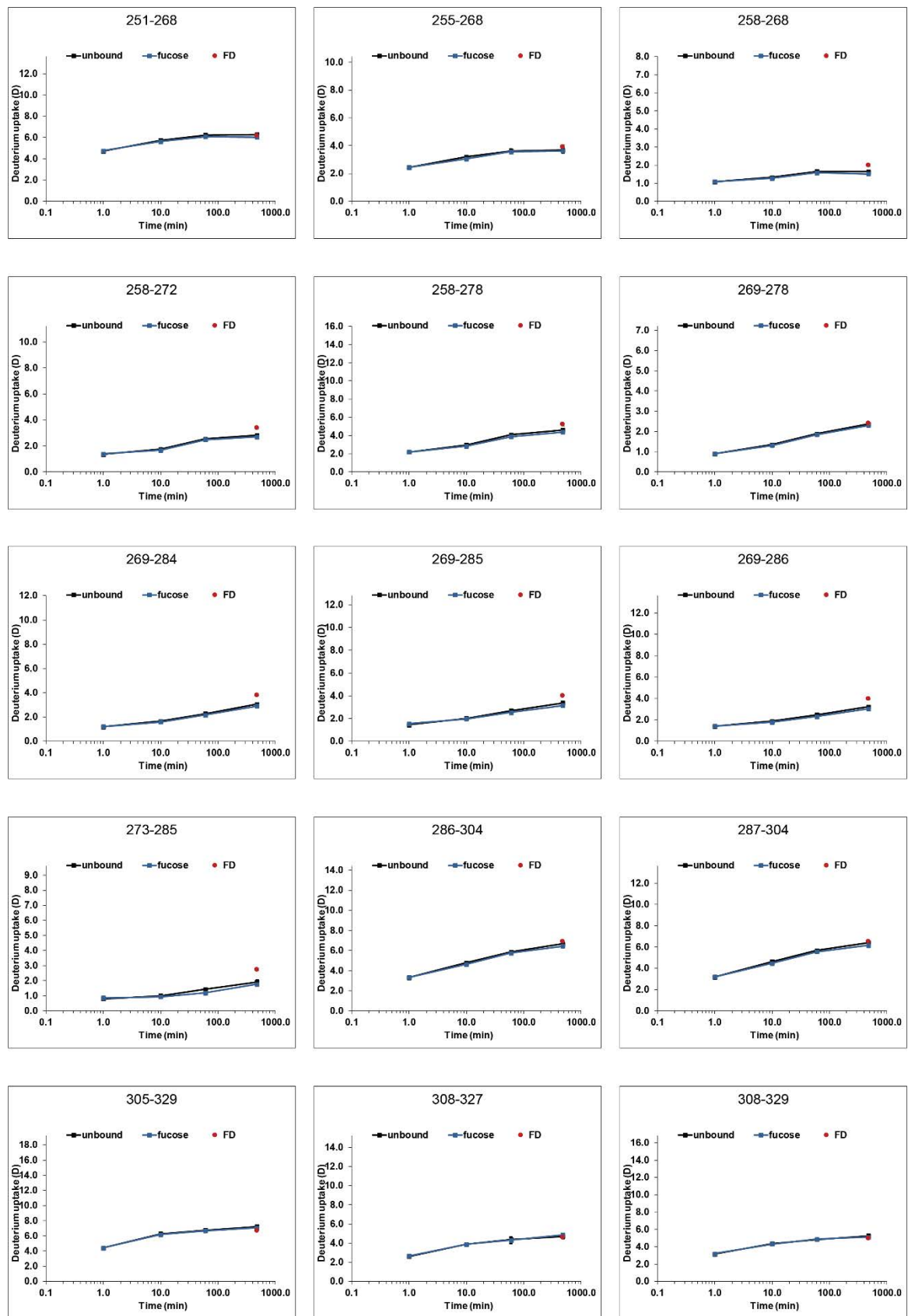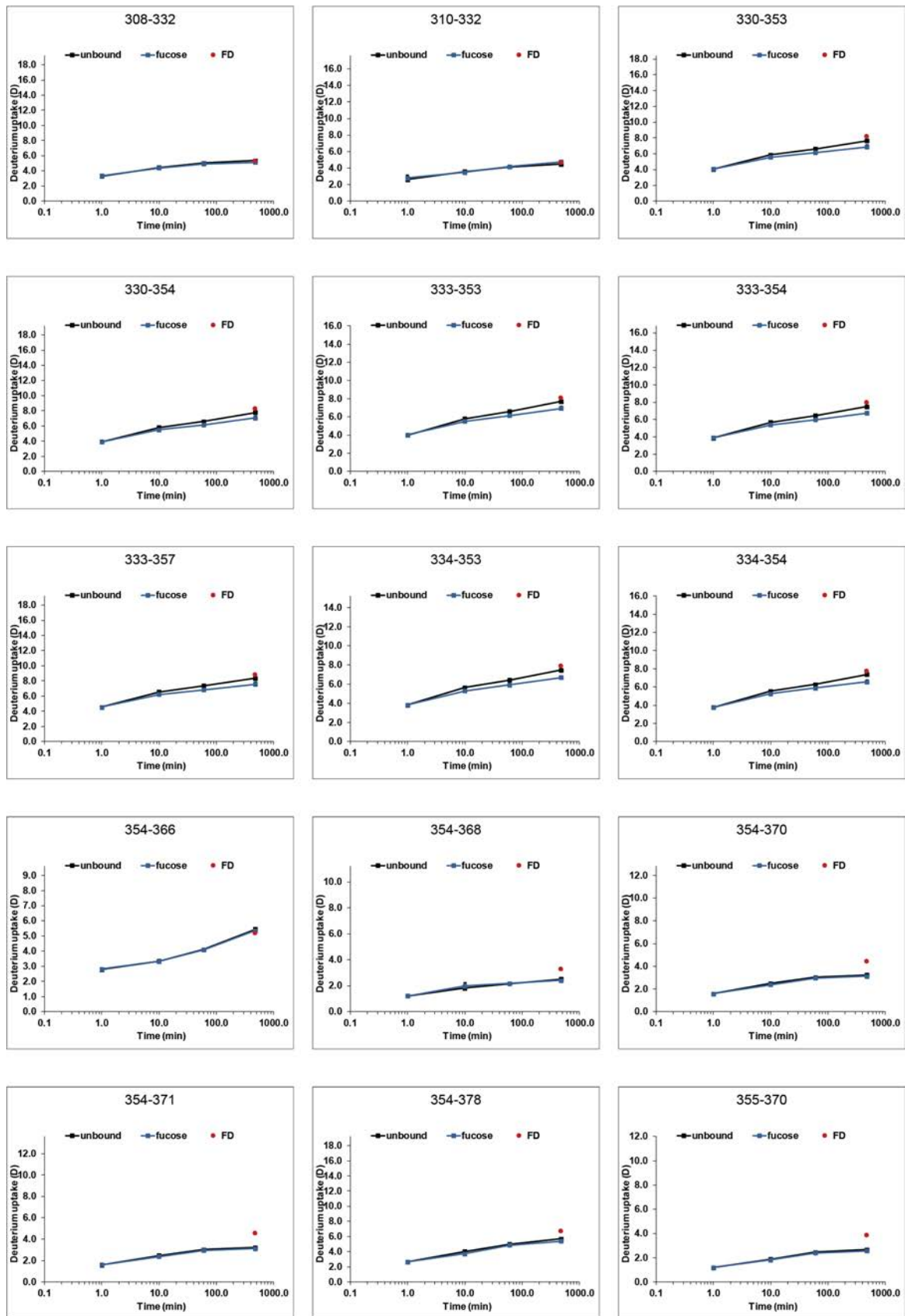
Figure S 40: continued
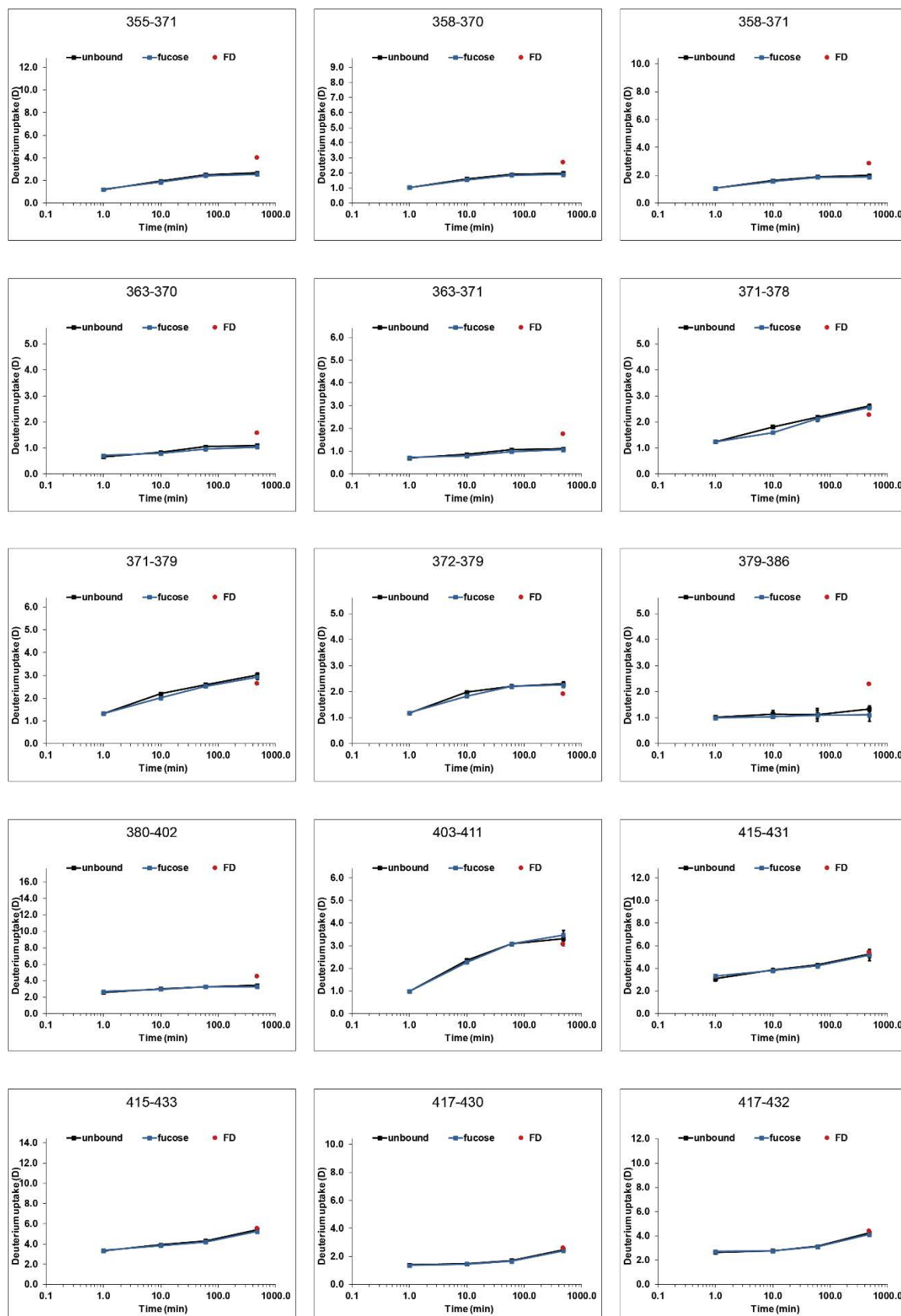
Figure S 40: continued
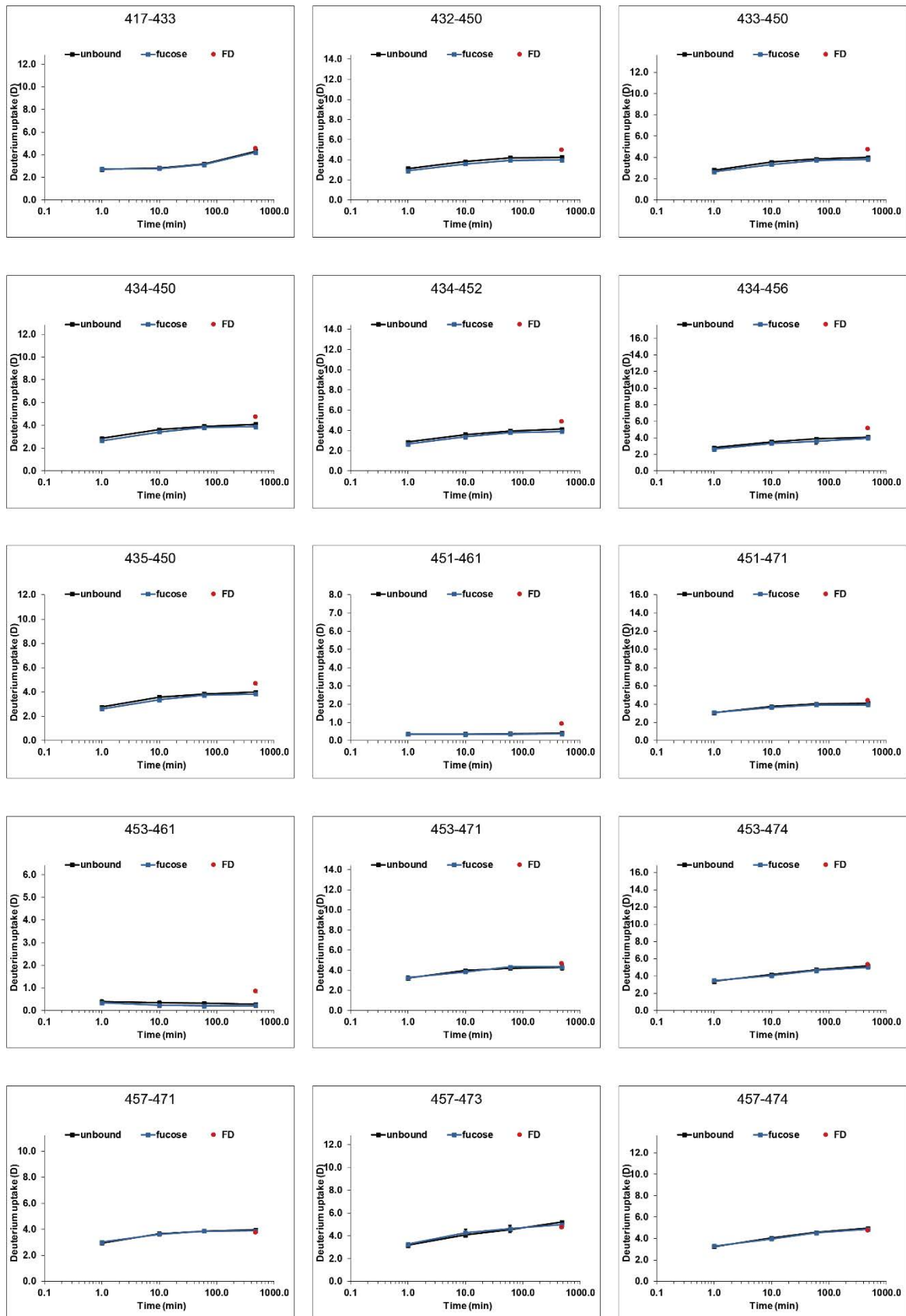
Figure S 40: continued

Figure S 40: continued
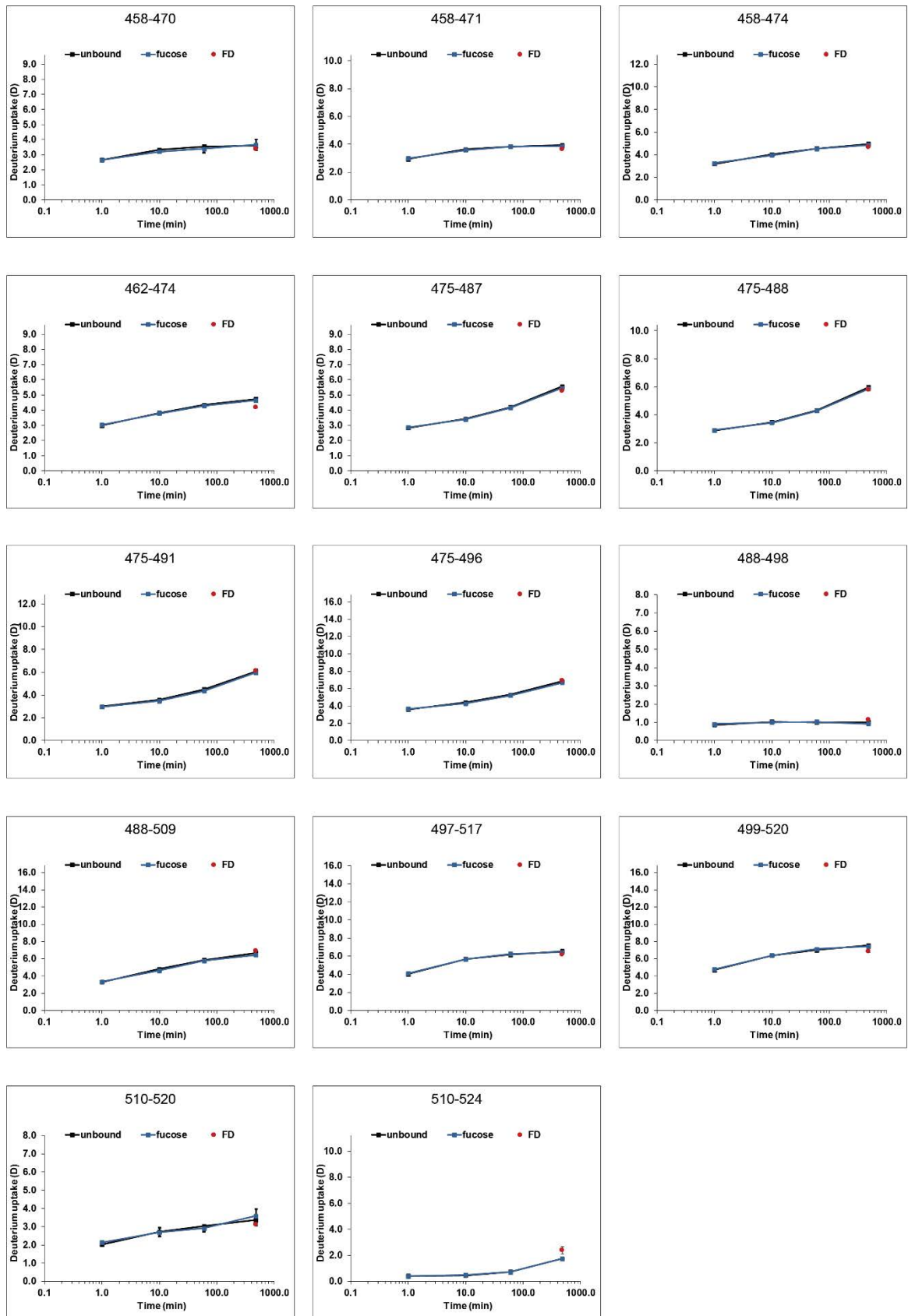
Figure S 40: continued
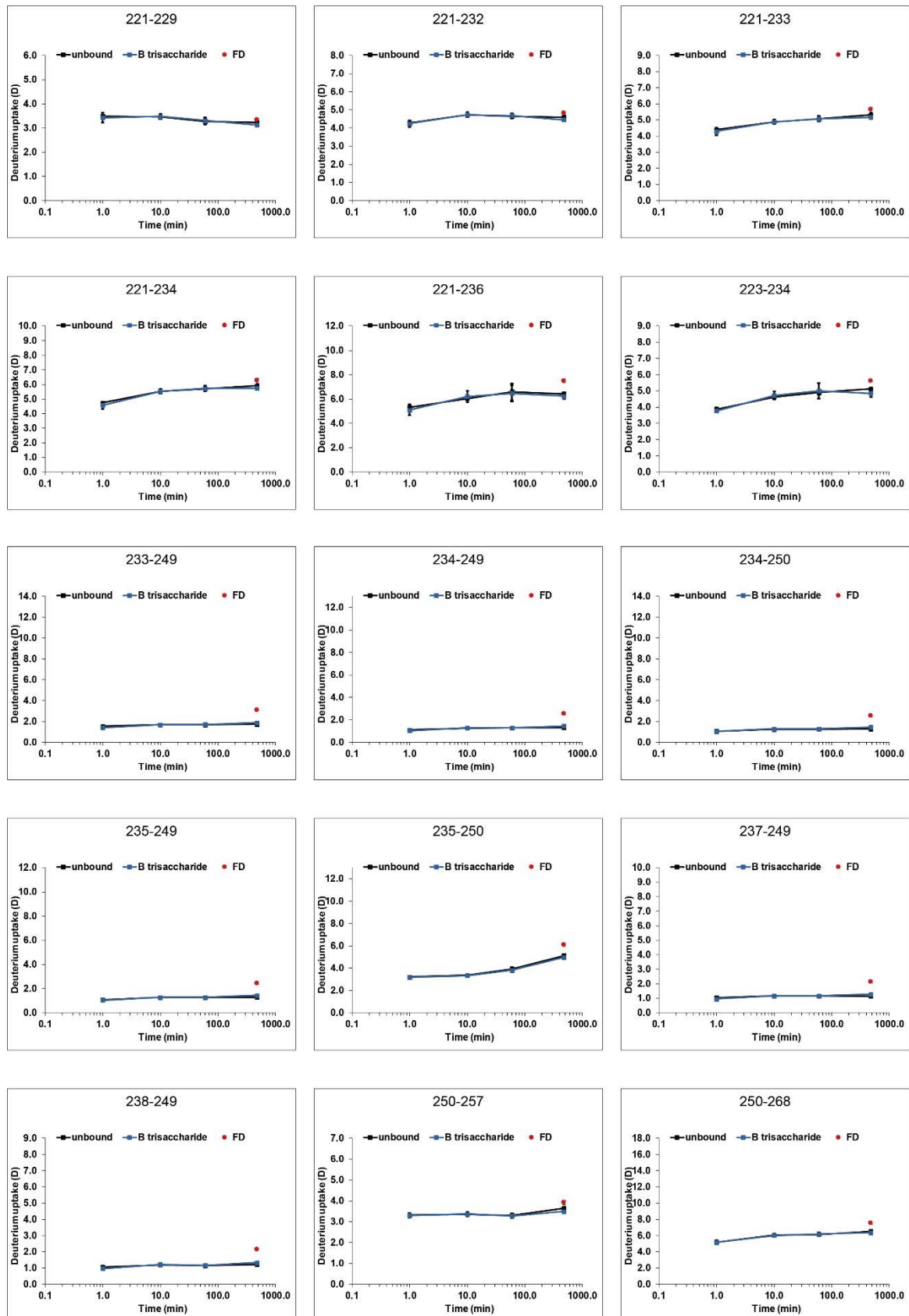
Figure S 41: GII.17 Kawasaki P dimer with B trisaccharide
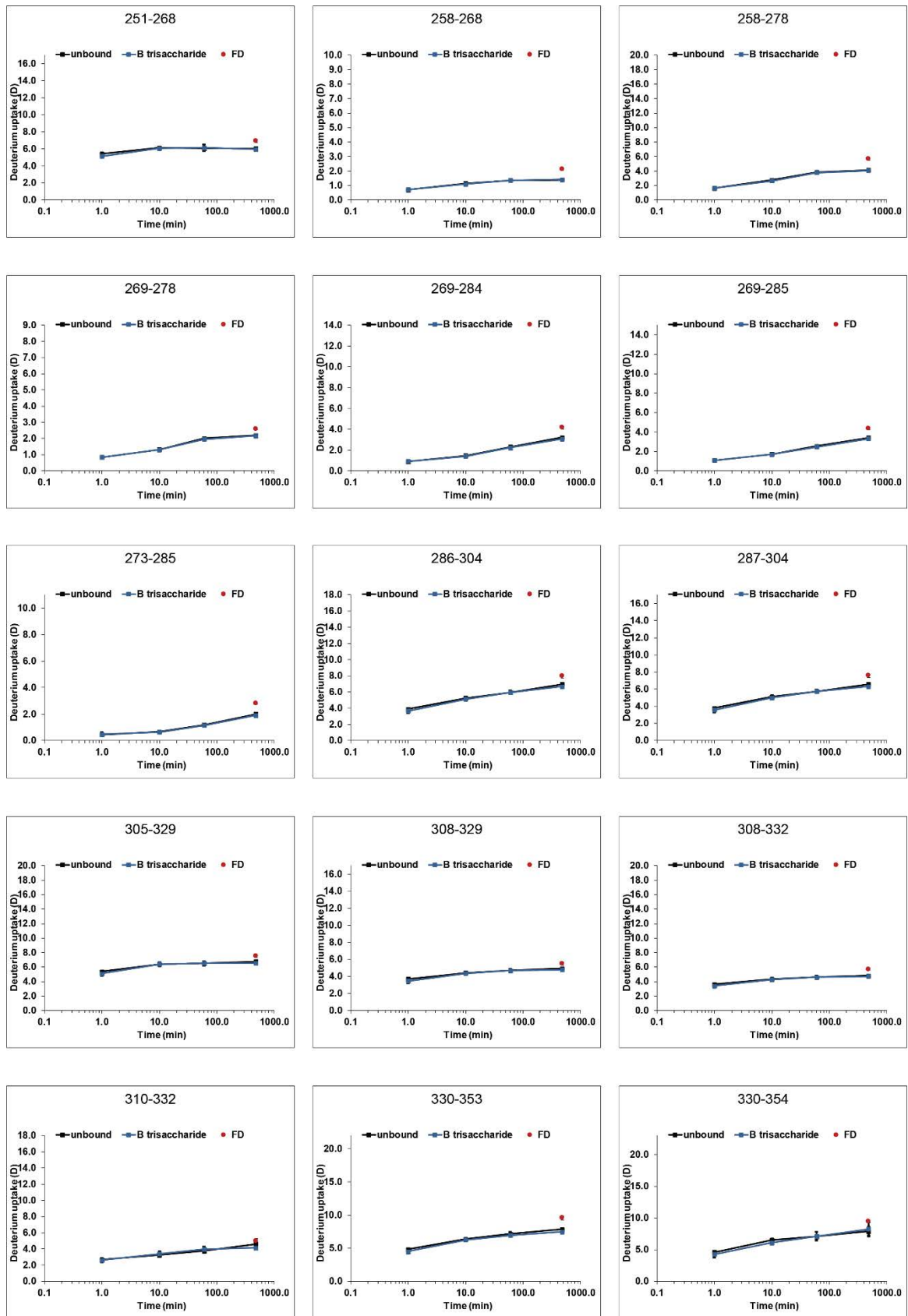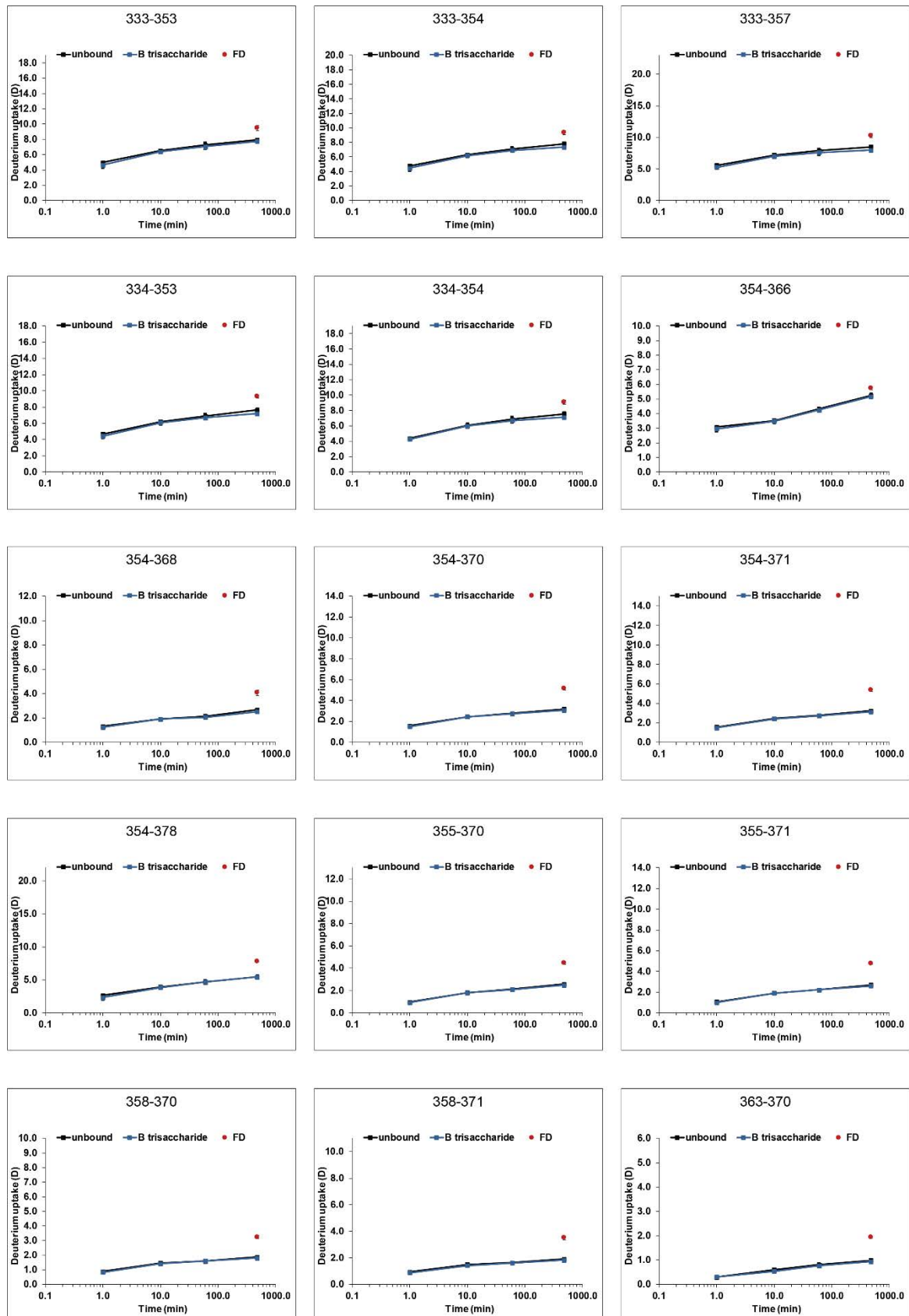
Figure S 41: continued
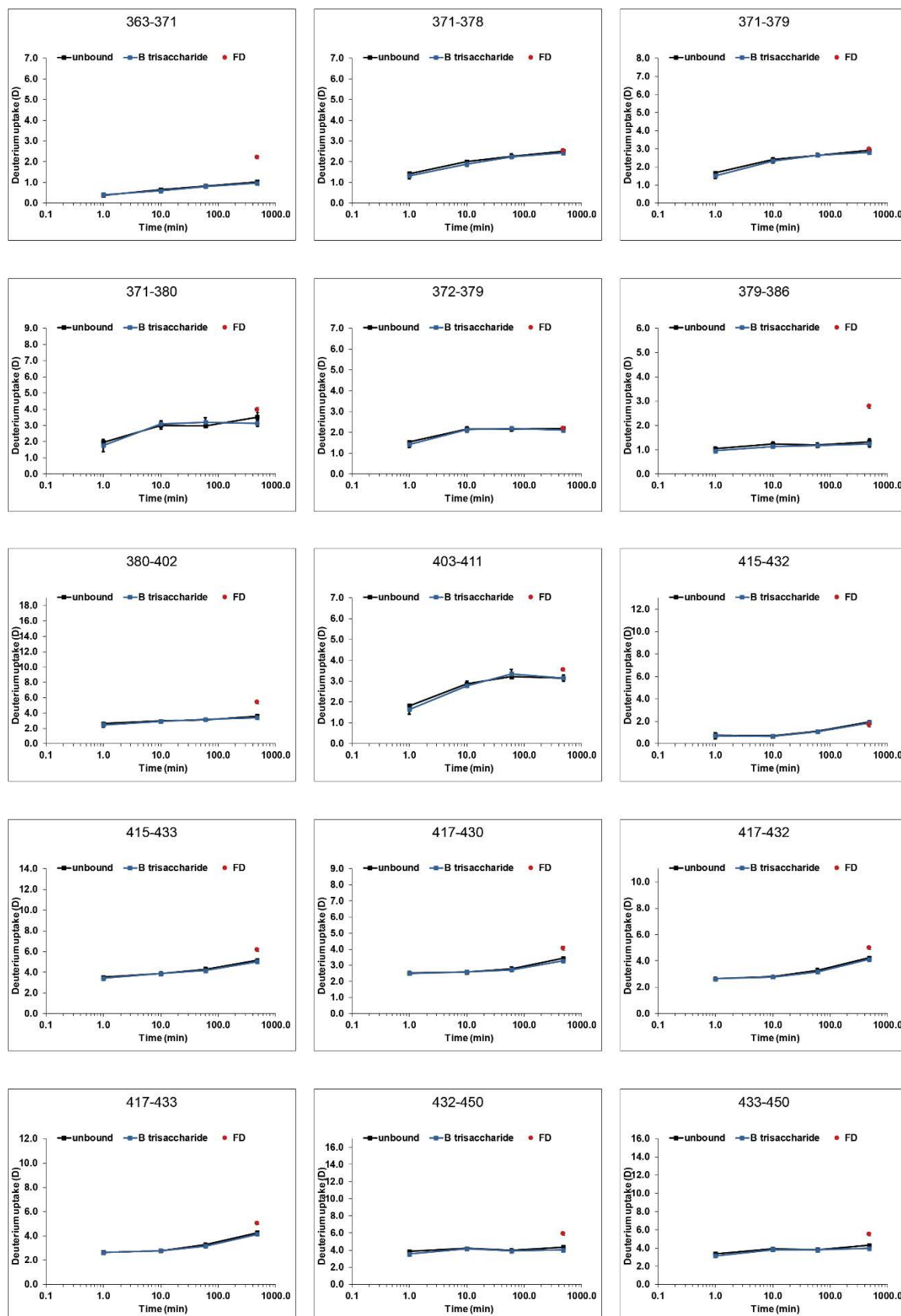
Figure S 41: continued
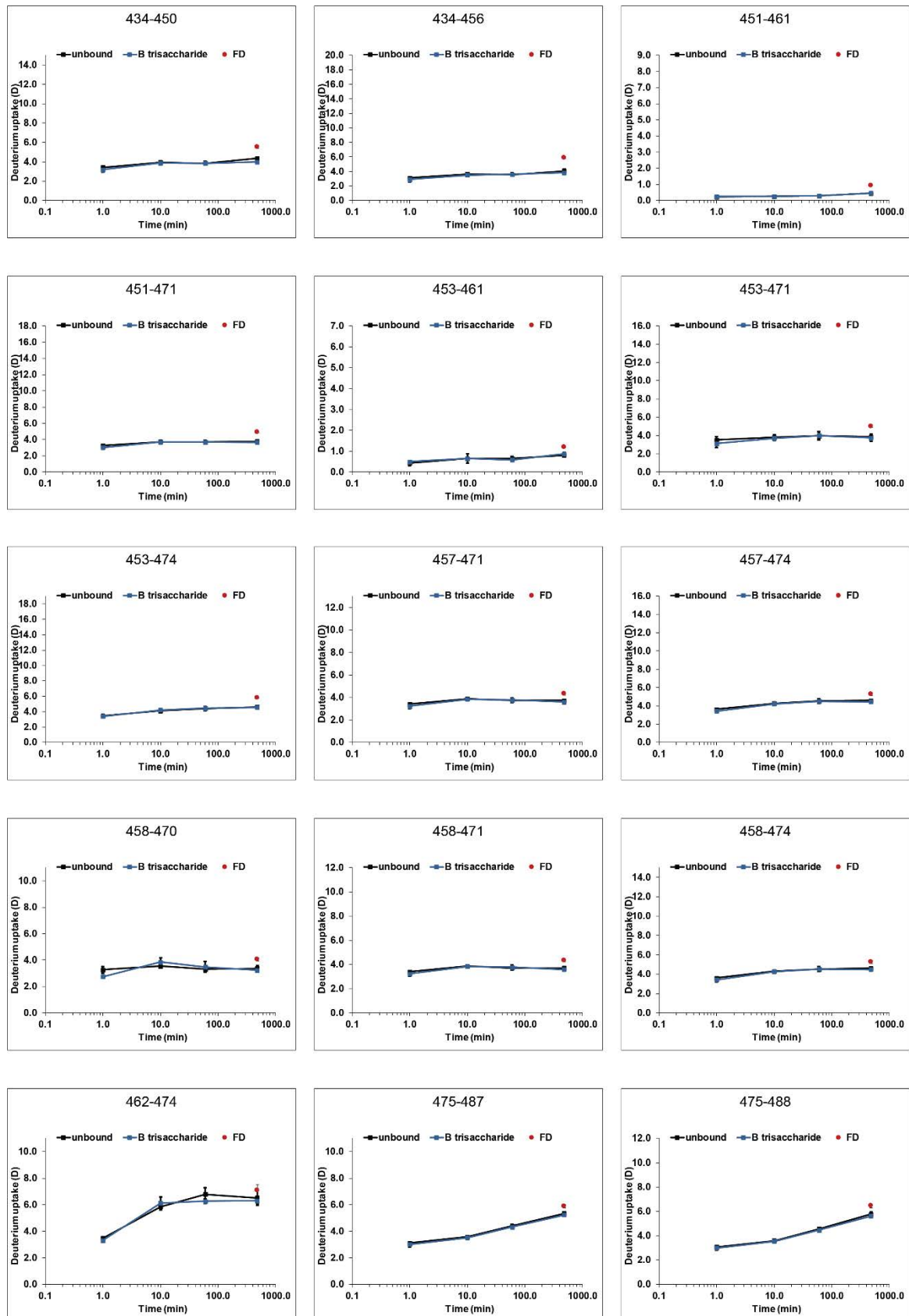
Figure S 41: continued

Figure S 41: continued
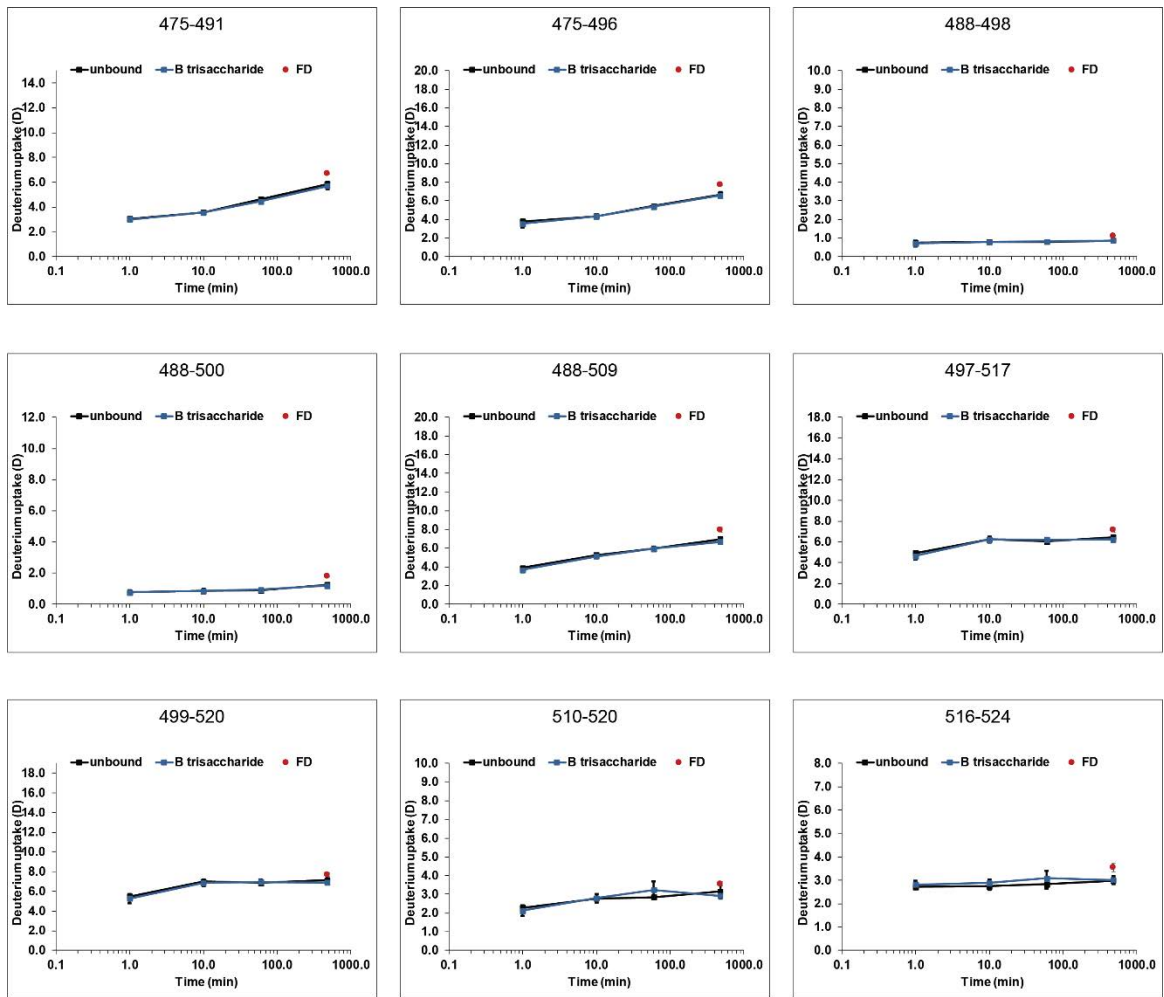
Figure S 41: continued
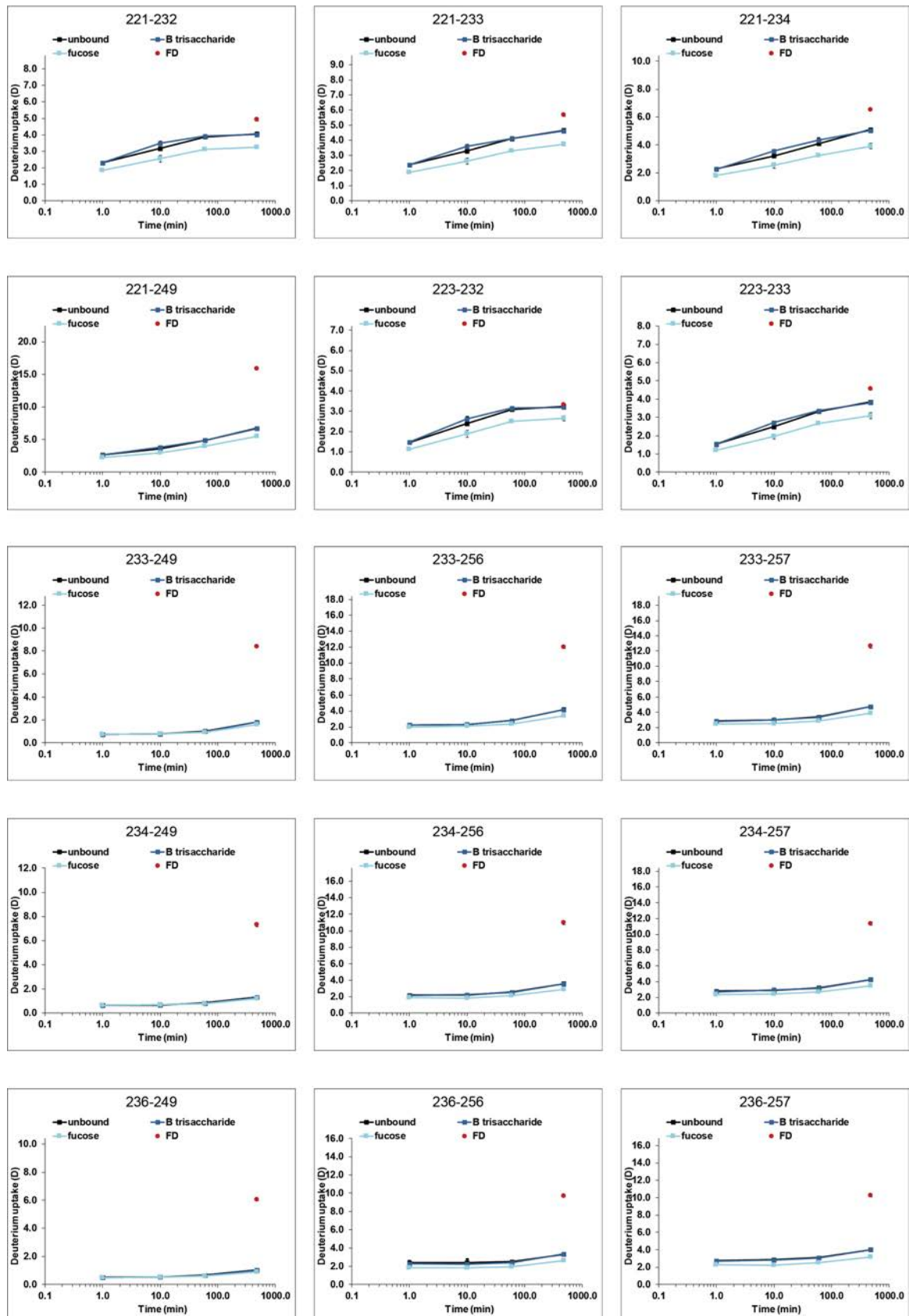
Figure S 42: GII.4 MI001 wt P dimer with 10 mM B trisaccharide, 100 mM fucose
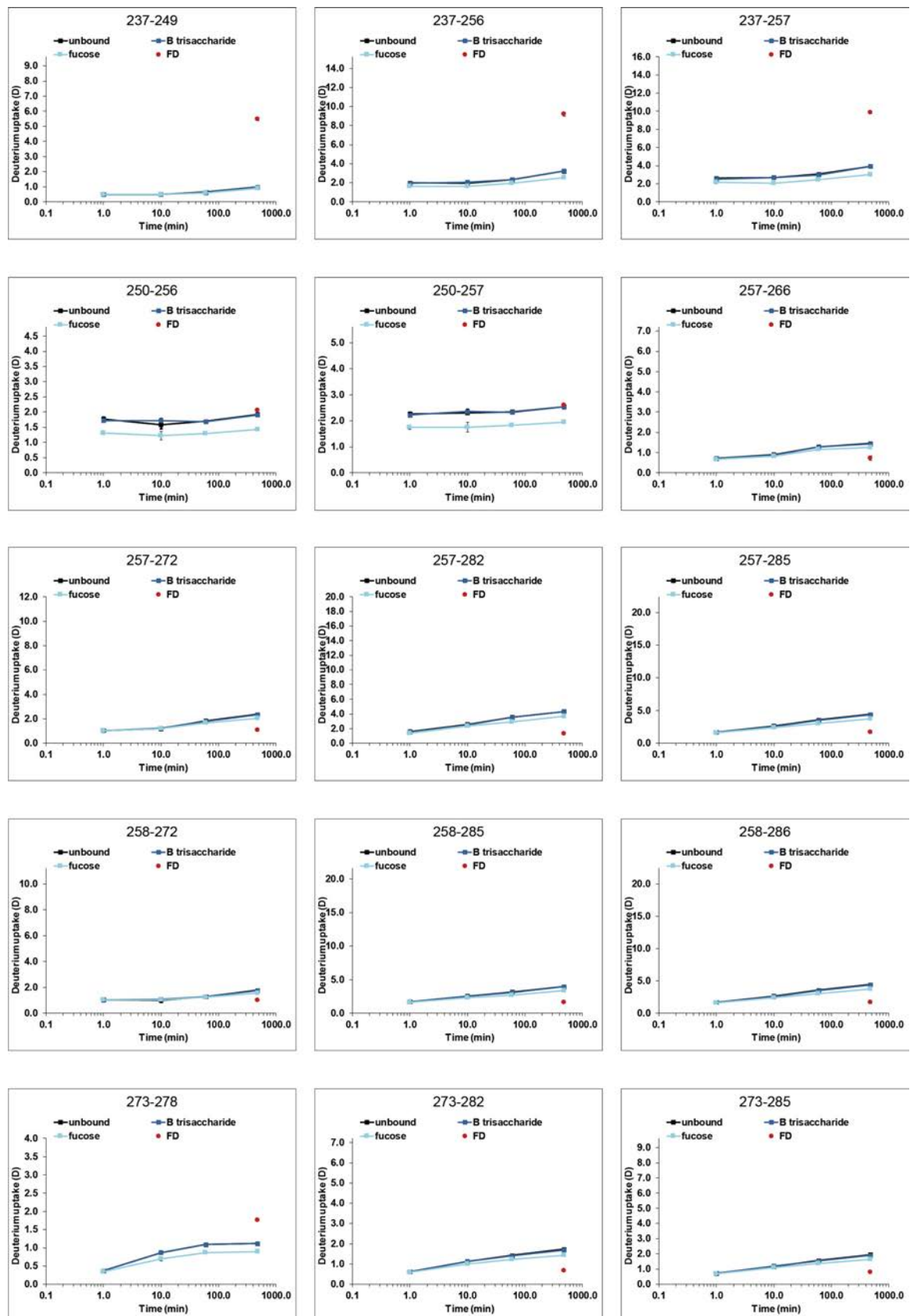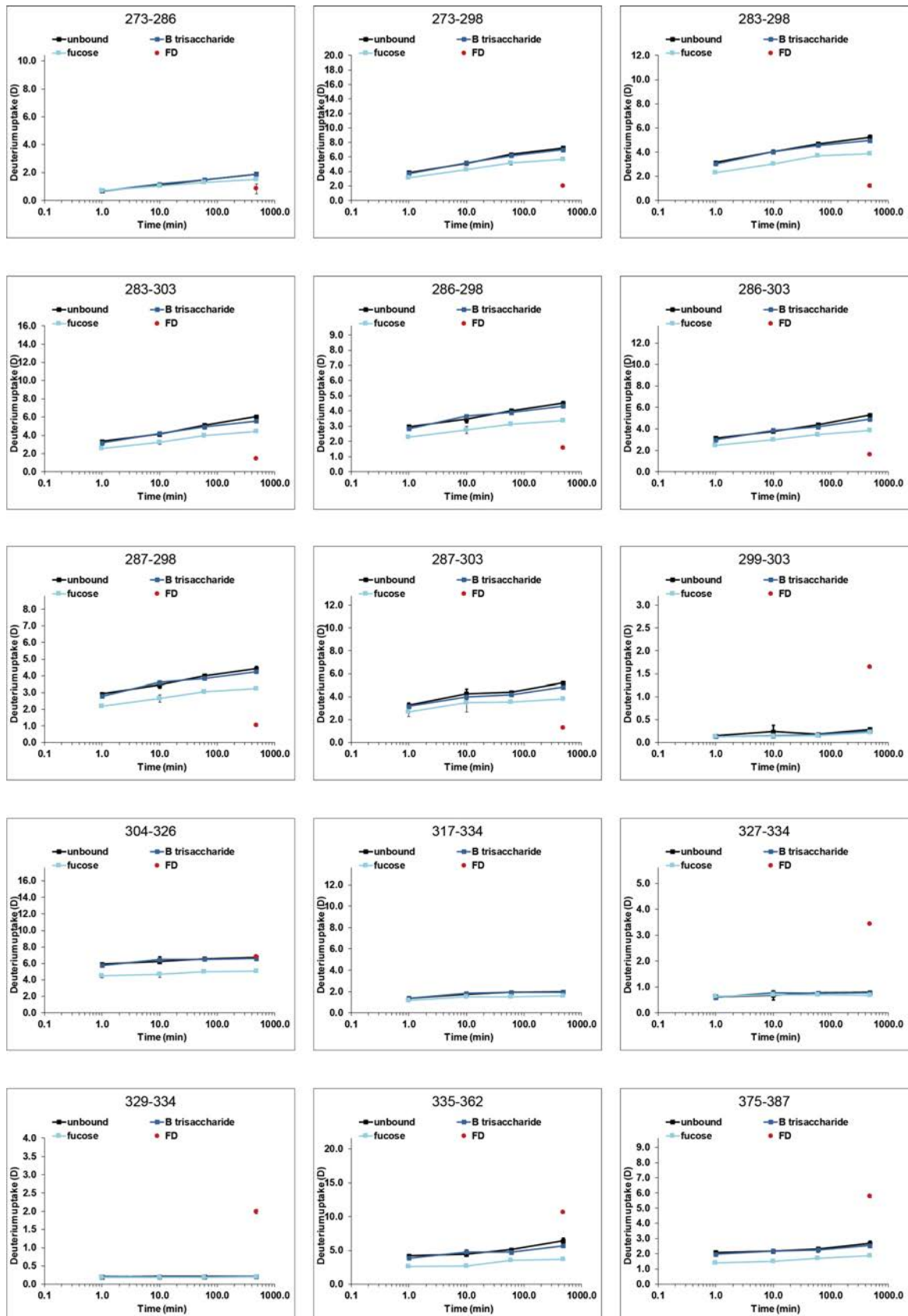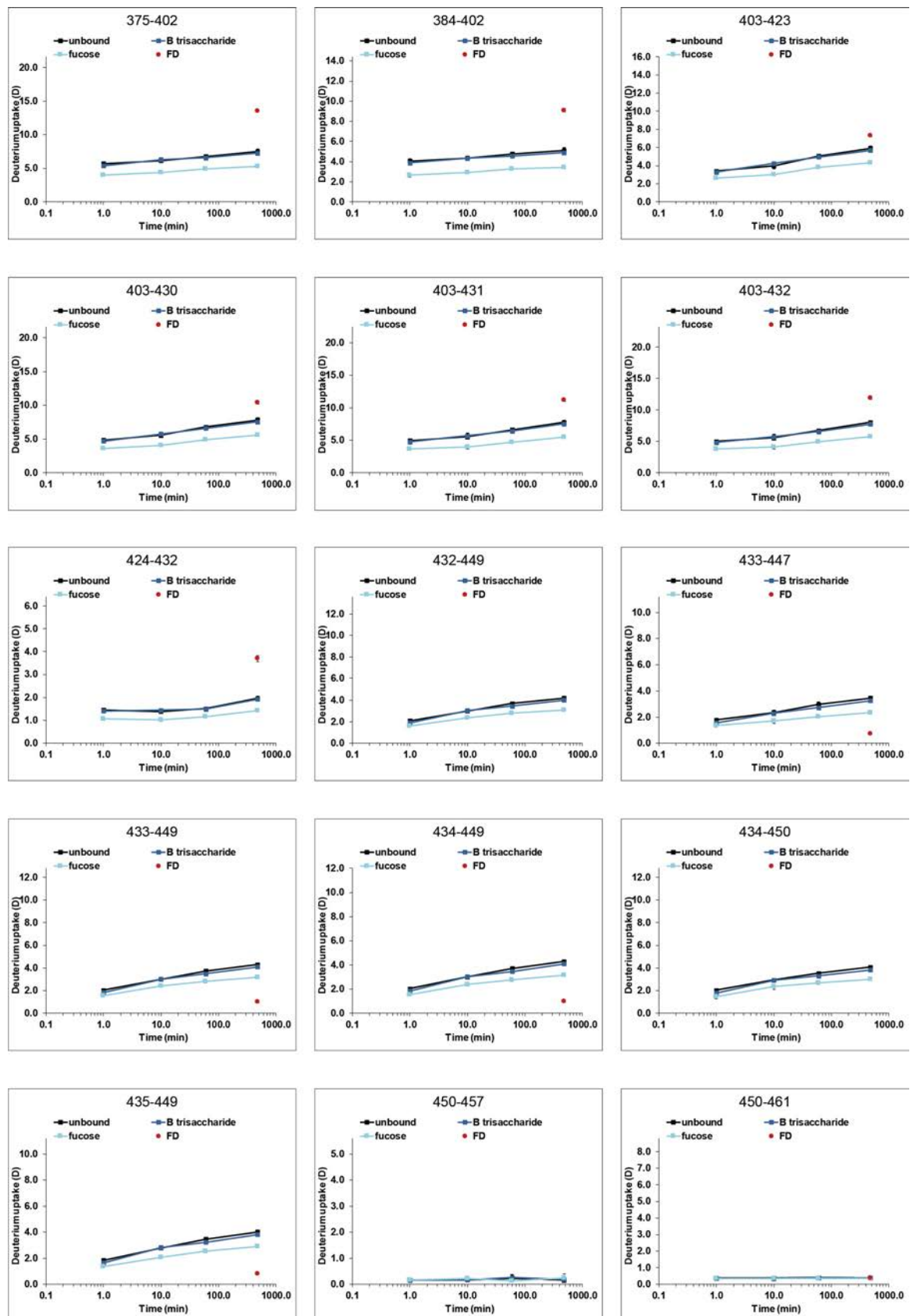
Figure S 42: continued
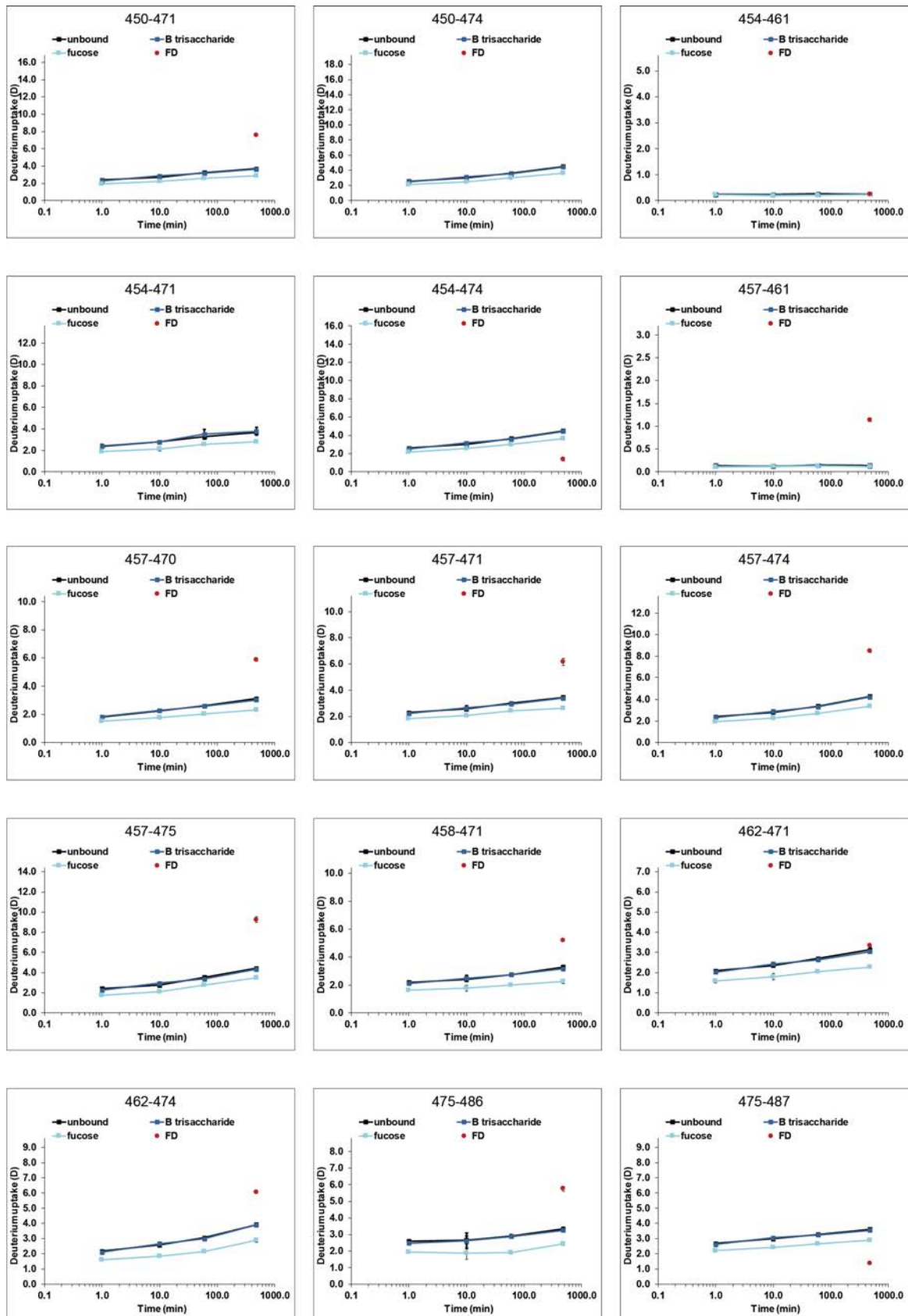
Figure S 42: continued

Figure S 42: continued
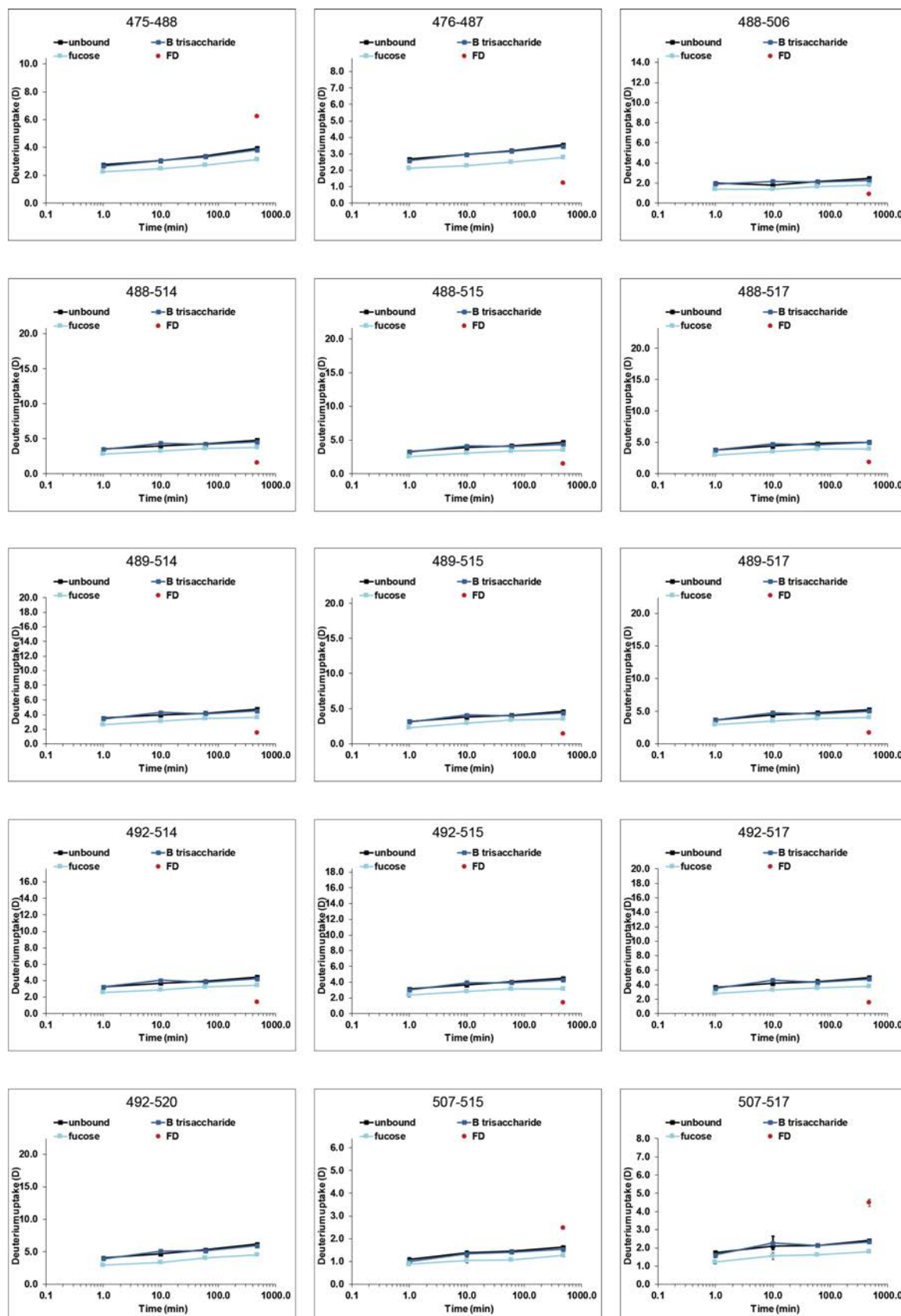
Figure S 42: continued
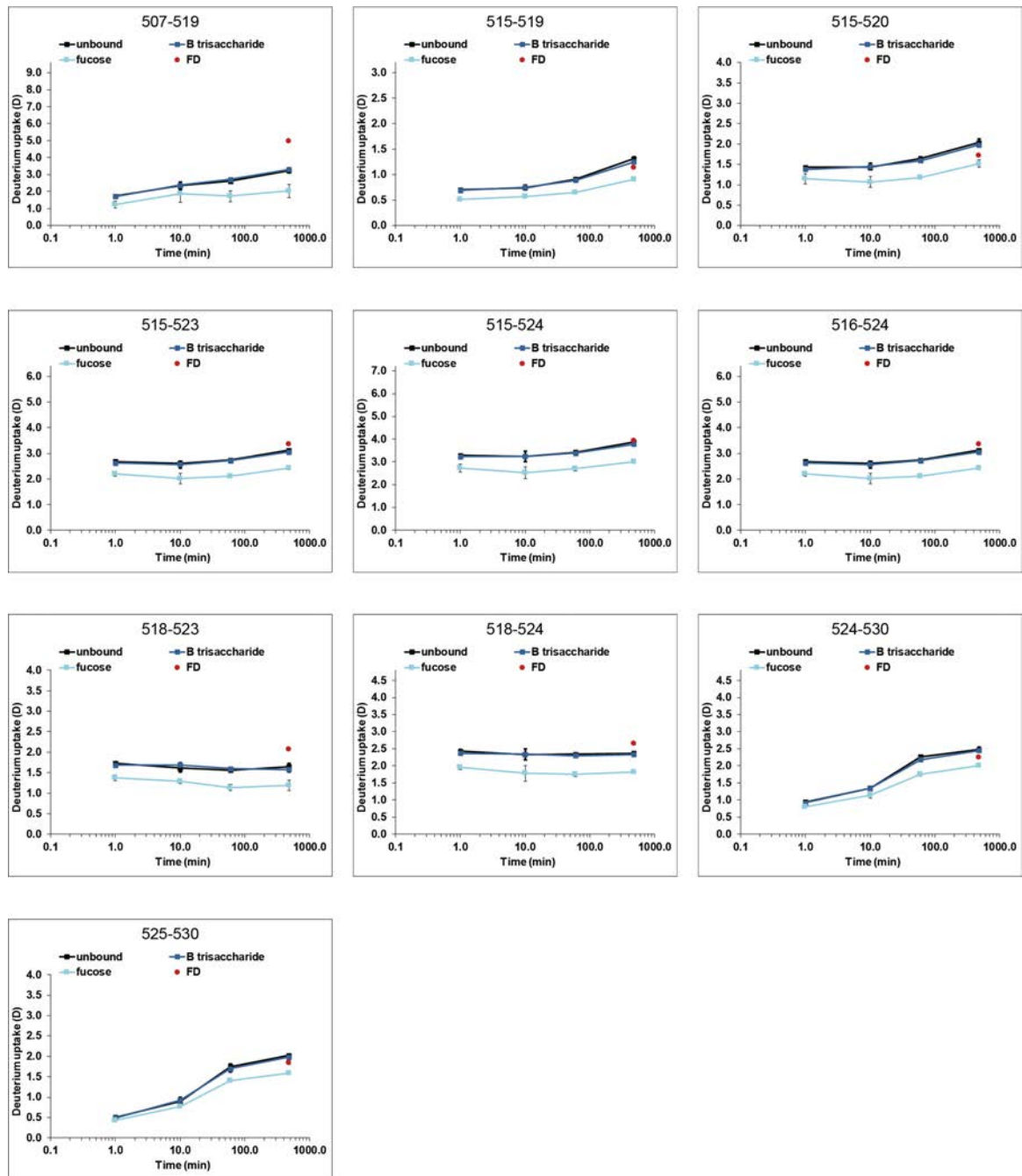
Figure S 42: continued

Figure S 42: continued
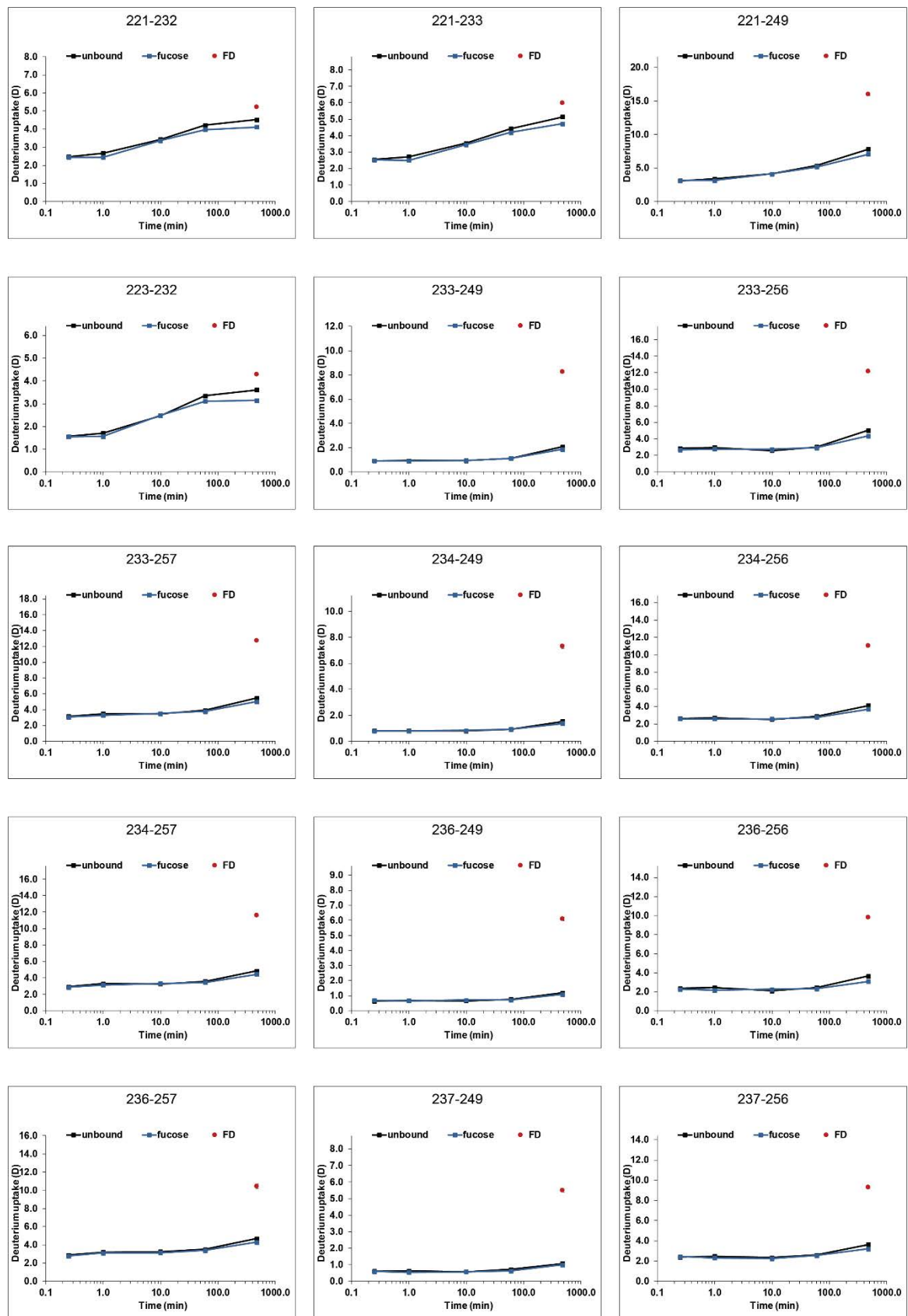
Figure S 43: GII.4 MI001 wild type P dimer with 100 mM fucose (single replicates)
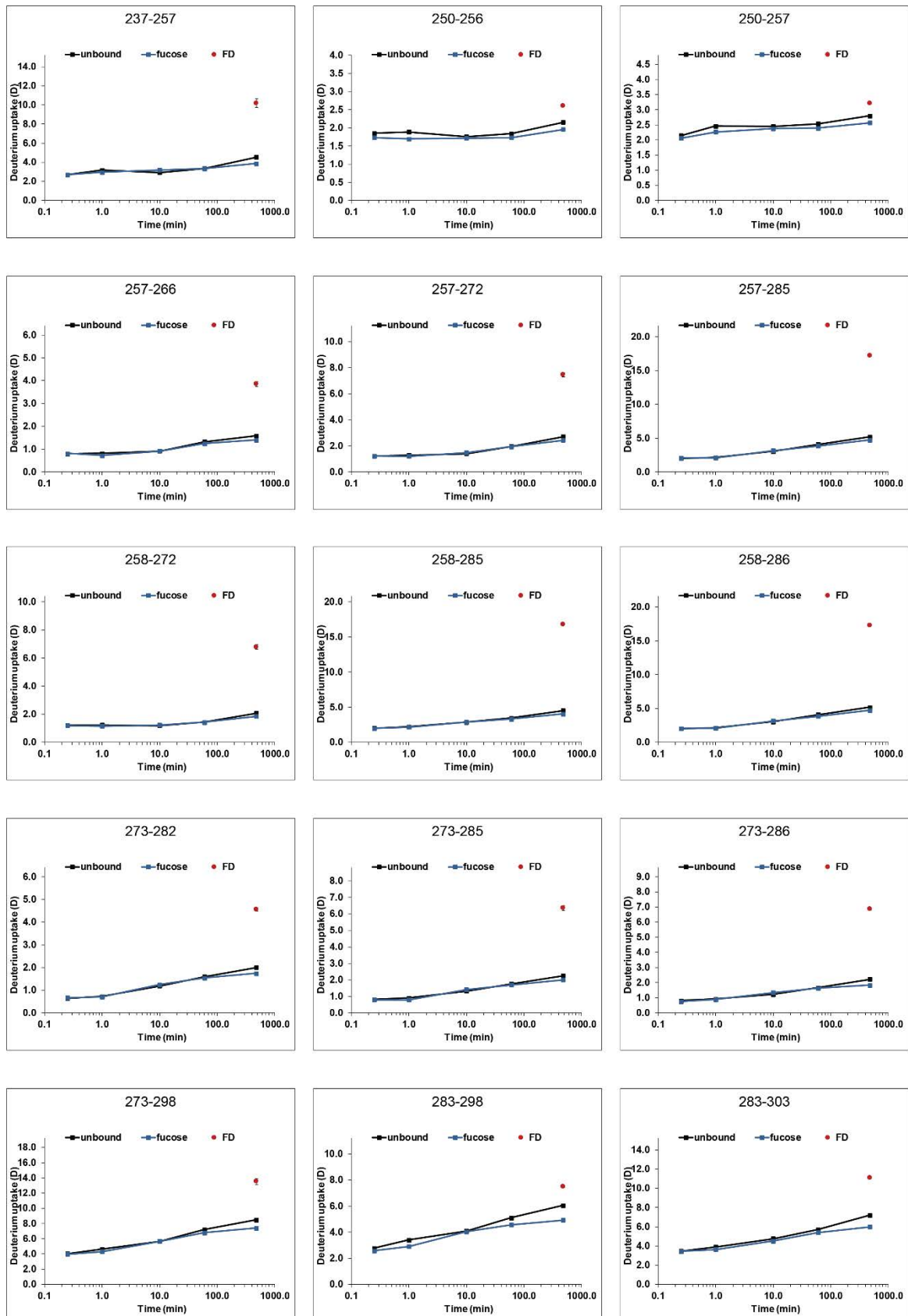
Figure S 43: continued
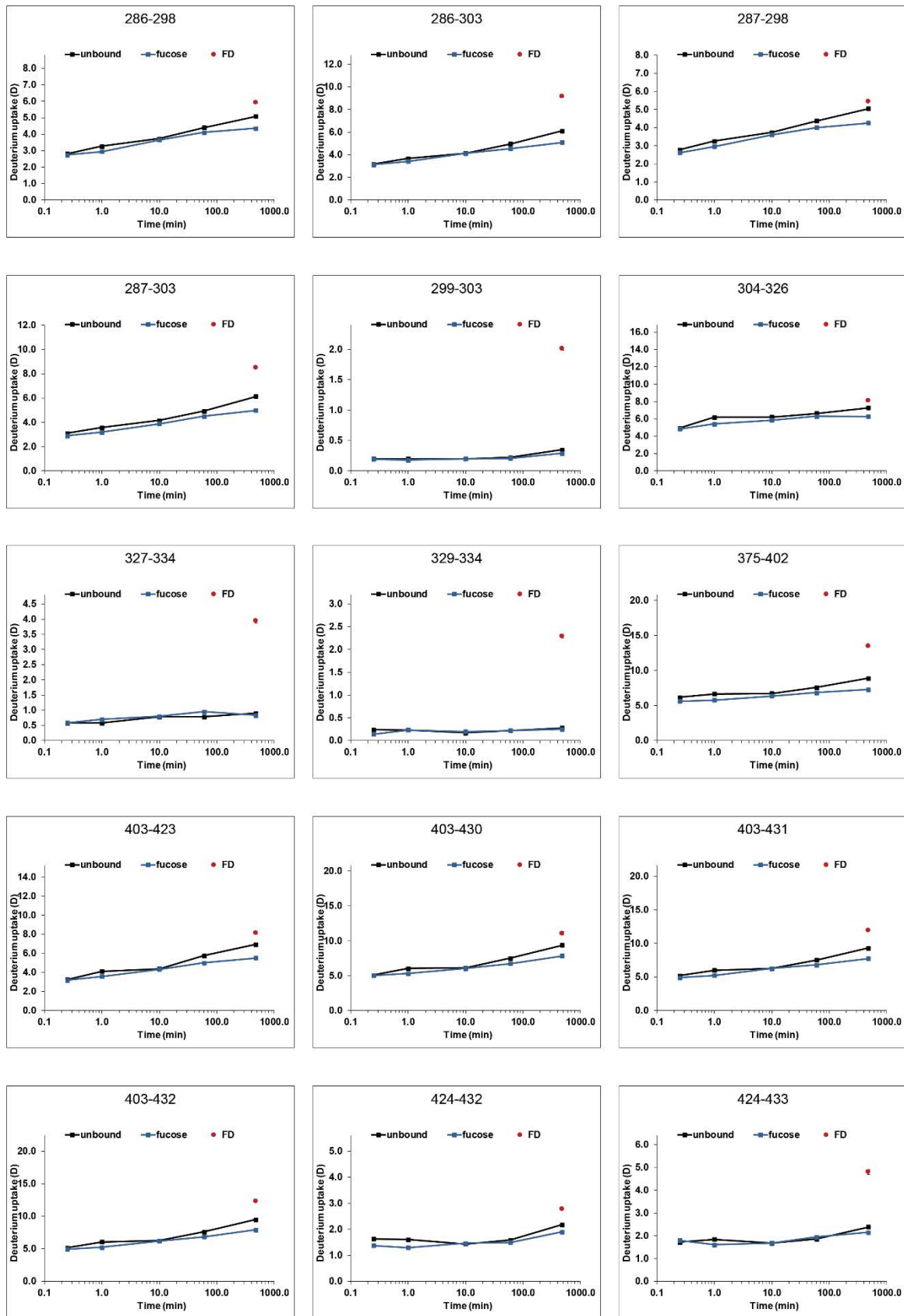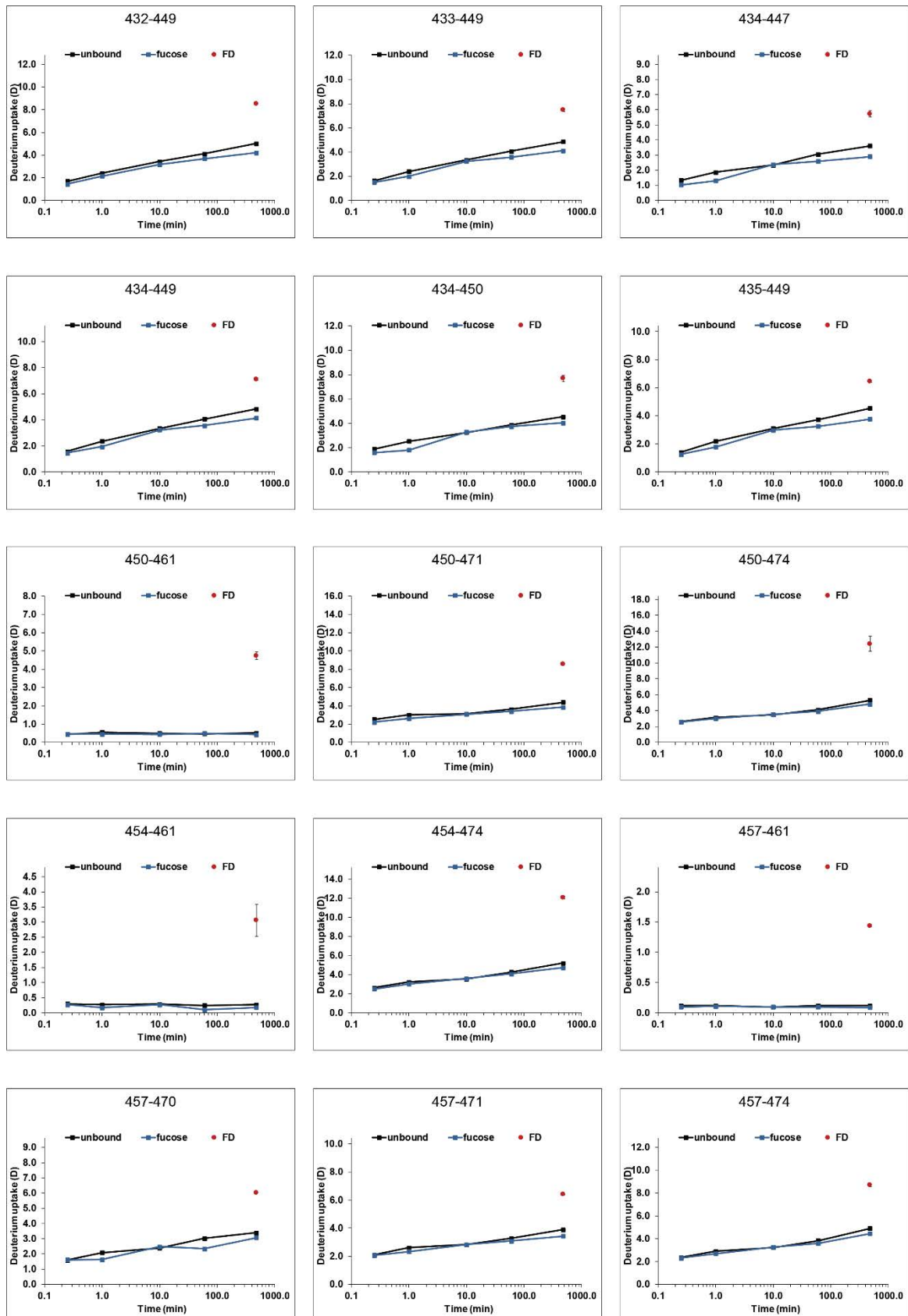
Figure S 43: continued
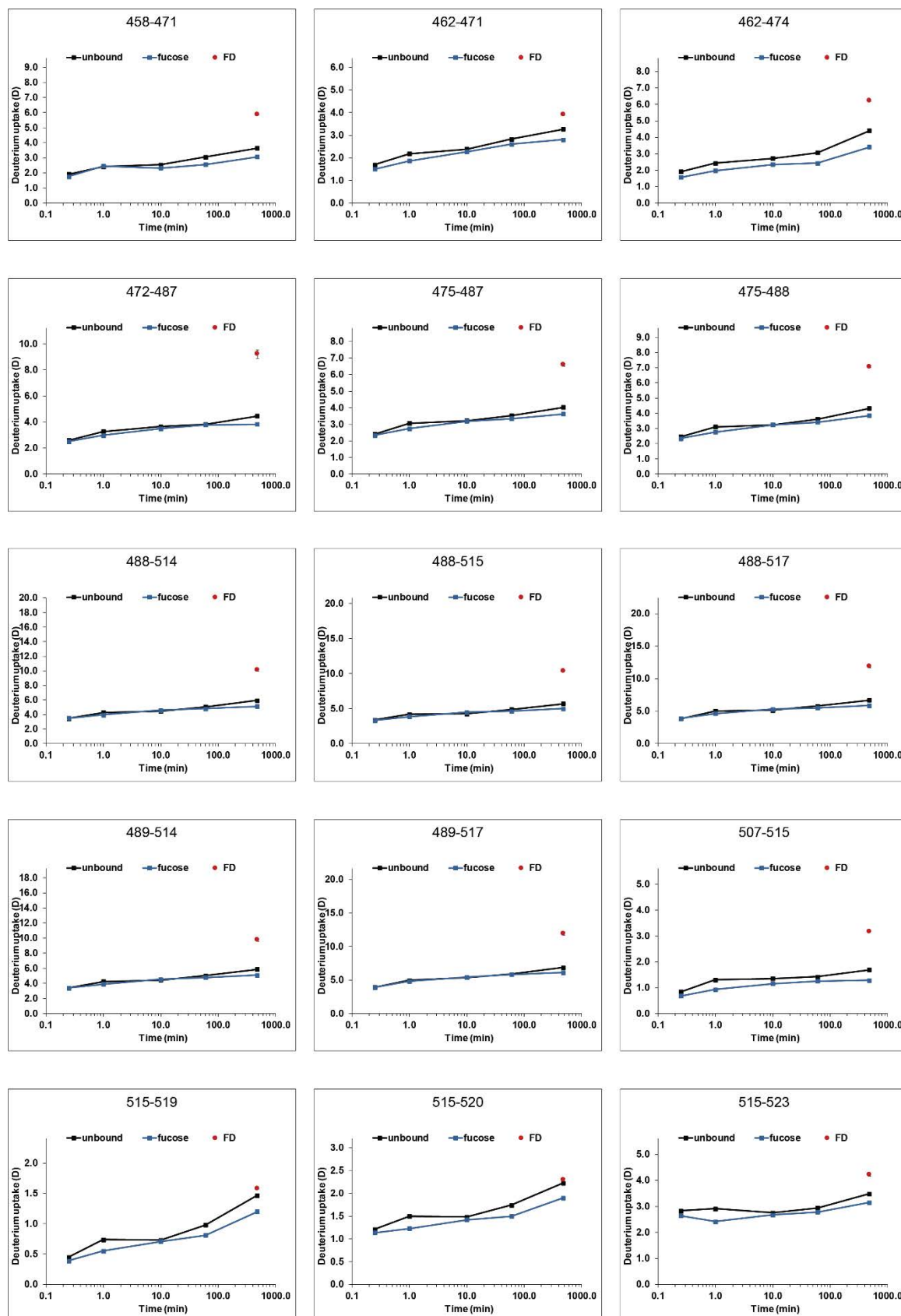
Figure S 43: continued
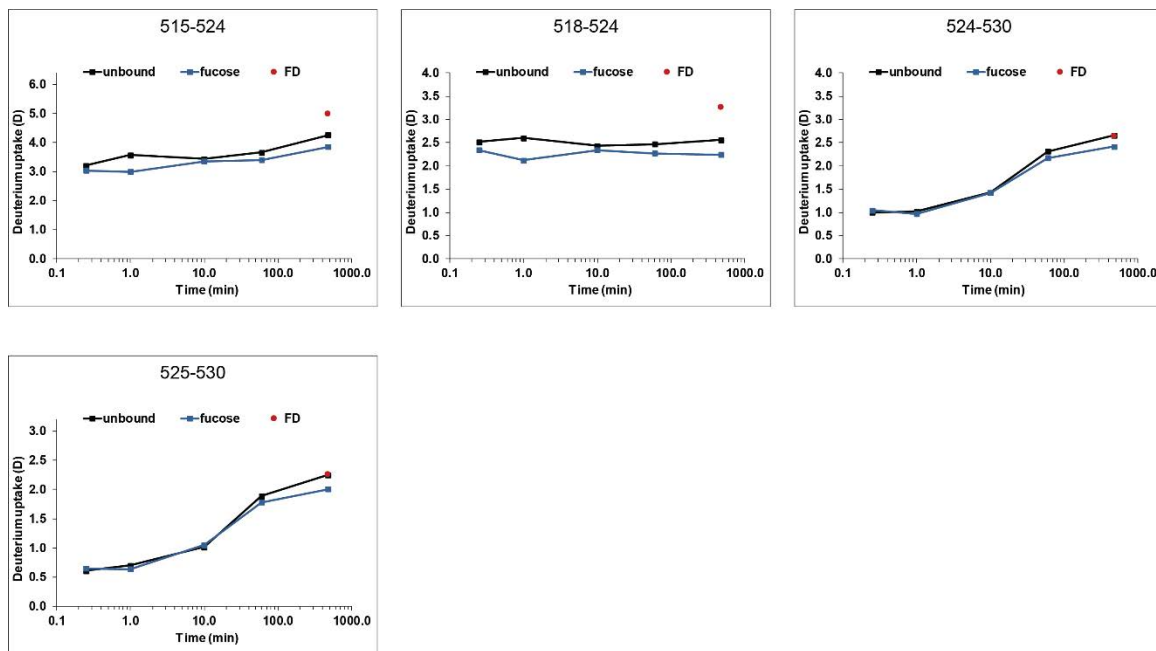
Figure S 43: continued

Figure S 43: continued
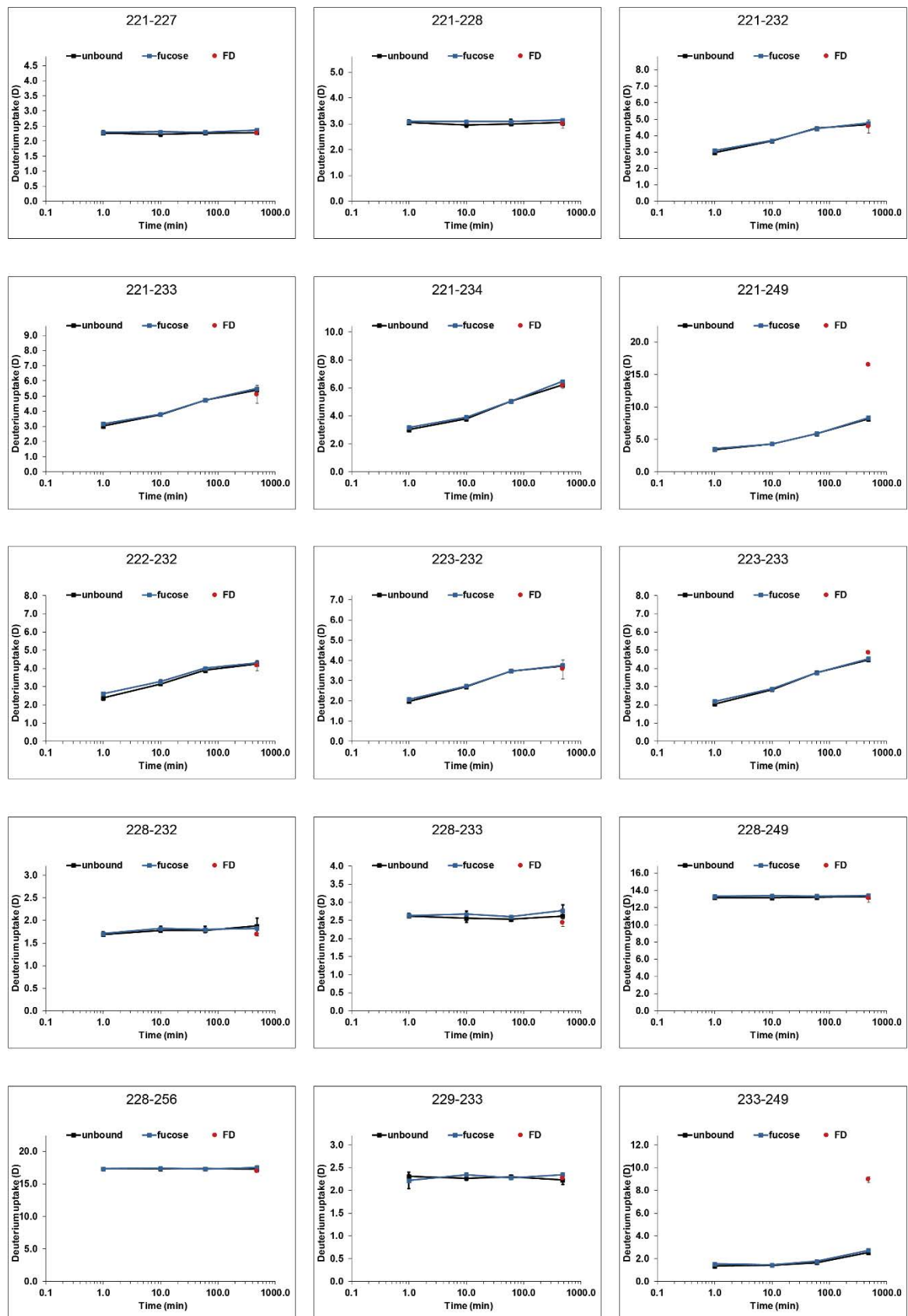
Figure S 44: GII.4 MI001 partially deamidated P dimer with 100 mM fucose
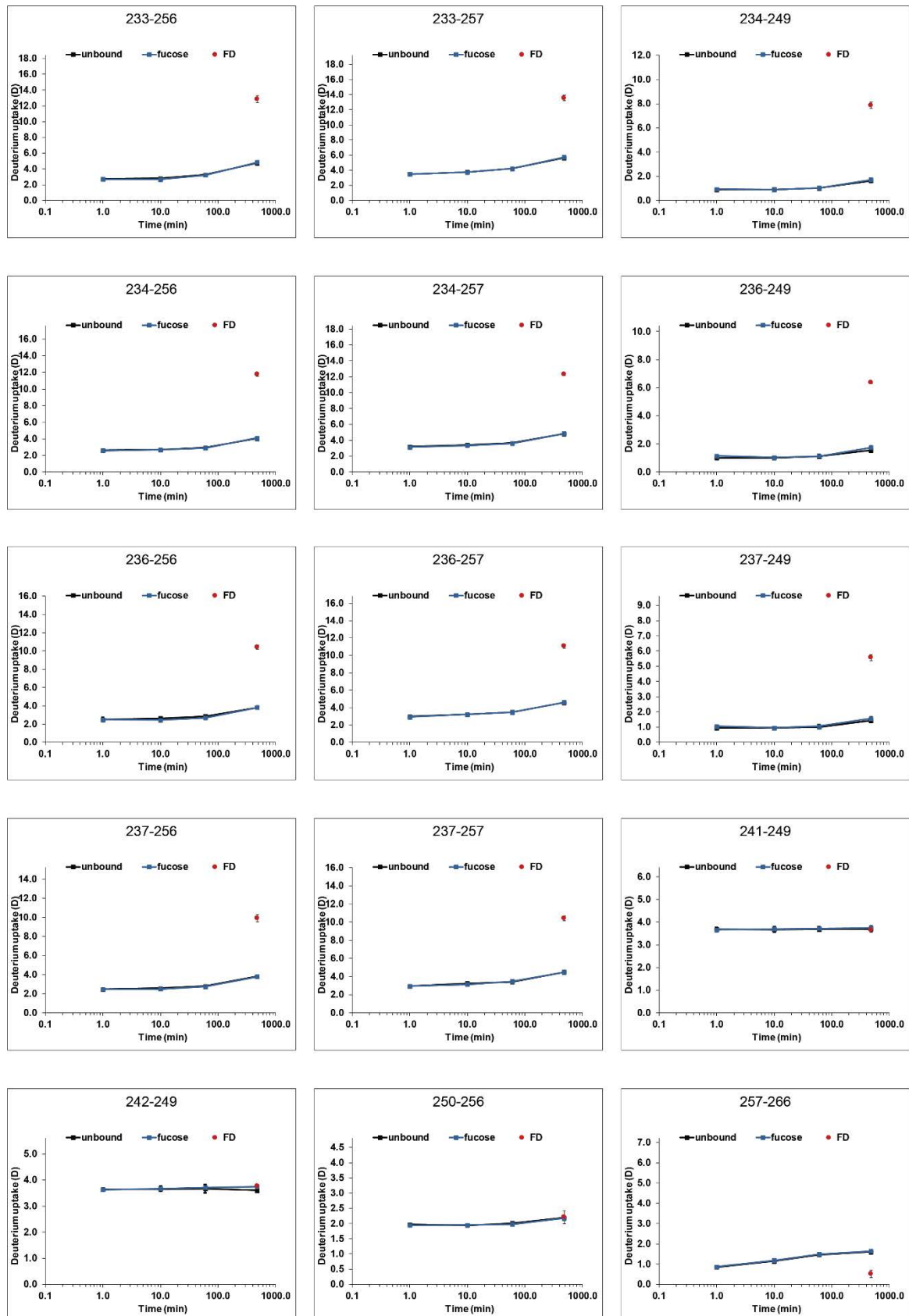
Figure S 44: continued
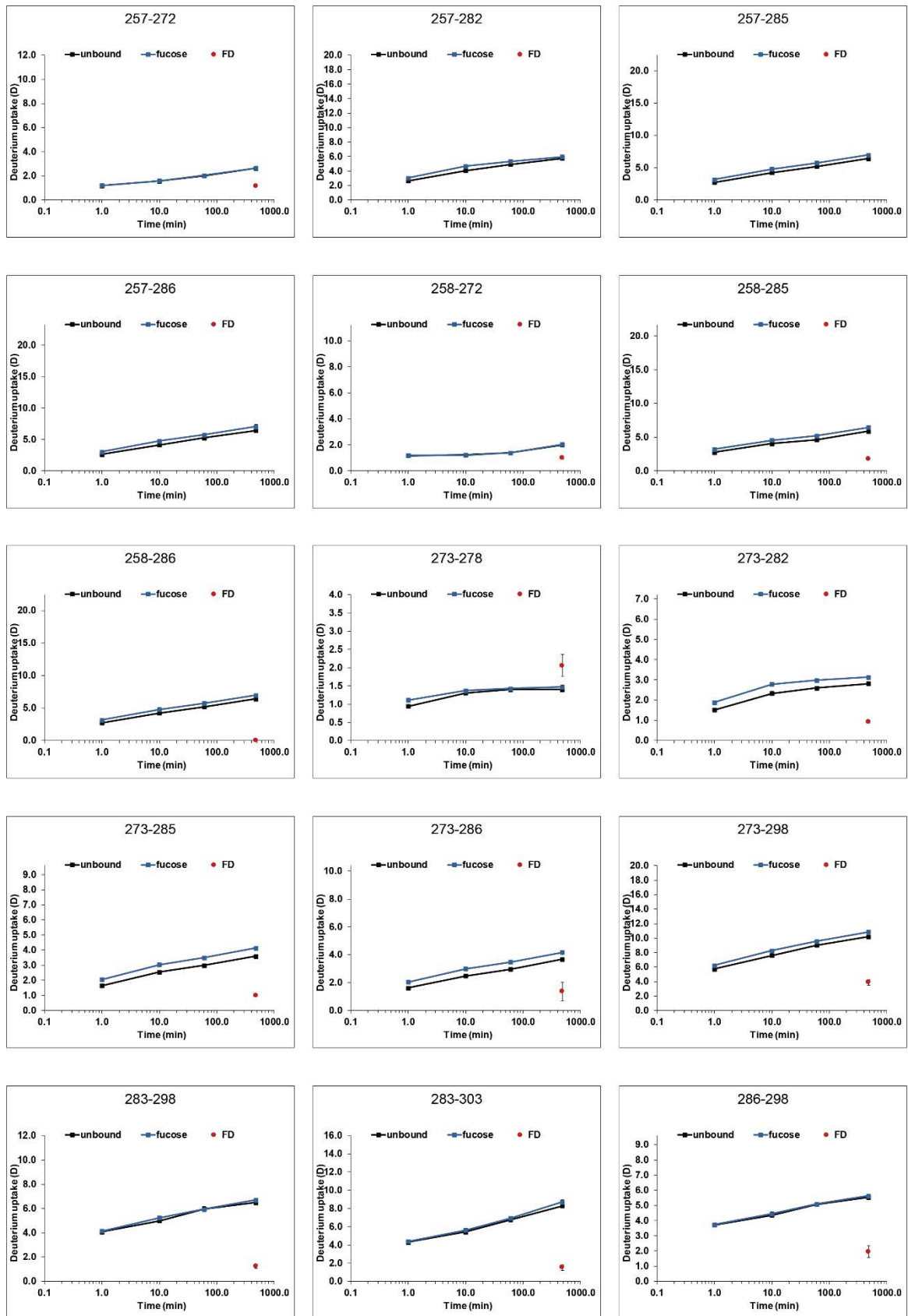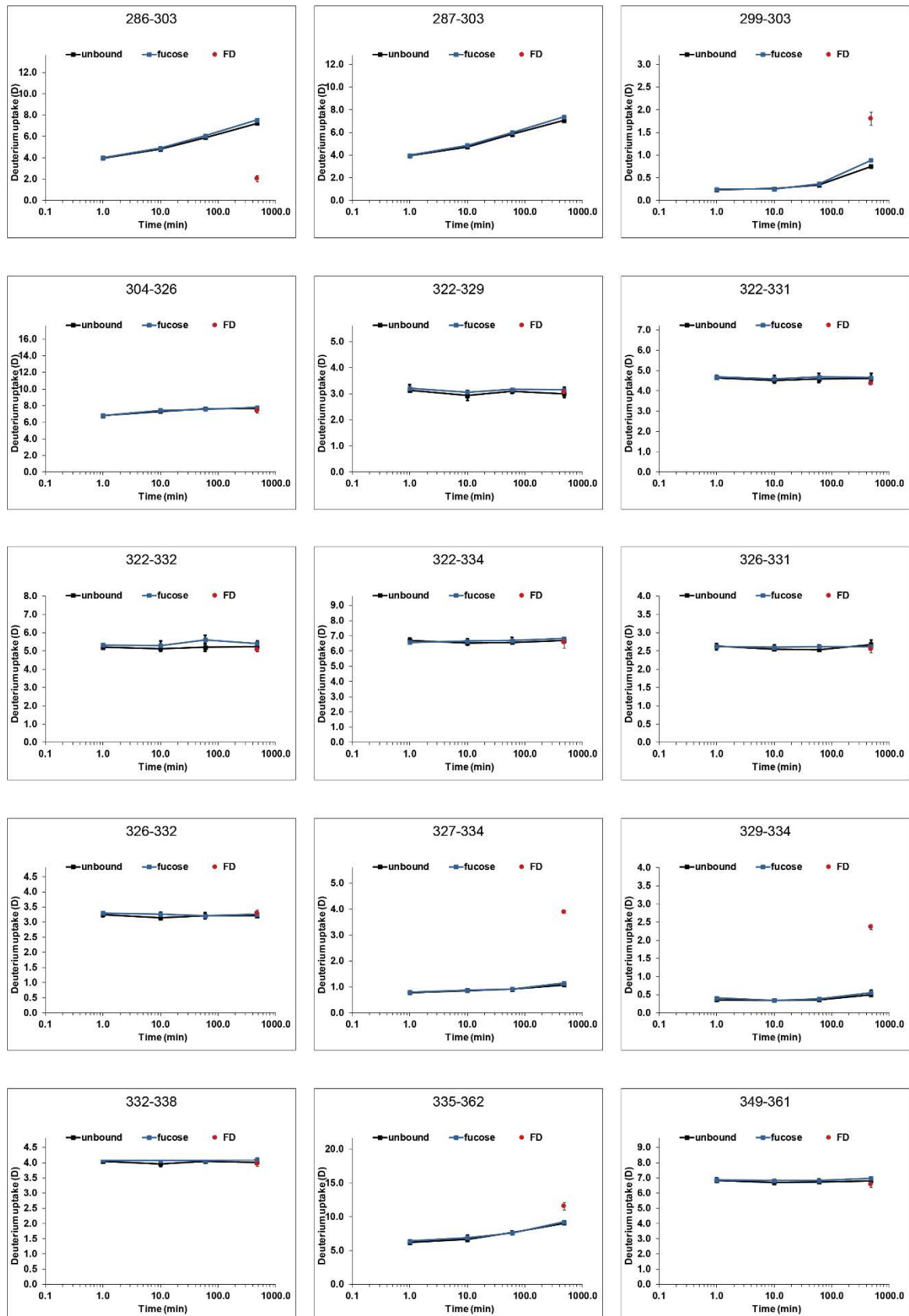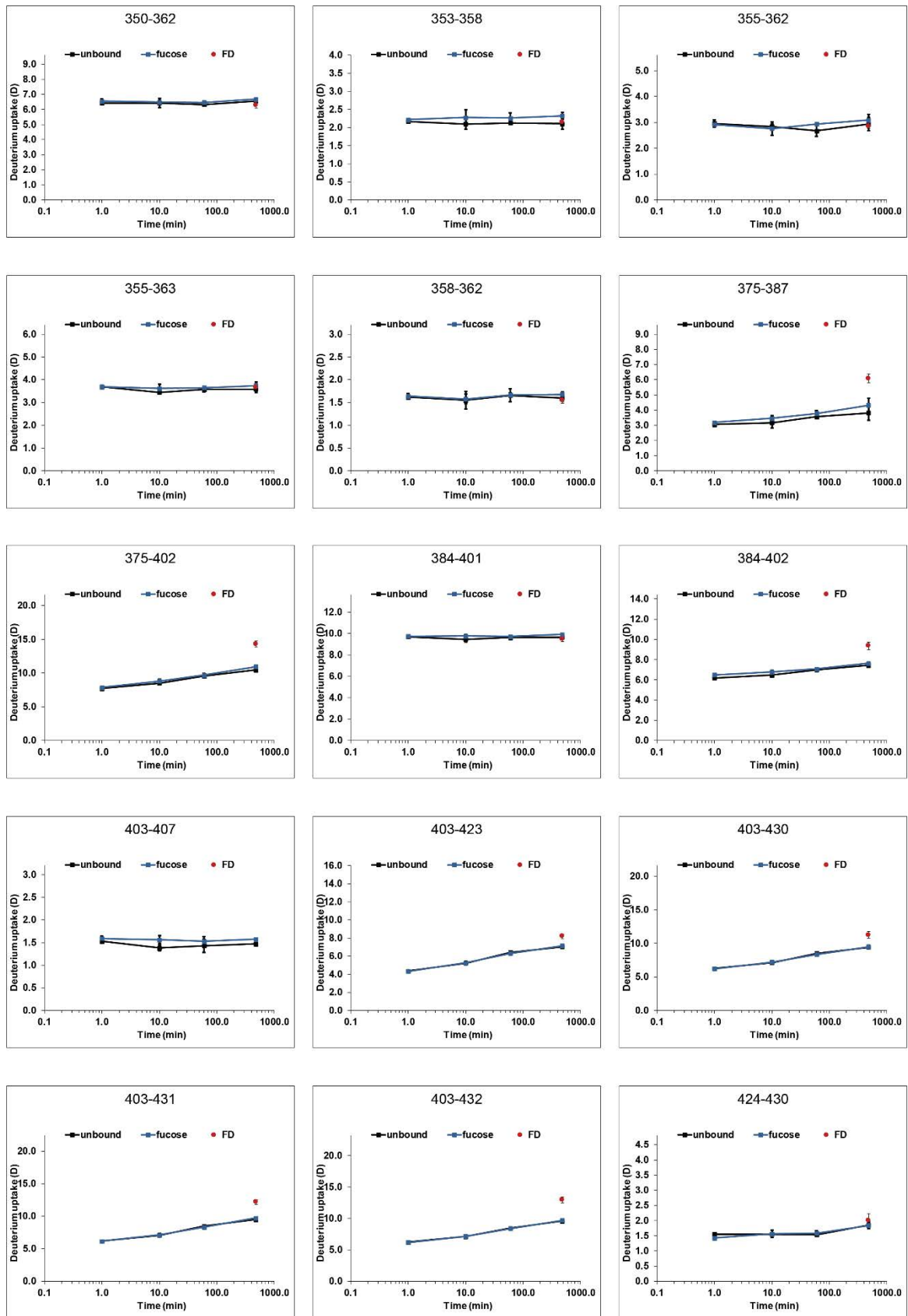
Figure S 44: continued
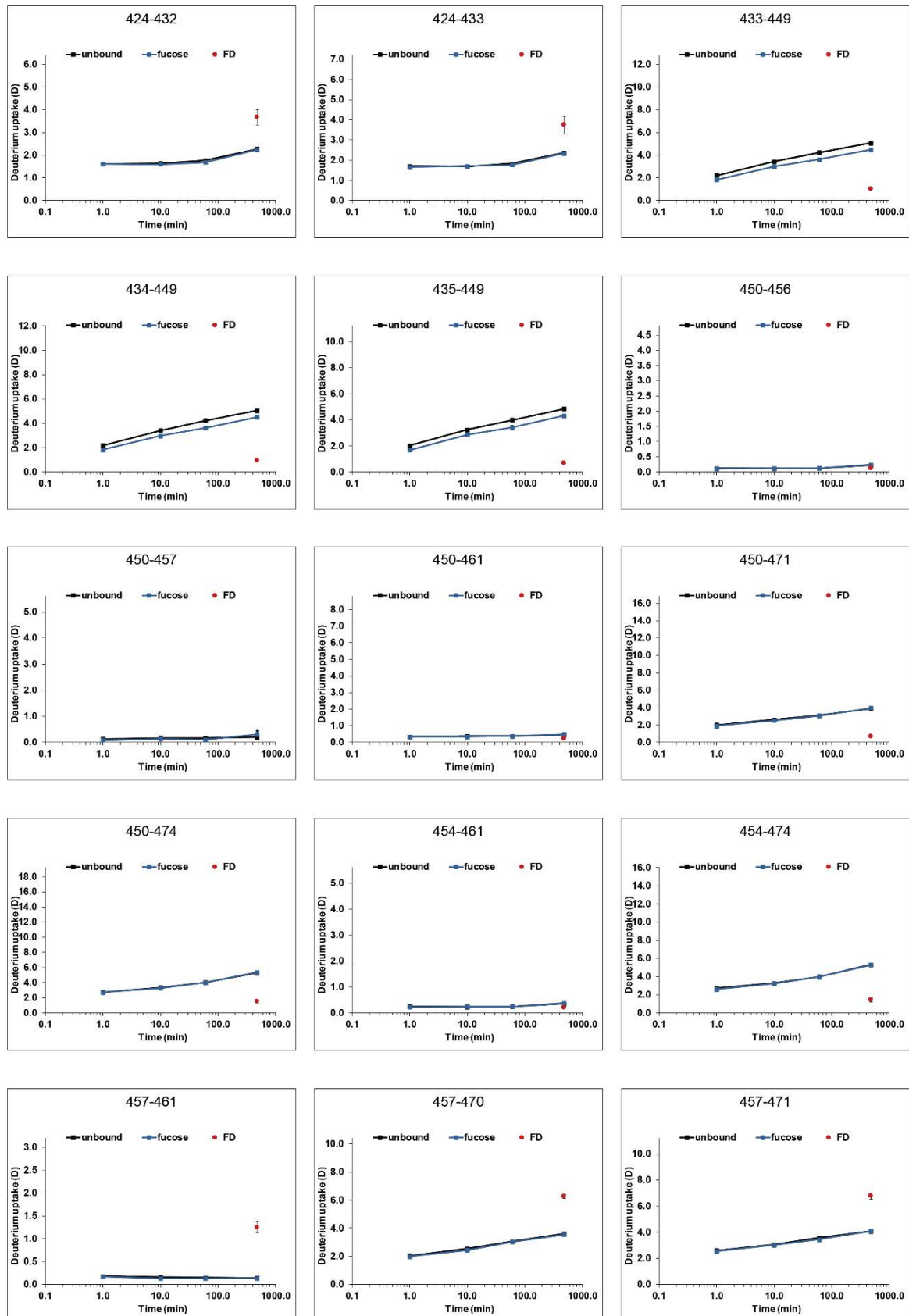
Figure S 44: continued

Figure S 44: continued
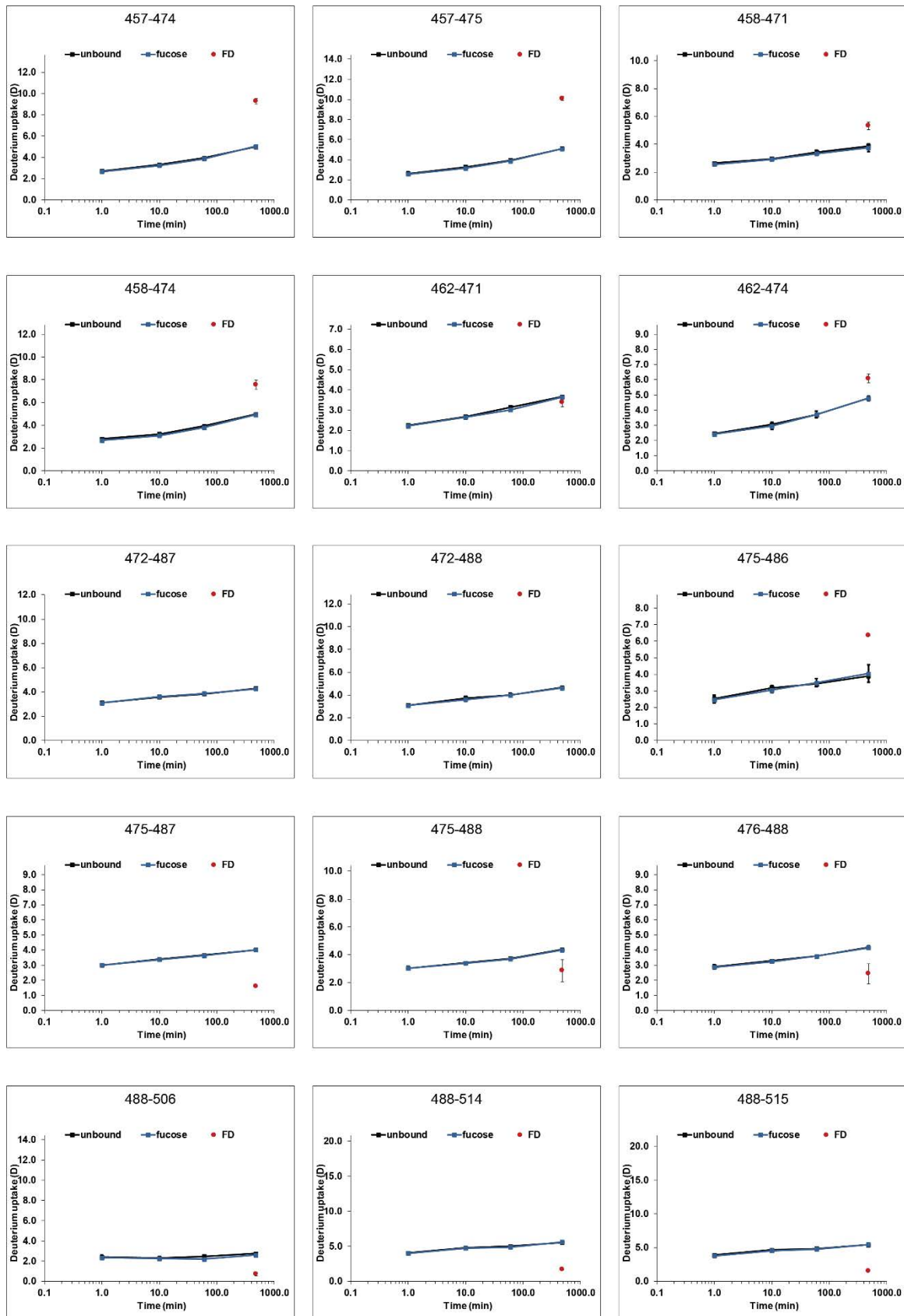
Figure S 44: continued
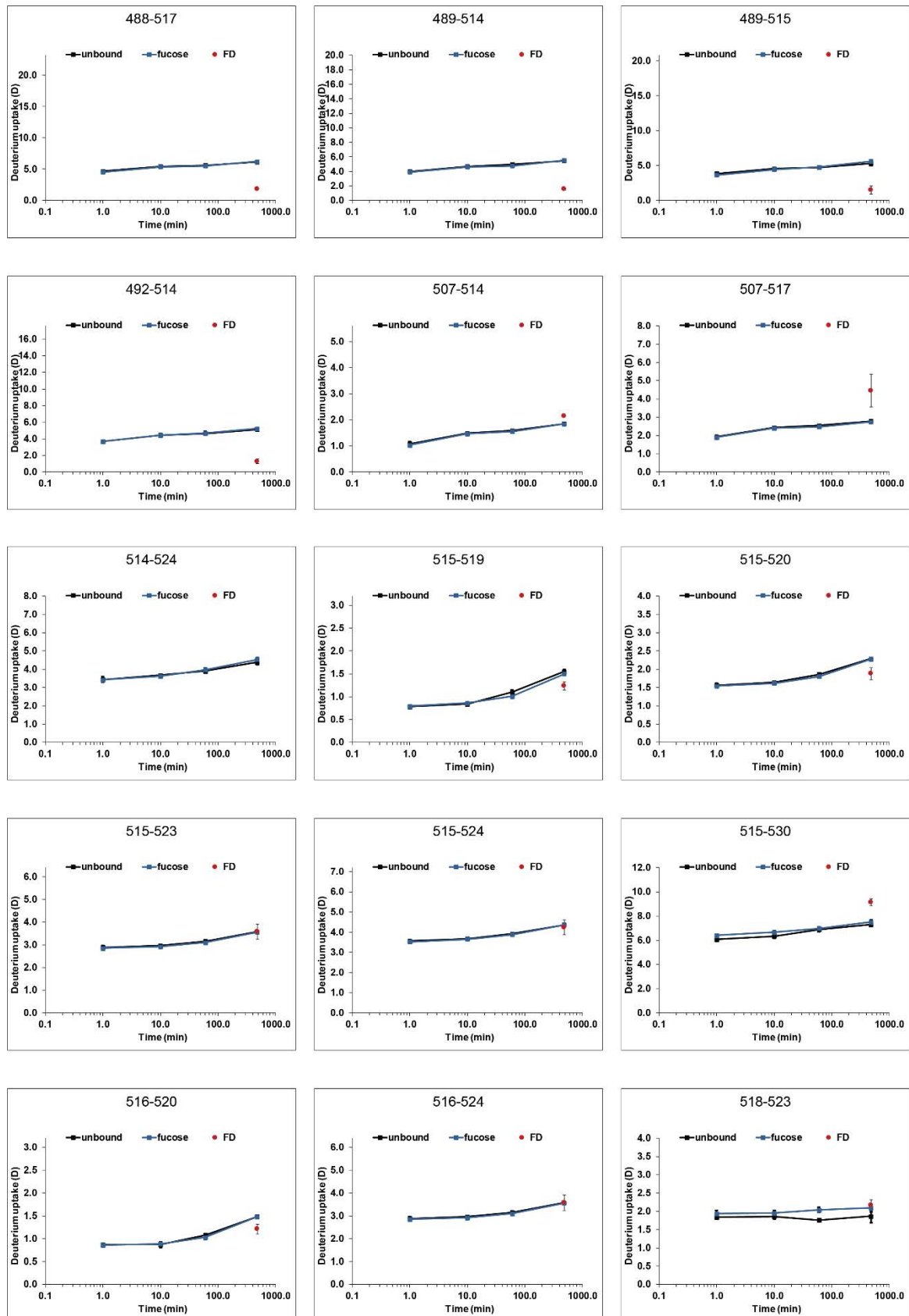
Figure S 44: continued

Figure S 44: continued
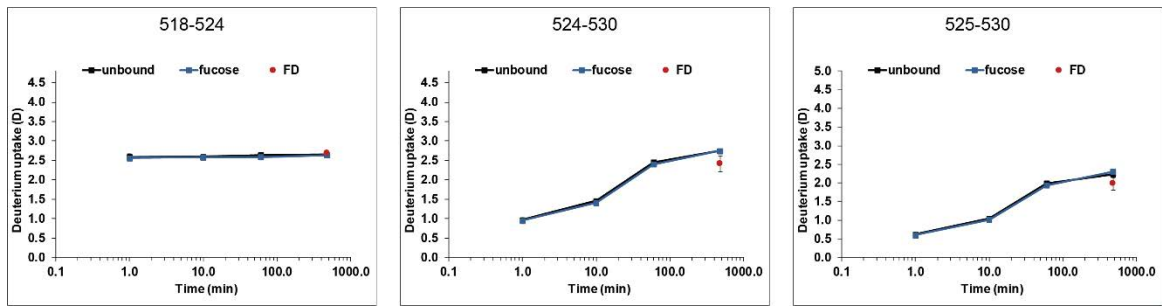
Figure S 44: continued
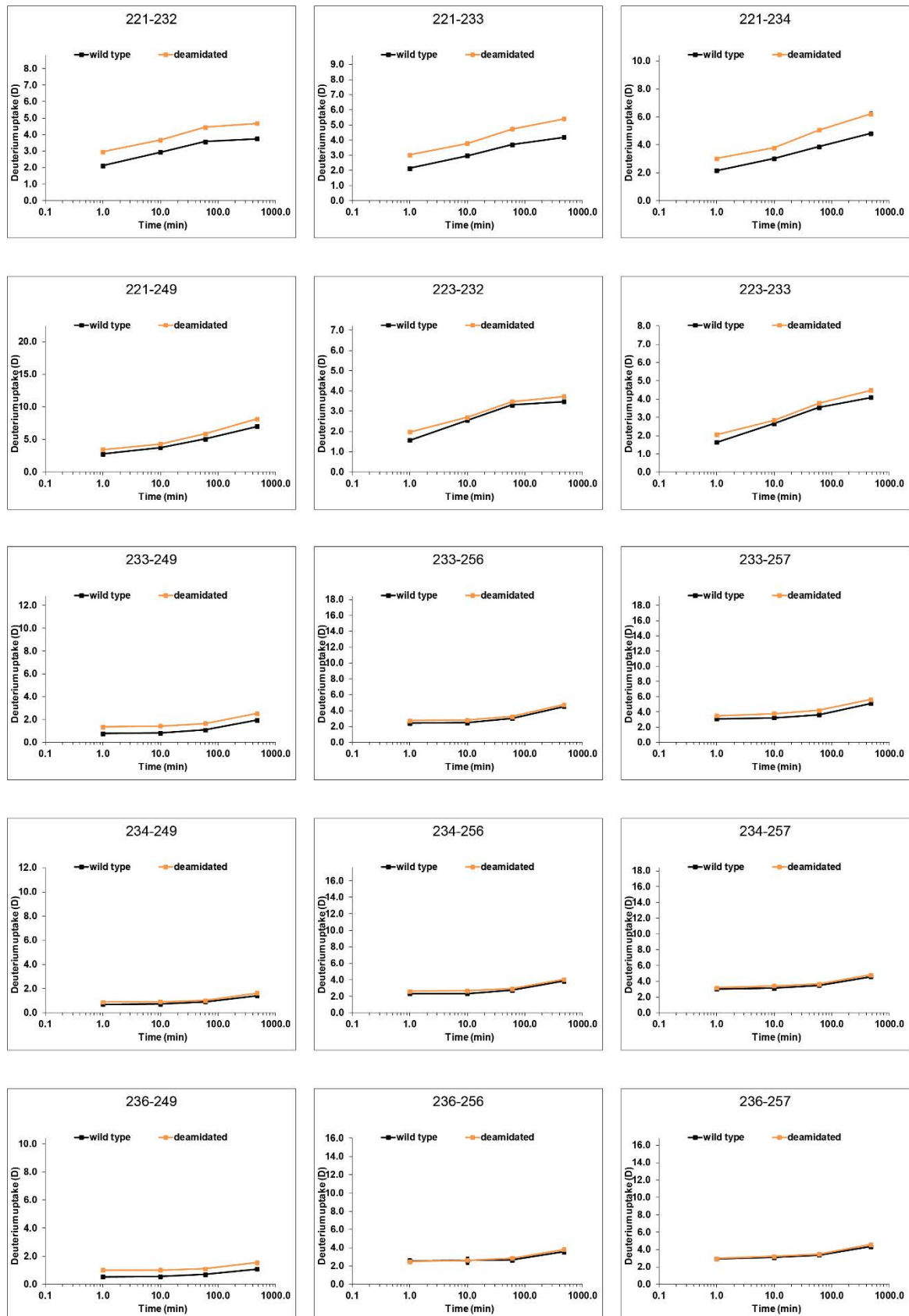
Figure S 45: GII.4 MI001 wild type versus partially deamidated P dimer
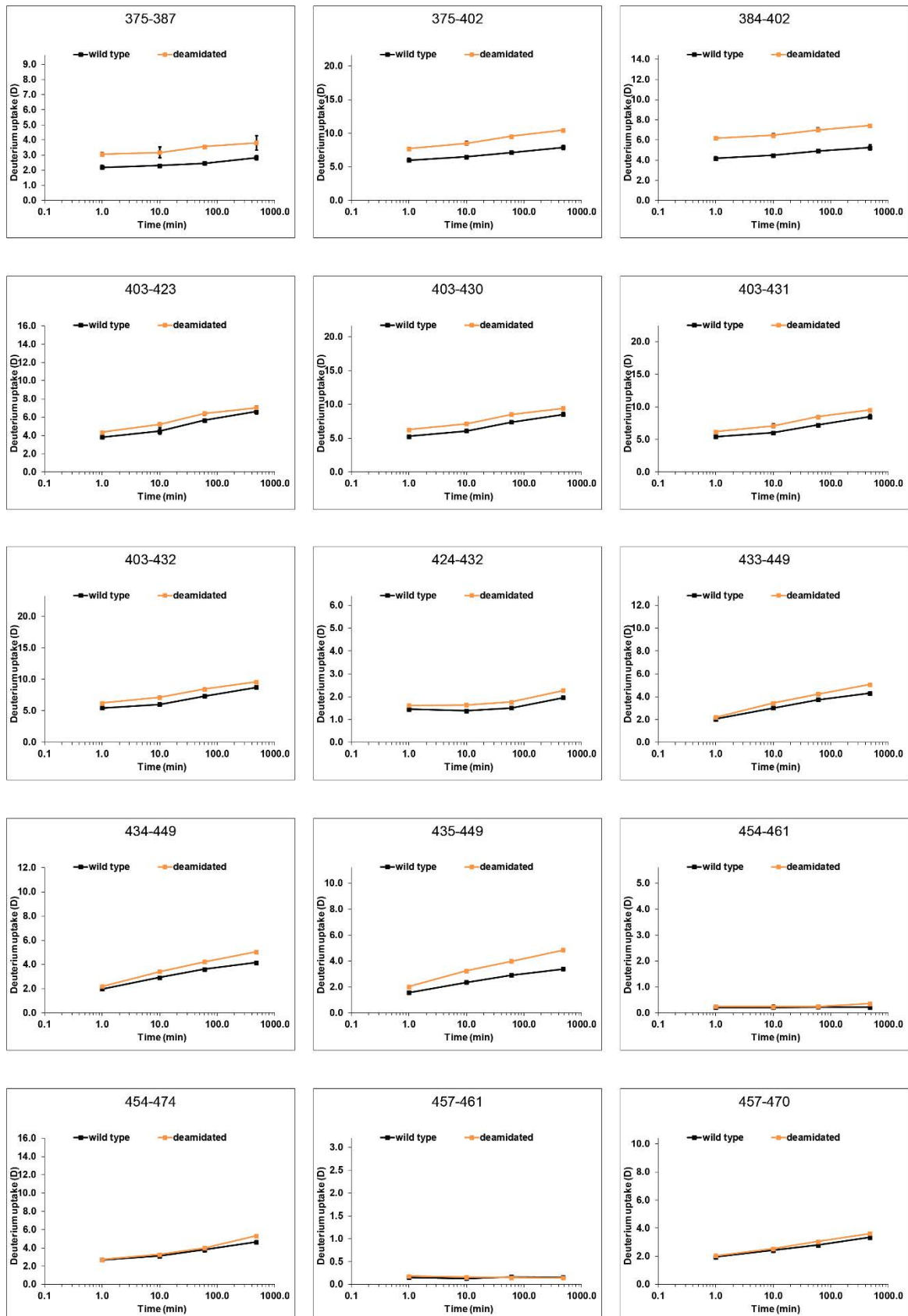
Figure S 45: continued
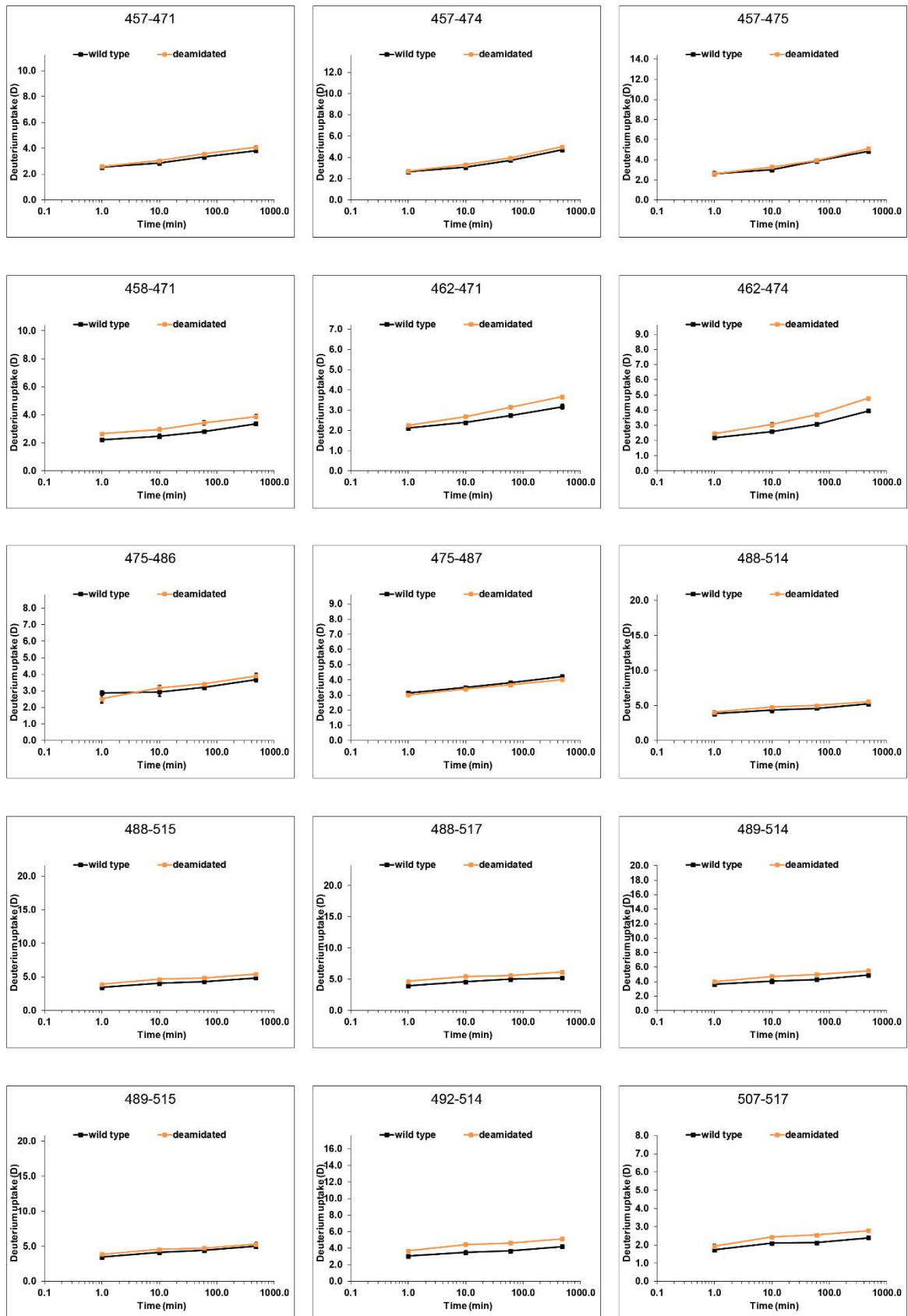
Figure S 45: continued
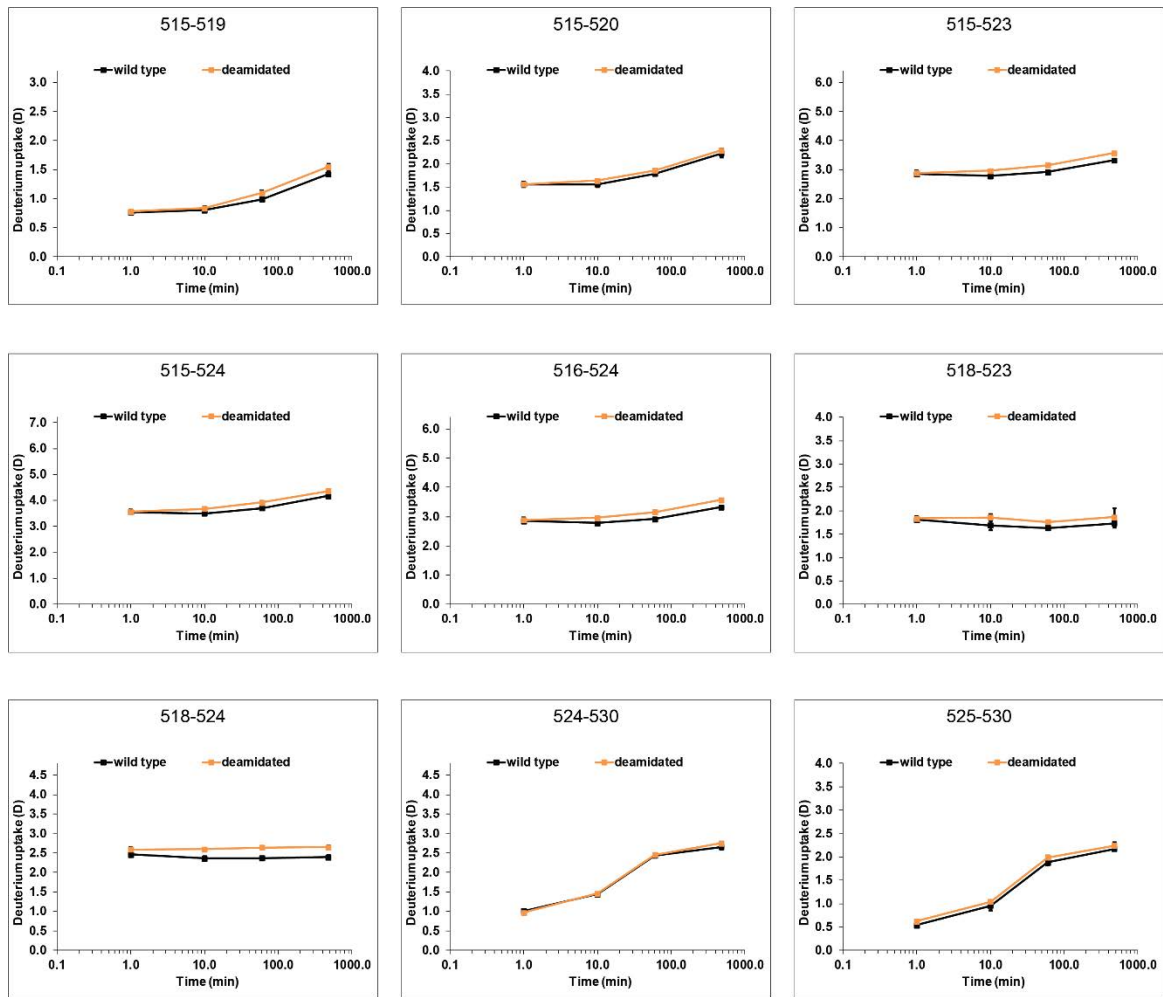
Figure S 45: continued

Figure S 45: continued
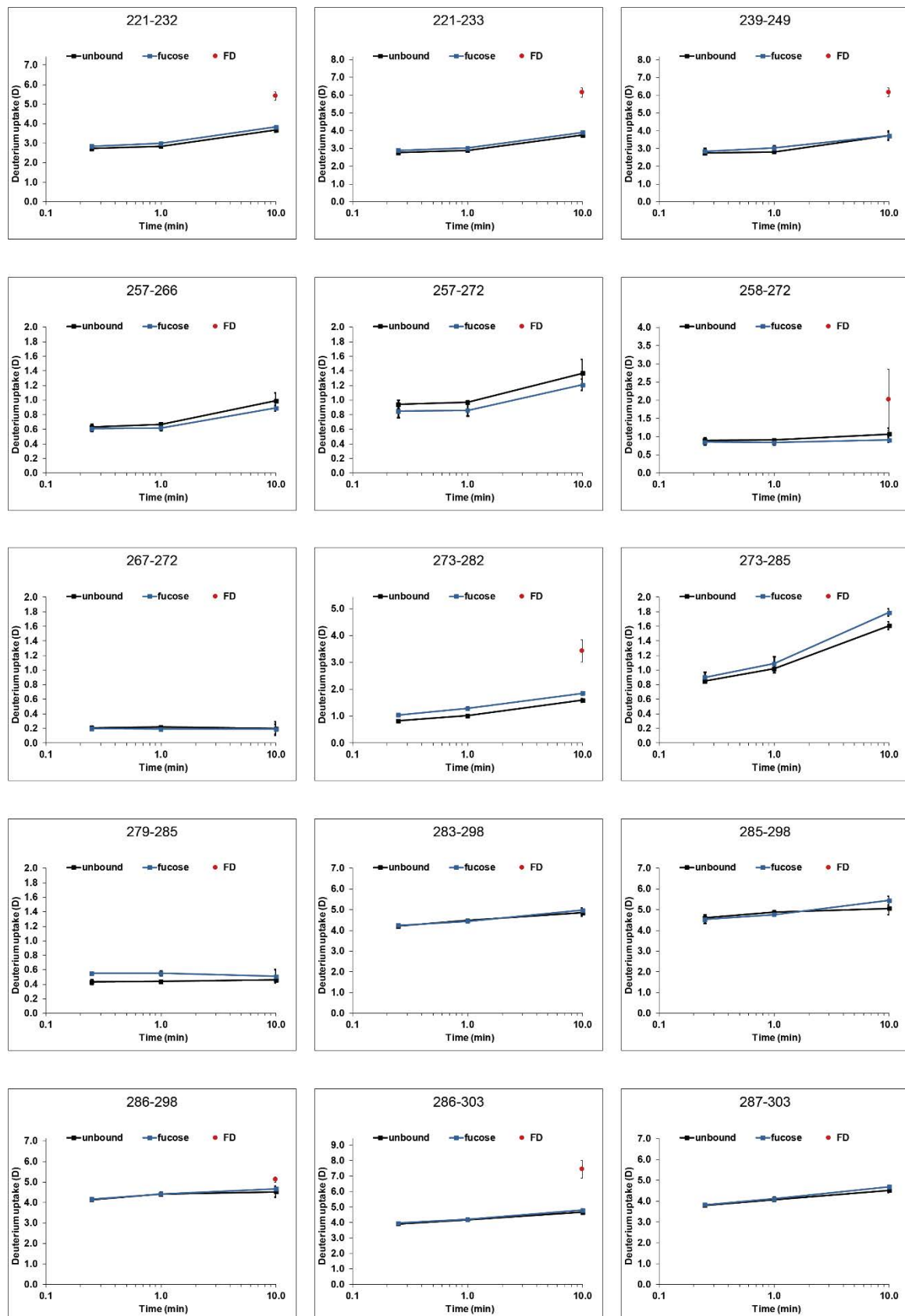
Figure S 46: GII.4 Saga partially deamidated P dimer with 100 mM fucose
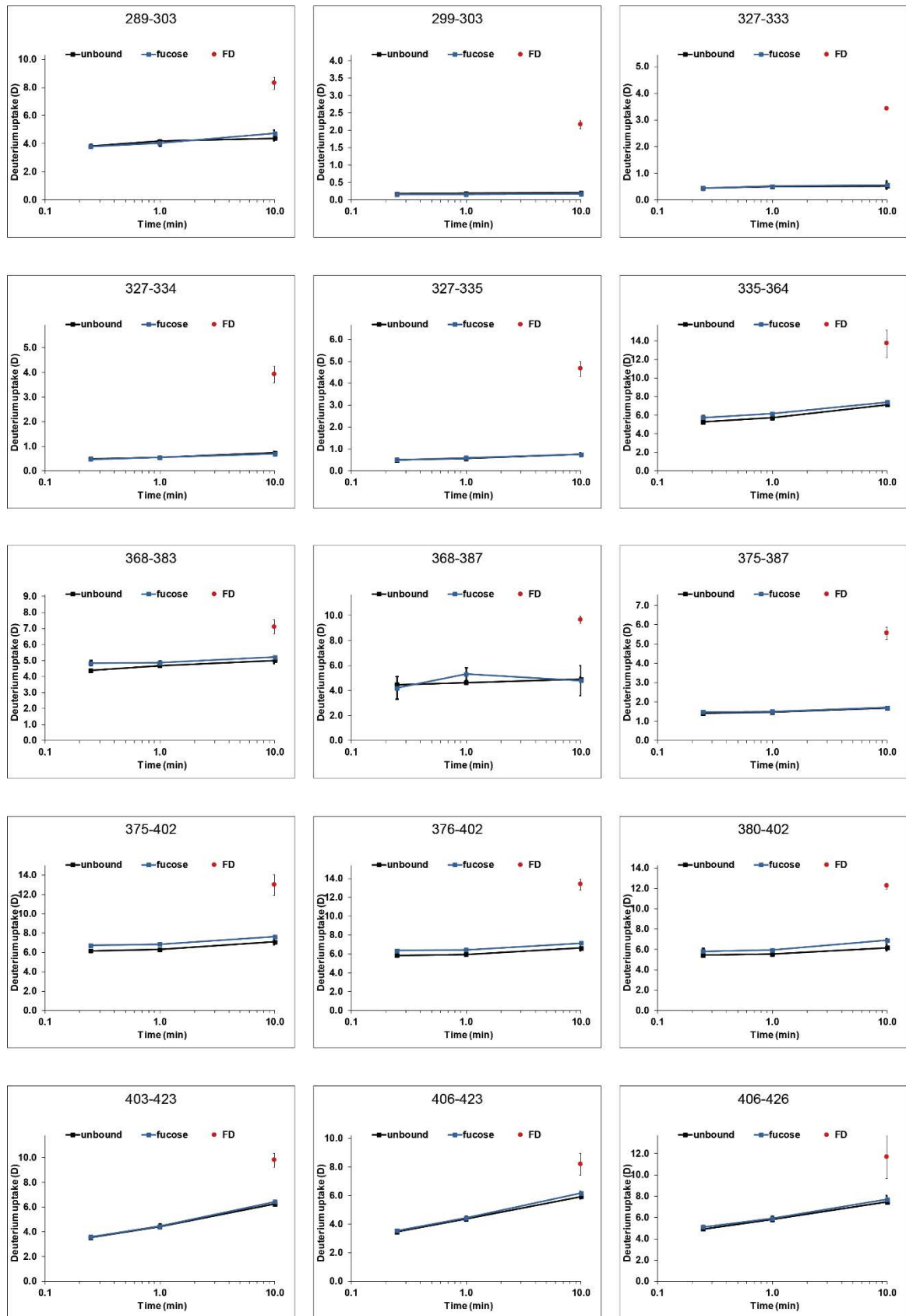
Figure S 46: continued
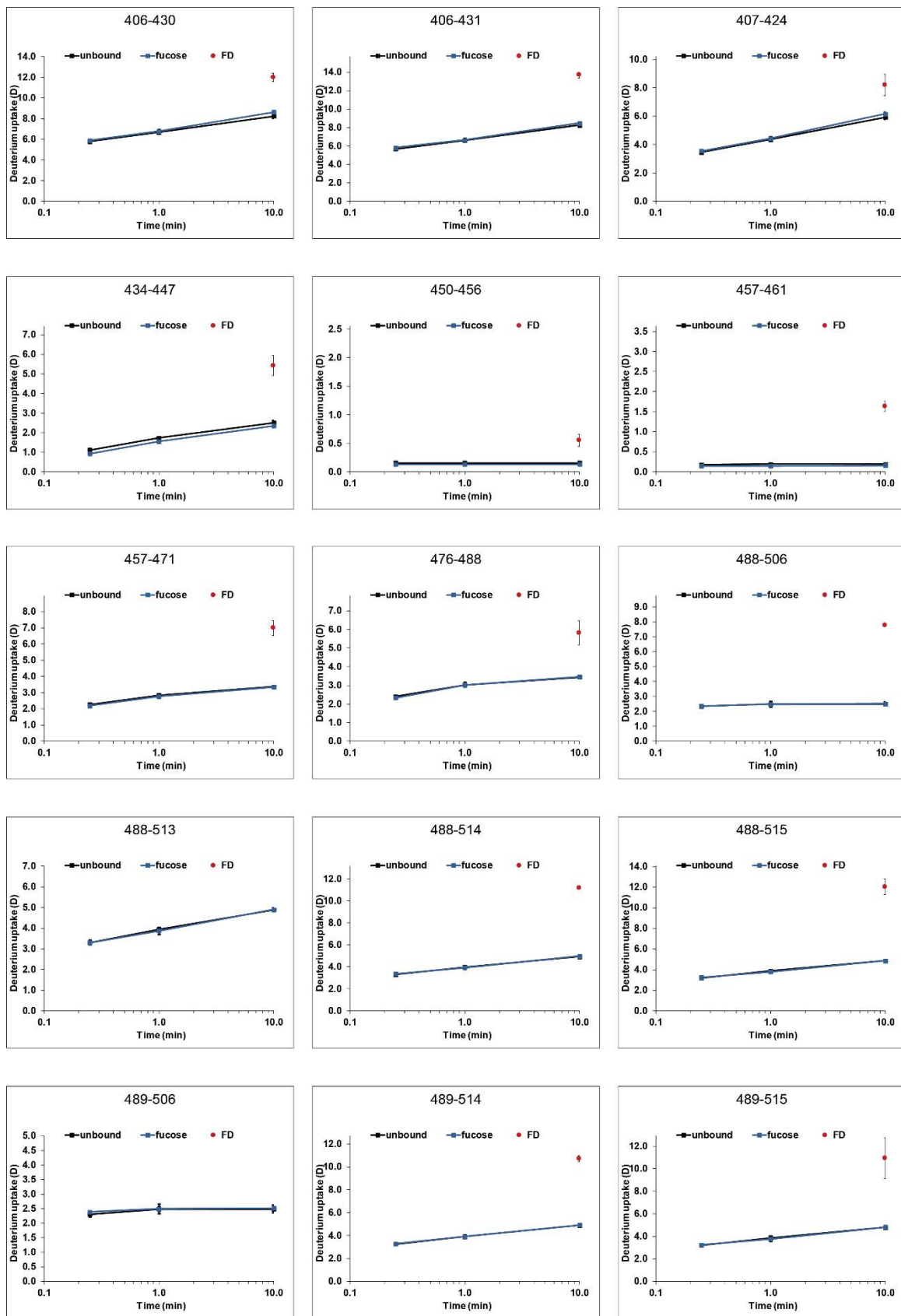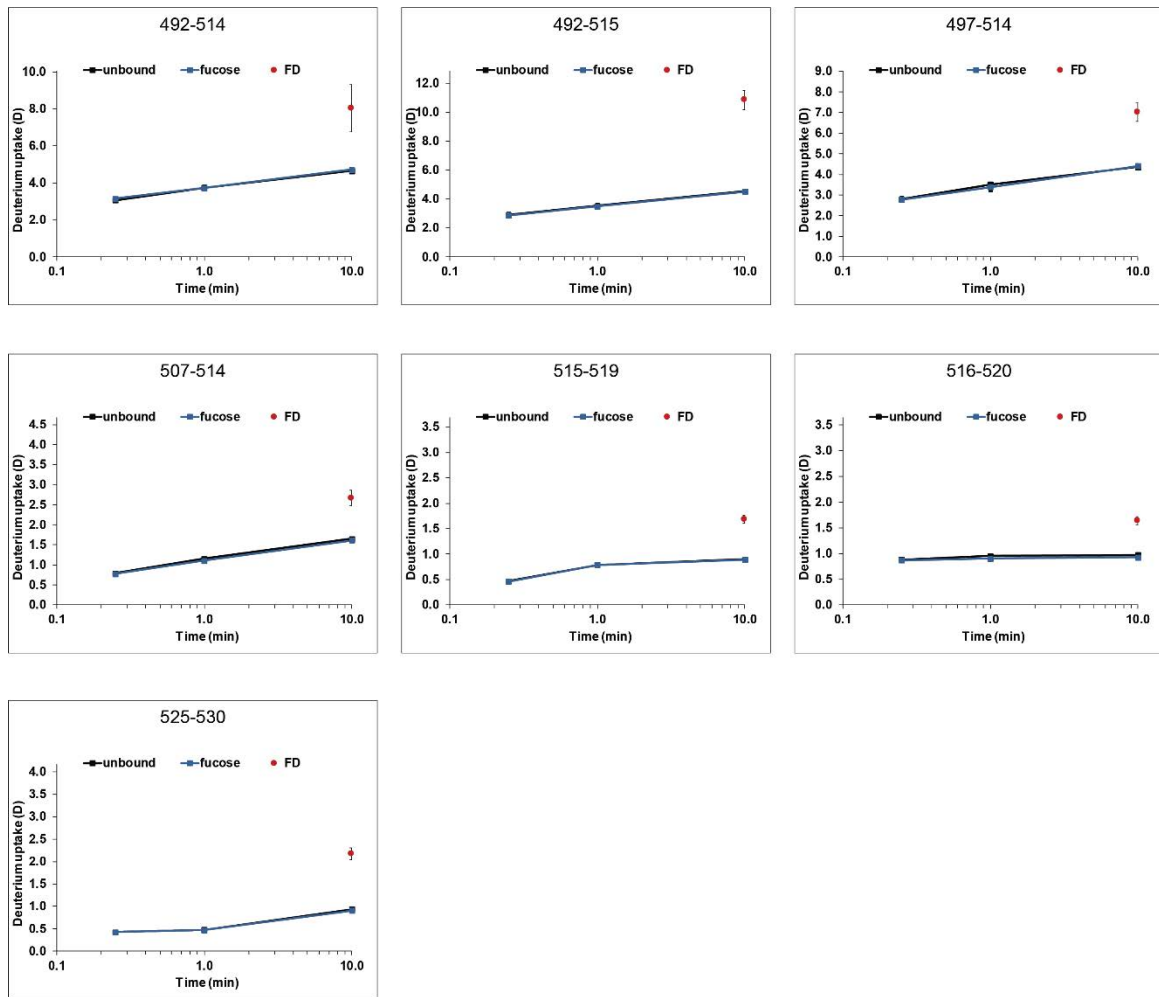
Figure S 46: continued

Figure S 46: continued
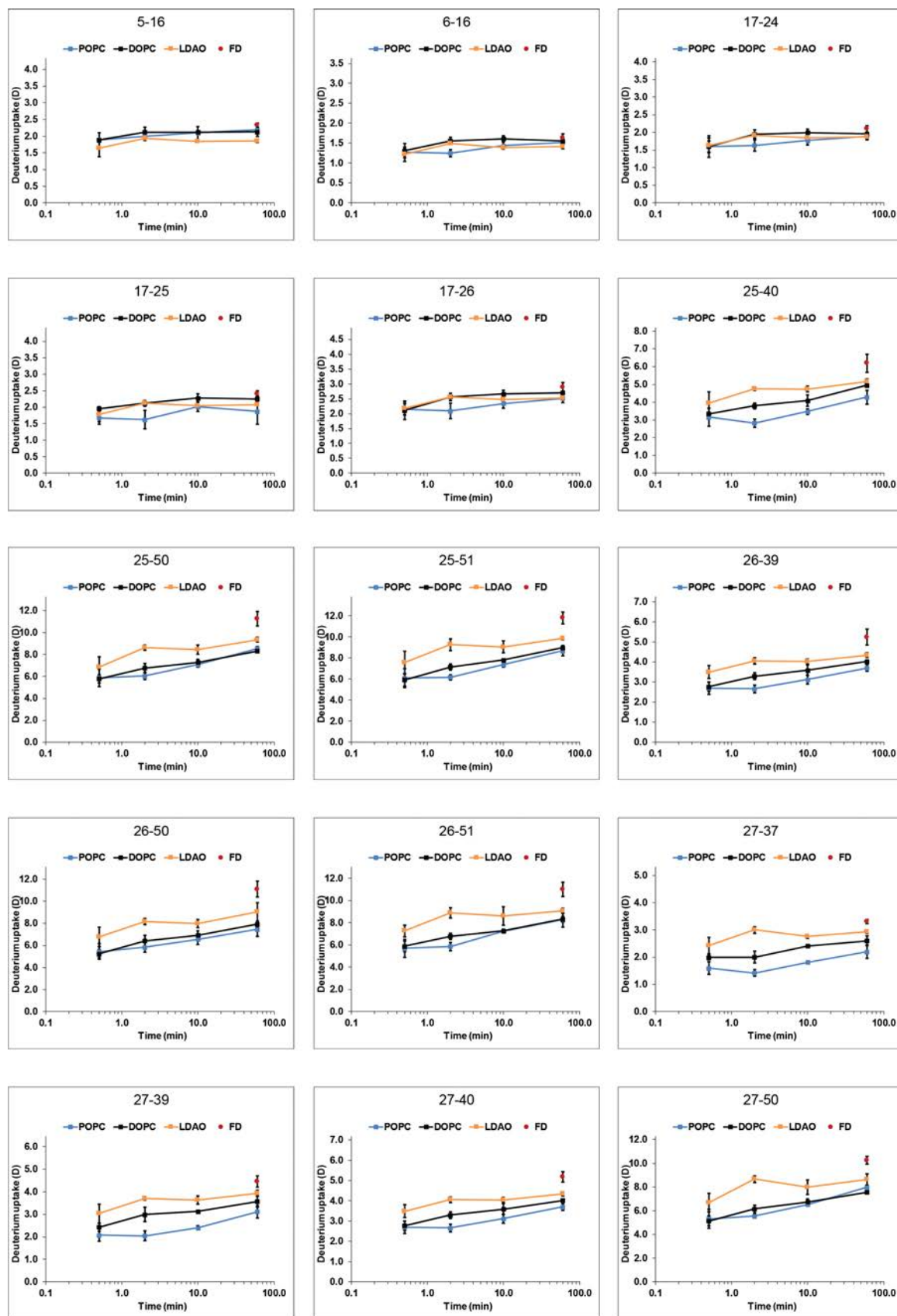
Figure S 47: Mistic in POPC, DOPC and LDAO
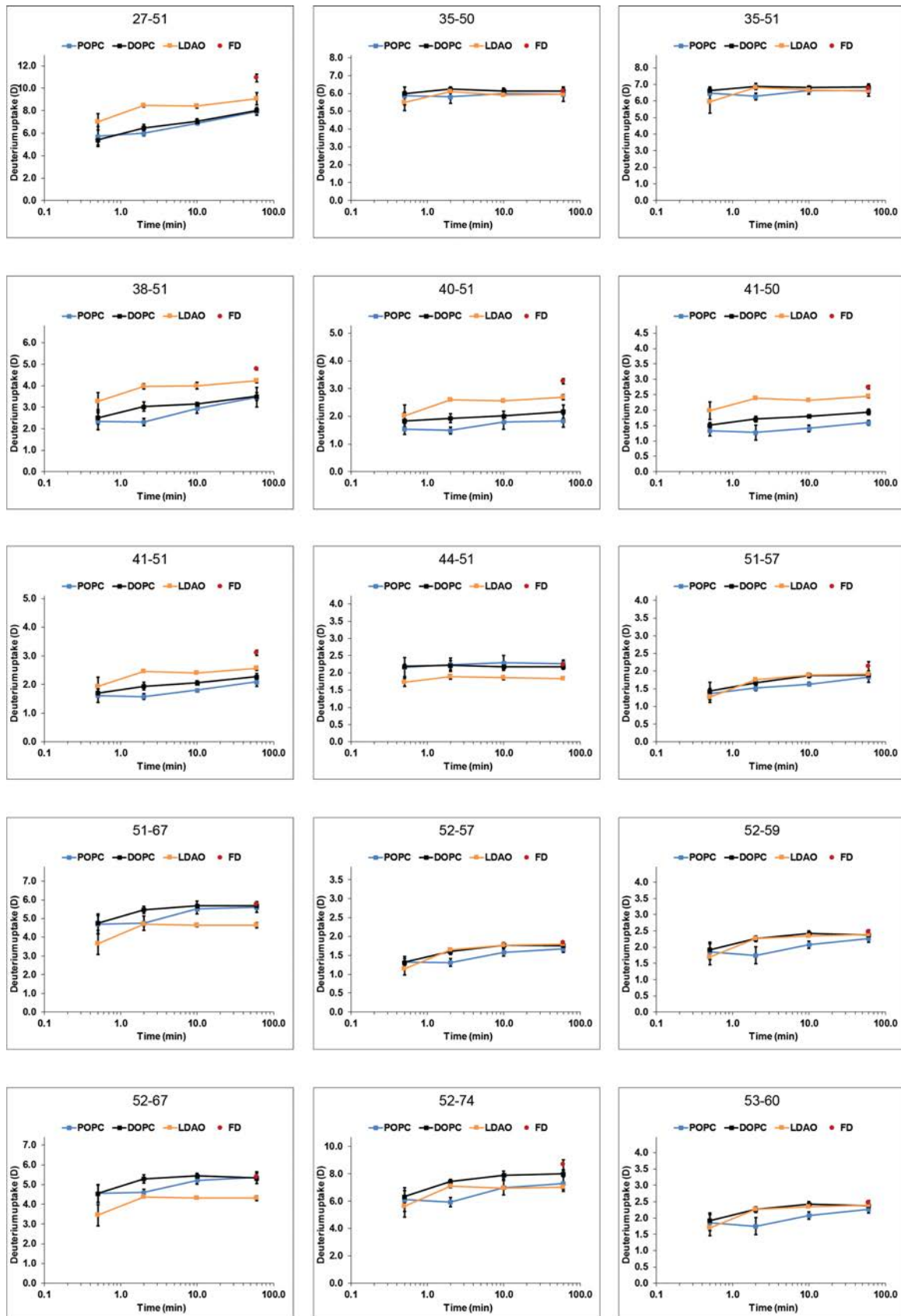
Figure S 47: continued

Figure S 47: continued

Figure S 47: continued
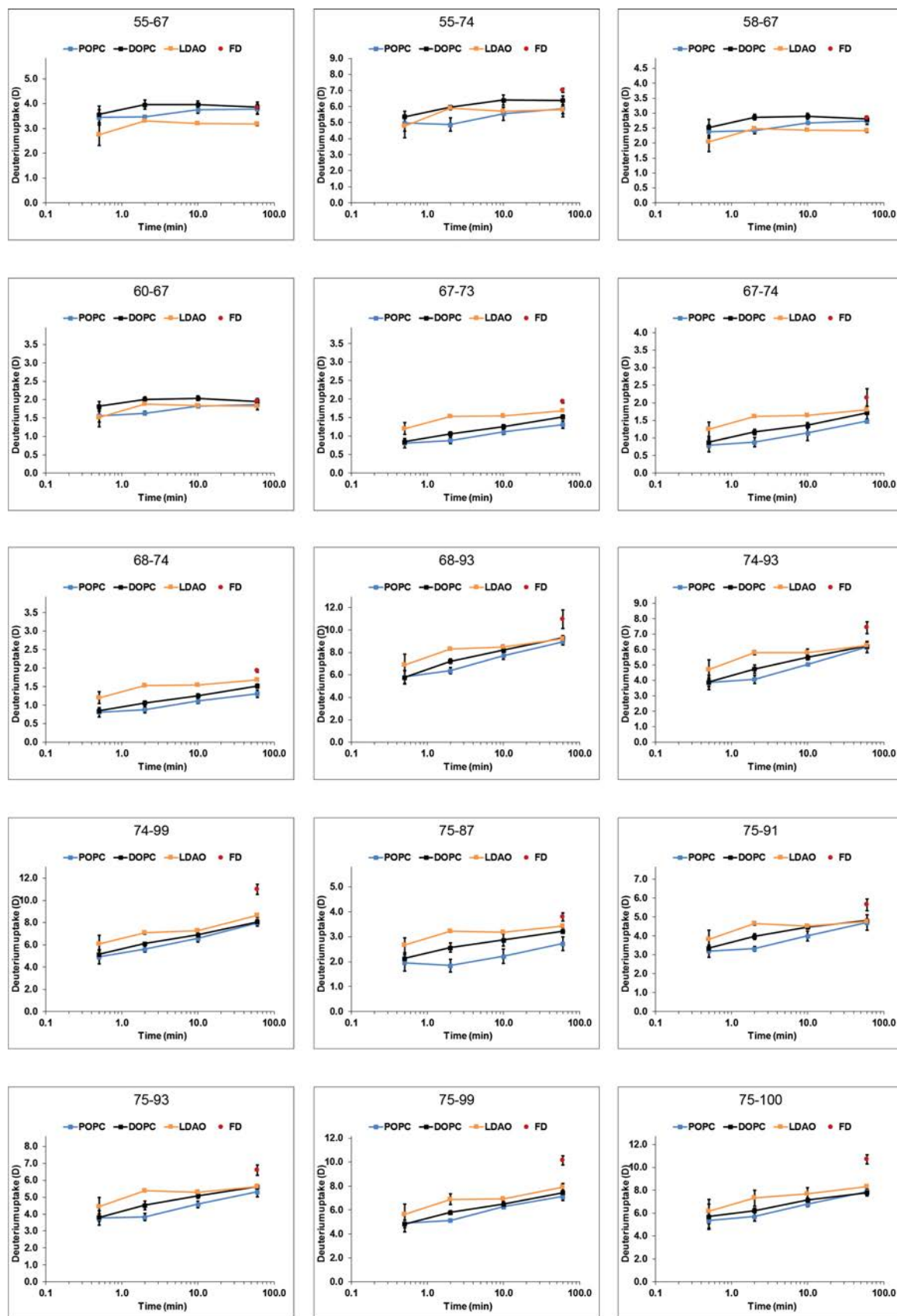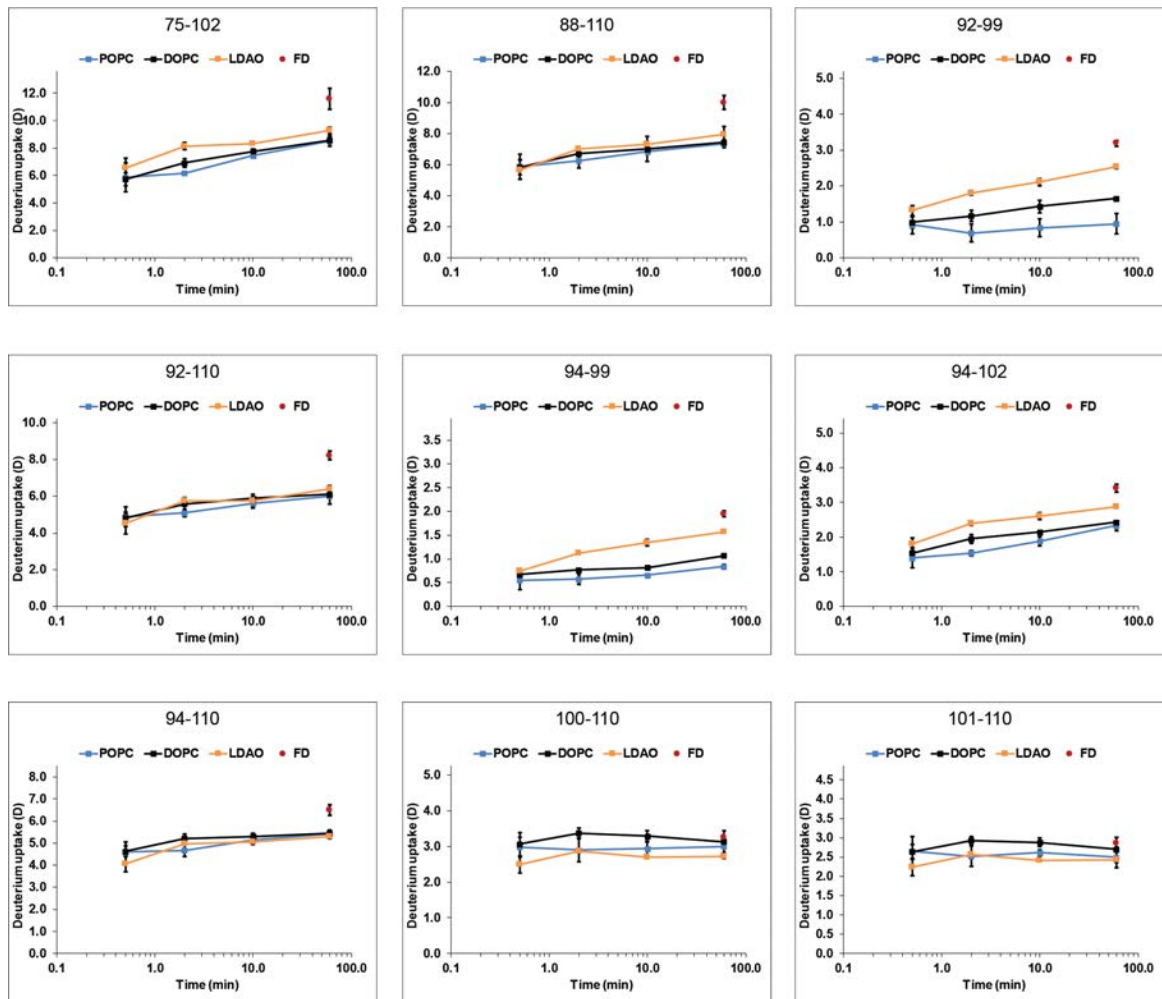
## 6.4 Protein and peptide sequences

**Human norovirus GII.4 Saga P domain**

GPGSKPFTVPILTVEEMTNSRFPIPLEKLFTGPSGAFVVQPQNGRCTTDGVLLGTTQL
SPVNICTFRGDVTHIAGSRNYTMNLASLNWNNYDPTEEIPAPLGTPDFVGKIQGLLTQ
TTKGDGSTRGHKATVYTGSAPFTPKLGSVQFSTDTENDFETHQNTKFTPVGVIQDG
STTHRNEPQQWVLPSYSGRNVHNVHLAPAVAPTFPGEQLLFFRSTMPGCSGYPNM
DLDCLLPQEWVQHFYQEAAPAQSDVALLRFVNPDTGRVLFECKLHKSGYVTVAHT
GQHDLVIPPNGYFRFDSWVNQFYTLAPM

**Human norovirus GII.4 MI001 P domain**

GPGSKPFSVPILTVEEMTNSRFPIPLEKLFTGPSSAFVVQPQNGRCTTDGVLLGTTQL
SPVNICTFRGDVTHIAGTQEYTMNLASQNWNNYDPTEEIPAPLGTPDFVGKIQGVLT
QTTRRDGSTRGHKATVSTGSVHFTPKLGRIQFSTDTSNDFETGQNTRFTPVGVVQD
GSTTHQNEPQQWVLPNYSGRDSHNVHLAPAVAPSFPGEQLLFFRSTMPGCSGYPN
MNLDCLLPQEWVQHFYQEAAPAQSDVALLRFVNPDTGRVLFECKLHKSGYVTVAH
TGQHDLVIPPNGYFRFDSWVNQFYTLAPM

**Human norovirus GII.17 Kawasaki 308 P domain**

GPGSKPFSLPILTLSELTNSRFPVPIDSLFTAQNNVLQVQCQNGRCTLDGELQGTTQL
LPTGICAFRGRVTAQINQRDRWHMQLQNLNGTTYDPTDDVPAPLGTPDFKGVVFG
MVSQRNVGNDAPGSTRAQQAWVSTYSPQFVPKLGSVNLRISDNDDFQFQPTKFTPV
GVNDDDDGHPFRQWELPNYSGELTLNMNLAPPVAPNFPGEQLLFFRSFVPCSGGY
NQGIIDCLIPQEWIQHFYQESAPSQSDVALIRYVNPDTGRTLFEAKLHRSGYITVAHS
GDYPLVVPANGHFRFDSWVNQFYSLAPM

**Human norovirus GII.10 Vietnam 026 P domain**

GPGSKPFTLPILTLGELTNSRFPLPIDVLYTNPNESAIVQCQNGRCTLDGELQGTTQL
LPTGICAFRGKVTQQVQDEHRGTHWNMTVTNLNGTPFDPTEDVPAPLGTPDFSGQI
YGVISQRNTNTVPGEGNLPANRAHEAVIATYSPKFTPKLGNIQFSTWETQDVSSGQP
TKFTPVGLASVDANSHFDQWTLPSYSGALTLNMNLAPSVAPVFPGECLLFFRSFIPLK
GGYGNPAIDCLMPQEWVQHLYQESAPSLSDVALVRYVNPETGRTLFEAKLHRNGFL
TVARNSAGPVVAPTNGYFRFDSWVNQFYTLAPM

**Murine norovirus GV/MNV07 P domain**

GPGSRMVDLPVLQPRLCTHARWPAPVYGLLVDPSLPSNPQWQNGRVHVDGTLLGT
TPVSGSWVSCFAAEAAYEFQSGTGEVATFTLIEQDGSAYVPGDRAAPLGYPDFSGQ
LEIEVQTETTKAGDKLKVTTFEMILGPTTNVDQAPYQGRVHASTSVTASLNLVDGRV
RAVPRSIYSFQDVVPEYNDGLLVPLAPPIGPFLPGEVLLRFRTYMRQIDSSDAAAEAI
DCALPQEFISWFASNAFTVQSEALLLRYRNTLTGQLLFECKLYSEGYIALSYSGSGPL
TFPTDGFFEVVSWVPRLYQLASV

**Murine norovirus GV/CR10 P domain**

GPGSRMVDLPVLQPRLCTHARWPAPIYGLLVDPSLPSNPQWQNGRVHVDGTLLGTT
PVSGSWVSCFAAEAAYEFQSGTGEVATFTLIEQDGSAYVPGDRAAPLGYPDFSGQL
EIEVQTETTKTGDKLKVTTFEMILGPTTNVDQAPYQGRVYASLTAVASLDLVDGRVR
AVPRSIYGFQDVIPEYNDGLLVPLAPPIGPFLPGEVLLRFRTYMRQLDTADAAAEAID
CALPQEFISWFASNAFTVQSDALLLRYRNTLTGQLLFECKLYSEGYIALSYSGSGPLT
FPTDGFFEVVSWVPRLFQLASV

**Mistic (*Bacillus subtilis)***

MFCTFFEKHHRKWDILLEKSTGVMEAMKVTSEEKEQLSTAIDRMNEGLDAFIQLYN
ESEIDEPLIQLDDDTAELMKQARDMYGQEKLNEKLNTIIKQILSISVSEEGEKE

**Angiotensin I**

DRVYIHPFHL

**Bradykinin**

RPPGFSPFR

## 6.5 Hazardous substances according to GHS

Table S 8: Hazard and precautionary statements of hazardous substances used in this work.

| Name | Symbols | Hazard Statements (H)/Precautionary Statements (P) |
|---|---|---|
| Acetic acid | | H226 - Flammable liquid and vapor |
| | | H314 - Causes severe skin burns and eye damage |
| | | P210 - Keep away from heat/sparks/open flames/hot surfaces. - No smoking |
| | | P303 + P361 + P353 - IF ON SKIN (or hair): Remove/ Take off immediately all contaminated clothing. Rinse skin with water/shower |
| | | P280 - Wear protective gloves/ protective clothing/ eye protection/ face protection |
| | | P301 + P330 + P331 - IF SWALLOWED: Rinse mouth. Do NOT induce vomiting |
| | | P305 + P351 + P338 - IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing |
| | | P310 - Immediately call a POISON CENTER or doctor/ physician |
| Acetonitrile | | H225 - Highly flammable liquid and vapor |
| | | H302 - Harmful if swallowed |
| | | H312 - Harmful in contact with skin |
| | | H319 - Causes serious eye irritation |
| | | H332 - Harmful if inhaled |
| | | P210 - Keep away from heat/sparks/open flames/hot surfaces. - No smoking |
| | | P280 - Wear protective gloves/ protective clothing/ eye protection/ face protection |
| | | P301 + P312 - IF SWALLOWED: Call a POISON CENTER or doctor/ physician if you feel unwell |
| | | P302 + P352 - IF ON SKIN: Wash with plenty of soap and water |
| | | P304 + P340 - IF INHALED: Remove to fresh air and keep at rest in a position comfortable for breathing |
| | | P305 + P351 + P338 - IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing |

| Name | Symbols | Hazard Statements (H)/Precautionary Statements (P) |
|------|---------|---------------------------------------------------|
| Acrylamide | | H301 - Toxic if swallowed |
| | | H312 + H332 - Harmful in contact with skin or if inhaled |
| | | H315 - Causes skin irritation |
| | | H319 - Causes serious eye irritation |
| | | H317 - May cause an allergic skin reaction |
| | | H372 - Causes damage to organs through prolonged or repeated exposure |
| | | H350 - May cause cancer |
| | | H340 - May cause genetic defects |
| | | H361 - Suspected of damaging fertility or the unborn child |
| | | P301 + P310 - IF SWALLOWED: Immediately call a POISON CENTRE or doctor/physician |
| | | P302 + P352 - IF ON SKIN: Wash with plenty of soap and water |
| | | P304 + P340 - IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing |
| | | P305 + P351 + P338 - IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing |
| | | P333 + P313 - If skin irritation or rash occurs: Get medical advice/attention |
| | | P260 - Do not breathe dust/fume/gas/mist/vapors/spray |
| | | P202 - Do not handle until all safety precautions have been read and understood |
| Ammonium hydroxide | | H290 May be corrosive to metals. |
| | | H314 Causes severe skin burns and eye damage. |
| | | H335 May cause respiratory irritation. |
| | | H400 Very toxic to aquatic life. |
| | | P273 Avoid release to the environment. |
| | | P280 Wear protective gloves/protective clothing/eye protection/face protection. |
| | | P303+P361+P353 IF ON SKIN (or hair): take off immediately all contaminated clothing. Rinse skin with water/shower. |
| | | P304+P340 IF INHALED: Remove person to fresh air and keep comfortable for breathing. |
| | | P305+P351+P338 IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. |
| | | P310 Immediately call a POISON CENTER/doctor. |

# 6 Supplement

| Name | Symbols | Hazard Statements (H)/Precautionary Statements (P) |
|------|---------|-----------------------------------------------------|
| Ammonium persulfate (APS) | | H272 - May intensify fire; oxidizer |
| | | H302 - Harmful if swallowed |
| | | H315 - Causes skin irritation |
| | | H317 - May cause an allergic skin reaction |
| | | H319 - Causes serious eye irritation |
| | | H334 - May cause allergy or asthma symptoms or breathing difficulties if inhaled |
| | | H335 - May cause respiratory irritation |
| | | P210 - Keep away from heat, hot surfaces, sparks, open flames and other ignition sources. No smoking |
| | | P220 - Keep/Store away from clothing/ combustible materials |
| | | P261 - Avoid breathing dust/fume/gas/mist/vapors/spray |
| | | P264 - Wash hands thoroughly after handling |
| | | P280 - Wear protective gloves/protective clothing/eye protection/face protection |
| | | P301 + P312 - IF SWALLOWED: Call a POISON CENTER or doctor/physician if you feel unwell |
| | | P302 + P352 - IF ON SKIN: Wash with plenty of soap and water |
| | | P304 + P340 - IF INHALED: Remove person to fresh air and keep comfortable for breathing |
| | | P305 + P351 + P338 - IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing |
| | | P333 + P313 - If skin irritation or rash occurs: Get medical advice/attention |
| | | P403 + P233 - Store in a well-ventilated place. Keep container tightly closed |
| | | P501 - Dispose of contents/ container to an approved waste disposal plant |
| Argon | | H280 Contains gas under pressure; may explode if heated. |
| | | P403 Store in a well-ventilated place. |
| Cesium Iodide | | H315 Causes skin irritation. |
| | | H317 May cause an allergic skin reaction. |
| | | H319 Causes serious eye irritation. |
| | | H335 May cause respiratory irritation. |
| | | H410 Very toxic to aquatic life with long lasting effects. |
| | | P280 Wear protective gloves. |
| | | P305+P351+P338 IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. |

| Name | Symbols | Hazard Statements (H)/Precautionary Statements (P) |
|---|---|---|
| N,N-Dimethyl-dodecylamine N-oxide (LDAO) | | H302 Harmful if swallowed. |
| | | H315 Causes skin irritation. |
| | | H318 Causes serious eye damage. |
| | | H410 Very toxic to aquatic life with long lasting effects. |
| | | P264 Wash skin thoroughly after handling. |
| | | P273 Avoid release to the environment. |
| | | P280 Wear protective gloves/ eye protection/ face protection. |
| | | P305 + P351 + P338 +P310 IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. Immediately call a POISON CENTER/doctor. |
| | | P391 Collect spillage. |
| | | P501 Dispose of contents/ container to an approved waste disposal plant. |
| Formic acid | | H226 Flammable liquid and vapour |
| | | H290 May be corrosive to metals |
| | | H302 Harmful if swallowed |
| | | H314 Causes severe skin burns and eye damage |
| | | H331 Toxic if inhaled |
| | | P210 Keep away from heat. No smoking. |
| | | P280 Wear protective gloves/protective clothing/eye protection/face protection. |
| | | P303+P361+P353 IF ON SKIN (or hair): take off immediately all contaminated clothing. Rinse skin with water/shower. |
| | | P304+P340 IF INHALED: Remove person to fresh air and keep comfortable for breathing. |
| | | P305+P351+P338 IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. |
| | | P310 Immediately call a POISON CENTER/doctor. |
| Guanidine Hydrochloride (GndHCl) | | H302 - Harmful if swallowed |
| | | H332 - Harmful if inhaled |
| | | H315 - Causes skin irritation |
| | | H319 - Causes serious eye irritation |
| | | May form combustible dust concentrations in air |
| | | P301 + P330 + P331 - IF SWALLOWED: rinse mouth. Do NOT induce vomiting |
| | | P312 - Call a POISON CENTER or doctor/ physician if you feel unwell |
| | | P304 + P340 - IF INHALED: Remove to fresh air and keep at rest in a position comfortable for breathing |
| | | P302 + P352 - IF ON SKIN: Wash with plenty of soap and water |
| | | P337 + P313 - If eye irritation persists: Get medical advice/ attention |
| | | P280 - Wear protective gloves/ protective clothing/ eye protection/ face protection |

# 6 Supplement

| Name | Symbols | Hazard Statements (H)/Precautionary Statements (P) |
|---|---|---|
| 2-Mercaptoethanol | | H317 - May cause an allergic skin reaction |
| | | P302 + P352 - IF ON SKIN: Wash with plenty of soap and water |
| | | P273 - Avoid release to the environment |
| | | P333 + P313 - If skin irritation or rash occurs: Get medical advice/attention |
| | | P280 - Wear protective gloves/protective clothing/eye protection/face protection |
| | | P261 - Avoid breathing dust/fume/gas/mist/vapors/spray |
| | | P272 - Contaminated work clothing should not be allowed out of the workplace |
| Methanol | | H225 - Highly flammable liquid and vapor |
| | | H301 - Toxic if swallowed |
| | | H311 - Toxic in contact with skin |
| | | H331 - Toxic if inhaled |
| | | H370 - Causes damage to organs |
| | | P210 - Keep away from heat/sparks/open flames/hot surfaces. - No smoking |
| | | P280 - Wear protective gloves/ protective clothing/ eye protection/ face protection |
| | | P301 + P310 - IF SWALLOWED: Immediately call a POISON CENTER or doctor/ physician |
| | | P302 + P350 - IF ON SKIN: Gently wash with plenty of soap and water |
| | | P304 + P340 - IF INHALED: Remove to fresh air and keep at rest in a position comfortable for breathing |
| | | P240 - Ground/Bond container and receiving equipment |
| Nitrogen | | H280 Contains gas under pressure; may explode if heated. |
| | | P403 Store in a well-ventilated place. |
| Pepsin immobilized on Resin beads | | H334 - May cause allergy or asthma symptoms or breathing difficulties if inhaled |
| | | P261 - Avoid breathing dust/fume/gas/mist/vapours/spray |
| | | P284 - In case of inadequate ventilation wear respiratory protection |
| | | P304 + P340 - IF INHALED: Remove person to fresh air and keep comfortable for breathing |
| | | P342 + P311 - If experiencing respiratory symptoms: Call a POISON CENTER or doctor/physician |
| | | P501 - Dispose of contents/ container to an approved waste disposal plant |

| Name | Symbols | Hazard Statements (H)/Precautionary Statements (P) |
|---|---|---|
| Pierce® LTQ Velos ESI Positive Ion Calibration Solution | | H225 - Highly flammable liquid and vapour |
| | | H301 + H311 + H331 - Toxic if swallowed, in contact with skin or if inhaled |
| | | H319 - Causes serious eye irritation |
| | | H370 - Causes damage to organs |
| | | P280 - Wear protective gloves/protective clothing/eye protection/face protection |
| | | P210 - Keep away from heat, hot surfaces, sparks, open flames and other ignition sources. No smoking |
| | | P241 - Use explosion-proof electrical/ ventilating/ lighting/ equipment |
| | | P260 - Do not breathe dust/fume/gas/mist/vapours/spray |
| | | P304 + P340 - IF INHALED: Remove person to fresh air and keep comfortable for breathing |
| | | P301 + P310 - IF SWALLOWED: Immediately call a POISON CENTER or doctor/physician |
| | | P303 + P361 + P353 - IF ON SKIN (or hair): Remove/Take off immediately all contaminated clothing. Rinse skin with water/shower |
| | | P403 + P235 - Store in a well-ventilated place. Keep cool |
| | | P501 - Dispose of contents/ container to an approved waste disposal plant |
| Phosphoric acid | | H290 - May be corrosive to metals |
| | | H314 - Causes severe skin burns and eye damage |
| | | P280 - Wear protective gloves/ protective clothing/ eye protection/ face protection |
| | | P301 + P330 + P331 - IF SWALLOWED: Rinse mouth. Do NOT induce vomiting |
| | | P303 + P361 + P353 - IF ON SKIN (or hair): Remove/ Take off immediately all contaminated clothing. Rinse skin with water/shower |
| | | P305 + P351 + P338 - IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing |
| | | P310 - Immediately call a POISON CENTER or doctor/ physician |

| Name | Symbols | Hazard Statements (H)/Precautionary Statements (P) |
|------|---------|---------------------------------------------------|
| Sodium Dodecyl Sulfate (SDS) | | H228 - Flammable solid |
| | | H302 - Harmful if swallowed |
| | | H332 - Harmful if inhaled |
| | | H315 - Causes skin irritation |
| | | H318 - Causes serious eye damage |
| | | H335 - May cause respiratory irritation |
| | | H412 - Harmful to aquatic life with long lasting effects |
| | | P301 + P312 - IF SWALLOWED: Call a POISON CENTRE or doctor/physician if you feel unwell |
| | | P302 + P352 - IF ON SKIN: Wash with plenty of soap and water |
| | | P305 + P351 + P338 - IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing |
| | | P304 + P340 - IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing |
| | | P273 - Avoid release to the environment |
| | | P370 + P378 - In case of fire, use water/water spray/water jet/carbon dioxide/sand/foam/alcohol resistant foam/chemical powder for extinction |
| Sodium hydroxide | | H290 - May be corrosive to metals |
| | | H314 - Causes severe skin burns and eye damage |
| | | P280 - Wear protective gloves/ protective clothing/ eye protection/ face protection |
| | | P301 + P330 + P331 - IF SWALLOWED: Rinse mouth. Do NOT induce vomiting |
| | | P303 + P361 + P353 - IF ON SKIN (or hair): Remove/ Take off immediately all contaminated clothing. Rinse skin with water/shower |
| | | P305 + P351 + P338 - IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing |
| | | P310 - Immediately call a POISON CENTER or doctor/ physician |

| Name | Symbols | Hazard Statements (H)/Precautionary Statements (P) |
|------|---------|----------------------------------------------------|
| *N,N,N',N'*-Tetramethyl ethylenediamine (TEMED) | | H225 - Highly flammable liquid and vapor |
| | | H301 - Toxic if swallowed |
| | | H314 - Causes severe skin burns and eye damage |
| | | H318 - Causes serious eye damage |
| | | H331 - Toxic if inhaled |
| | | P210 - Keep away from heat, hot surfaces, sparks, open flames and other ignition sources. No smoking |
| | | P280 - Wear protective gloves/protective clothing/eye protection/face protection |
| | | P301 + P310 - IF SWALLOWED: Immediately call a POISON CENTRE or doctor/physician |
| | | P303 + P361 + P353 - IF ON SKIN (or hair): Remove/Take off immediately all contaminated clothing. Rinse skin with water/shower |
| | | P304 + P340 - IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing |
| | | P403 + P235 - Store in a well-ventilated place. Keep cool |
| | | P501 - Dispose of contents/ container to an approved waste disposal plant |

# LIST OF FIGURES

## Main figures

**Supplementary figures**

## LIST OF TABLES

### Main tables

### Supplementary tables

## CONTRIBUTIONS

This thesis is a result of collaborative projects and involved contributions from several people. The presented data are mass spectrometry-centered views on the projects. Supportive results from other techniques were added to this thesis to underline and accentuate the MS findings. Unless specified differently, all work has been performed by me, Jasmin Dülfer (JD). The introduction chapter is in part taken from two review articles, in which I was first and second author. The following list is an overview over other author's contributions:

**Chapter 1.1:** JD, AK, JDK and BK did literature research and wrote the paper. JD prepared the figures. JD, AK and CU took care of critical revision of the manuscript.

Reference: Dülfer et al. (2019)

**Chapter 1.2:** JD and RP did literature research and wrote the paper. JD prepared figures and tables. RP, JD and CU took care of critical revision of the manuscript.

Reference: Pogan et al. (2018a)

**Chapter 3.1:** AM, RC and JMO produced and purified the proteins and performed IEX and NMR measurements. BB and PM performed crystallography experiments. JD performed MS measurements and data analysis. JD wrote the chapter. The results, discussion and associated method section was adapted from the published paper, originally written by AM, RC, TP and JD.

Reference: Mallagaray et al. (2019)

**Chapter 3.2:** AM and RC produced and purified the proteins. JD and HY performed native MS measurements (HY: Saga, Vietnam and Kawasaki strain, JD: MI001 strain). HY and JD analyzed native MS data. JD and HY performed HDX-MS measurements (HY: Kawasaki strain and Vietnam strain + fucose, JD: MI001 strain, Saga strain and Vietnam strain + HBGA B trisaccharide). JD analyzed and interpreted the HDX-MS data. JD wrote the chapter.

Reference: Dülfer et al. (in preparation)

**Chapter 3.3:** RC produced and purified the proteins. JD performed native MS measurements and data analysis. JD wrote the chapter.

Reference: Creutznacher et al. (in preparation)

**Chapter 3.4:** MB produced, purified and reconstituted the protein. MB performed reconstitution quality control by DLS. JD performed HDX-MS method optimizations and measured the HDX-MS data. JD analyzed and interpreted the data and wrote the chapter. MB provided information for the accompanying methods sections. The Mistic expression and purification section was adapted from Hartmann et al. (2015).

Reference: Krainer et al. (in preparation)

**Chapter 4:** GH and JMG produced and purified VLPs. JK and KTB produced MNV and inactivated the virus. JD performed pepsin digestion tests with inactivated MNV and MS experiments with VLPs.

**Initials and names:**

| | | | |
|---|---|---|---|
| AK | Alan Kádek | JDK | Janine-Denise Kopicki |
| AM | Alvaro Mallagaray | JK | Jan Knickmann |
| BB | Bärbel Blaum | JMG | Jürgen Müller-Guhl |
| BK | Boris Krichel | JMO | Jose Maria Orduña |
| CU | Charlotte Uetrecht | MB | Marta Batet |
| GH | Grant Hansman | RC | Robert Creutznacher |
| HY | Hao Yan | RP | Ronja Pogan |
| KTB | Katja Thiele-Bössel | TP | Thomas Peters |

# REFERENCES

Ahmed, S. M., Hall, A. J., Robinson, A. E., Verhoef, L., Premkumar, P., Parashar, U. D., Koopmans, M. & Lopman, B. A. 2014. Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis. *Lancet Infect Dis,* 14**,** 725-730.

Ahmed, S. M., Lopman, B. A. & Levy, K. 2013. A systematic review and meta-analysis of the global seasonality of norovirus. *PLoS One,* 8**,** e75922.

Anderson, K. W., Gallagher, E. S. & Hudgens, J. W. 2018. Automated Removal of Phospholipids from Membrane Proteins for H/D Exchange Mass Spectrometry Workflows. *Anal Chem,* 90**,** 6409-6412.

Arora, J., Hickey, J. M., Majumdar, R., Esfandiary, R., Bishop, S. M., Samra, H. S., Middaugh, C. R., Weis, D. D. & Volkin, D. B. 2015. Hydrogen exchange mass spectrometry reveals protein interfaces and distant dynamic coupling effects during the reversible self-association of an IgG1 monoclonal antibody. *MAbs,* 7**,** 525-39.

Avantipolarlipids. 2020. *Preparing Large, Unilamellar Vesicles by Extrusion (LUVET)* [Online]. Alabama, US. Available: https://avantilipids.com/tech-support/liposome-preparation/luvet [Accessed 14.05.2020].

Baclayon, M., Shoemaker, G. K., Uetrecht, C., Crawford, S. E., Estes, M. K., Prasad, B. V., Heck, A. J., Wuite, G. J. & Roos, W. H. 2011. Prestress strengthens the shell of Norwalk virus nanoparticles. *Nano Lett,* 11**,** 4865-9.

Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. 1993. Primary structure effects on peptide group hydrogen exchange. *Proteins,* 17**,** 75-86.

Benhaim, M. A. & Lee, K. K. 2020. New Biophysical Approaches Reveal the Dynamics and Mechanics of Type I Viral Fusion Machinery and Their Interplay with Membranes. *Viruses,* 12.

Benhaim, M. A., Mangala Prasad, V., Garcia, N. K., Guttman, M. & Lee, K. K. 2020. Structural monitoring of a transient intermediate in the hemagglutinin fusion machinery on influenza virions. *Sci Adv,* 6**,** eaaz8822.

Benkert, P., Biasini, M. & Schwede, T. 2011. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics,* 27**,** 343-50.

Bereszczak, J. Z., Barbu, I. M., Tan, M., Xia, M., Jiang, X., Van Duijn, E. & Heck, A. J. 2012. Structure, stability and dynamics of norovirus P domain derived protein complexes studied by native mass spectrometry. *J Struct Biol,* 177**,** 273-82.

Bereszczak, J. Z., Rose, R. J., Van Duijn, E., Watts, N. R., Wingfield, P. T., Steven, A. C. & Heck, A. J. 2013. Epitope-distal effects accompany the binding of two distinct antibodies to hepatitis B virus capsids. *J Am Chem Soc,* 135**,** 6504-12.

# References

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. 2000. The Protein Data Bank. *Nucleic Acids Res,* 28**,** 235-42.

Bertolotti-Ciarlet, A., Crawford, S. E., Hutson, A. M. & Estes, M. K. 2003. The 3 ' end of Norwalk virus mRNA contains determinants that regulate the expression and stability of the viral capsid protein VP1: a novel function for the VP2 protein. *Journal of Virology,* 77**,** 11603-11615.

Bertolotti-Ciarlet, A., White, L. J., Chen, R., Prasad, B. V. & Estes, M. K. 2002. Structural requirements for the assembly of Norwalk virus-like particles. *J Virol,* 76**,** 4044-55.

Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L. & Schwede, T. 2017. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep,* 7**,** 10480.

Boeri Erba, E. & Petosa, C. 2015. The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes. *Protein Sci,* 24**,** 1176-92.

Bozkurt, H., D'souza, D. H. & Davidson, P. M. 2013. Determination of the Thermal Inactivation Kinetics of the Human Norovirus Surrogates, Murine Norovirus and Feline Calicivirus. *Journal of Food Protection,* 76**,** 79-84.

Brodbelt, J. S. 2014. Photodissociation mass spectrometry: new tools for characterization of biological molecules. *Chem Soc Rev,* 43**,** 2757-83.

Brodbelt, J. S. 2016. Ion Activation Methods for Peptides and Proteins. *Anal Chem,* 88**,** 30-51.

Broecker, J., Fiedler, S., Gimpl, K. & Keller, S. 2014. Polar interactions trump hydrophobicity in stabilizing the self-inserting membrane protein Mistic. *J Am Chem Soc,* 136**,** 13761-8.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T. L. 2009. BLAST+: architecture and applications. *BMC Bioinformatics,* 10**,** 421.

Cannon, J. L., Papafragkou, E., Park, G. W., Osborne, J., Jaykus, L. A. & Vinje, J. 2006. Surrogates for the study of norovirus stability and inactivation in the environment: A comparison of murine norovirus and feline calicivirus. *Journal of Food Protection,* 69**,** 2761-2765.

Cao, S., Lou, Z., Tan, M., Chen, Y., Liu, Y., Zhang, Z., Zhang, X. C., Jiang, X., Li, X. & Rao, Z. 2007. Structural Basis for the Recognition of Blood Group Trisaccharides by Norovirus. *Journal of Virology,* 81**,** 5949-5957.

Chalmers, M. J., Busby, S. A., Pascal, B. D., West, G. M. & Griffin, P. R. 2011. Differential hydrogen/deuterium exchange mass spectrometry analysis of protein-ligand interactions. *Expert Rev Proteomics,* 8**,** 43-59.

Chan, M. C. W., Hu, Y., Chen, H., Podkolzin, A. T., Zaytseva, E. V., Komano, J., Sakon, N., Poovorawan, Y., Vongpunsawad, S., Thanusuwannasak, T., Hewitt, J.,

Croucher, D., Collins, N., Vinje, J., Pang, X. L., Lee, B. E., De Graaf, M., Van Beek, J., Vennema, H., Koopmans, M. P. G., Niendorf, S., Poljsak-Prijatelj, M., Steyer, A., White, P. A., Lun, J. H., Mans, J., Hung, T. N., Kwok, K., Cheung, K., Lee, N. & Chan, P. K. S. 2017. Global Spread of Norovirus GII.17 Kawasaki 308, 2014-2016. *Emerg Infect Dis,* 23**,** 1359-1354.

Chavez, J. D., Lee, C. F., Caudal, A., Keller, A., Tian, R. & Bruce, J. E. 2018. Chemical Crosslinking Mass Spectrometry Analysis of Protein Conformations and Supercomplexes in Heart Tissue. *Cell Syst,* 6**,** 136-141 e5.

Chazin, W. J., Koerdel, J., Thulin, E., Hofmann, T., Drakenberg, T. & Forsen, S. 1989. Identification of an isoaspartyl linkage formed upon deamidation of bovine calbindin D9k and structural characterization by 2D proton NMR. *Biochemistry,* 28**,** 8646-8653.

Chen, B., Vogan, E. M., Gong, H., Skehel, J. J., Wiley, D. C. & Harrison, S. C. 2005. Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature,* 433**,** 834-41.

Chen, M. W., Tan, Y. B., Zheng, J., Zhao, Y., Lim, B. T., Cornvik, T., Lescar, J., Ng, L. F. P. & Luo, D. 2017. Chikungunya virus nsP4 RNA-dependent RNA polymerase core domain displays detergent-sensitive primer extension and terminal adenylyltransferase activities. *Antiviral Res,* 143**,** 38-47.

Chernushevich, I. V., Loboda, A. V. & Thomson, B. A. 2001. An introduction to quadrupole-time-of-flight mass spectrometry. *J Mass Spectrom,* 36**,** 849-65.

Choi, J. M., Hutson, A. M., Estes, M. K. & Prasad, B. V. 2008. Atomic resolution structural characterization of recognition of histo-blood group antigens by Norwalk virus. *Proc Natl Acad Sci U S A,* 105**,** 9175-80.

Chorev, D. S., Baker, L. A., Wu, D., Beilsten-Edmands, V., Rouse, S. L., Zeev-Ben-Mordehai, T., Jiko, C., Samsudin, F., Gerle, C., Khalid, S., Stewart, A. G., Matthews, S. J., Grunewald, K. & Robinson, C. V. 2018. Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry. *Science,* 362**,** 829-834.

Conley, M. J., Mcelwee, M., Azmi, L., Gabrielsen, M., Byron, O., Goodfellow, I. G. & Bhella, D. 2019. Calicivirus VP2 forms a portal-like assembly following receptor engagement. *Nature,* 565**,** 377-381.

Cravello, L., Lascoux, D. & Forest, E. 2003. Use of different proteases working in acidic conditions to improve sequence coverage and resolution in hydrogen/deuterium exchange of large proteins. *Rapid Commun Mass Spectrom,* 17**,** 2387-93.

Creutznacher, R., Feldmann, C., Dülfer, J., Maass, T., Knickmann, J., Uetrecht, C., Peters, T. & Taube, S. in preparation. Murine norovirus capsid plasticity probed by NMR - Bile acid binding stabilizes and rearranges P domain dimers and triggers immune escape.

# References

Creutznacher, R., Schulze, E., Wallmann, G., Peters, T., Stein, M. & Mallagaray, A. 2019. Chemical-Shift Perturbations Reflect Bile Acid Binding to Norovirus Coat Protein: Recognition Comes in Different Flavors. *Chembiochem.*

Creutznacher, R., Wallmann, G., Maass, T., Ogrissek, P., Feldmann, C., Peters, T. & Mallagaray, A. submitted. NMR experiments shed new light on glycan recognition by human and murine norovirus capsid proteins

De Graaf, M., Van Beek, J. & Koopmans, M. P. 2016. Human norovirus transmission and evolution in a changing world. *Nat Rev Microbiol,* 14**,** 421-33.

De Graaf, M., Van Beek, J., Vennema, H., Podkolzin, A. T., Hewitt, J., Bucardo, F., Templeton, K., Mans, J., Nordgren, J., Reuter, G., Lynch, M., Rasmussen, L. D., Iritani, N., Chan, M. C., Martella, V., Ambert-Balay, K., Vinje, J., White, P. A. & Koopmans, M. P. 2015. Emergence of a novel GII.17 norovirus - End of the GII.4 era? *Eurosurveillance,* 20**,** 8-15.

Debnath, D. K., Basaiawmoit, R. V., Nielsen, K. L. & Otzen, D. E. 2011. The role of membrane properties in Mistic folding and dimerisation. *Protein Eng Des Sel,* 24**,** 89-97.

Deutsch, E. W., Csordas, A., Sun, Z., Jarnuczak, A., Perez-Riverol, Y., Ternent, T., Campbell, D. S., Bernal-Llinares, M., Okuda, S., Kawano, S., Moritz, R. L., Carver, J. J., Wang, M., Ishihama, Y., Bandeira, N., Hermjakob, H. & Vizcaino, J. A. 2017. The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. *Nucleic Acids Res,* 45**,** D1100-D1106.

Deverman, B. E., Cook, B. L., Manson, S. R., Niederhoff, R. A., Langer, E. M., Rosová, I., Kulans, L. A., Fu, X., Weinberg, J. S., Heinecke, J. W., Roth, K. A. & Weintraub, S. J. 2002. Bcl-xL Deamidation Is a Critical Switch in the Regulation of the Response to DNA Damage. *Cell,* 111**,** 51-62.

Donaldson, E. F., Lindesmith, L. C., Lobue, A. D. & Baric, R. S. 2010. Viral shape-shifting: norovirus evasion of the human immune system. *Nat Rev Microbiol,* 8**,** 231-41.

Douglas, D. J., Frank, A. J. & Mao, D. 2005. Linear ion traps in mass spectrometry. *Mass Spectrom Rev,* 24**,** 1-29.

Duc, N. M., Du, Y., Zhang, C., Lee, S. Y., Thorsen, T. S., Kobilka, B. K. & Chung, K. Y. 2015. Effective application of bicelles for conformational analysis of G protein-coupled receptors by hydrogen/deuterium exchange mass spectrometry. *J Am Soc Mass Spectrom,* 26**,** 808-817.

Dülfer, J., Kadek, A., Kopicki, J. D., Krichel, B. & Uetrecht, C. 2019. Structural mass spectrometry goes viral *In:* REY, F. A. (ed.) *Virus Structure and Function.* Elsevier.

Dülfer, J., Yan, H., Brodmerkel, M., Creutznacher, R., Mallagaray, A., Peters, T., Caleman, C., Marklund, E. G. & Uetrecht, C. in preparation. Glycan-induced structural dynamics in human norovirus P domains depend on virus strain and deamidation status.

Dvir, H., Lundberg, M. E., Maji, S. K., Riek, R. & Choe, S. 2009. Mistic: cellular localization, solution behavior, polymerization, and fibril formation. *Protein Sci,* 18**,** 1564-70.

Dykeman, E. C., Stockley, P. G. & Twarock, R. 2013. Packaging signals in two single-stranded RNA viruses imply a conserved assembly mechanism and geometry of the packaged genome. *J Mol Biol,* 425**,** 3235-49.

Eakin, C. M., Miller, A., Kerr, J., Kung, J. & Wallace, A. 2014. Assessing analytical methods to monitor isoAsp formation in monoclonal antibodies. *Front Pharmacol,* 5**,** 87.

Eden, J. S., Tanaka, M. M., Boni, M. F., Rawlinson, W. D. & White, P. A. 2013. Recombination within the Pandemic Norovirus GII.4 Lineage. *Journal of Virology,* 87**,** 6270-6282.

Eisinger, M. L., Dorrbaum, A. R., Michel, H., Padan, E. & Langer, J. D. 2017. Ligand-induced conformational dynamics of the Escherichia coli Na(+)/H(+) antiporter NhaA revealed by hydrogen/deuterium exchange mass spectrometry. *Proc Natl Acad Sci U S A,* 114**,** 11691-11696.

Eliuk, S. & Makarov, A. 2015. Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annu Rev Anal Chem (Palo Alto Calif),* 8**,** 61-80.

Engen, J. R. 2009. Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS. *Anal Chem,* 81**,** 7870-5.

Engen, J. R. & Smith, D. L. 2000. Investigating the higher order structure of proteins. Hydrogen exchange, proteolytic fragmentation, and mass spectrometry. *Methods Mol Biol,* 146**,** 95-112.

Engen, J. R. & Wales, T. E. 2015. Analytical Aspects of Hydrogen Exchange Mass Spectrometry. *Annu Rev Anal Chem (Palo Alto Calif),* 8**,** 127-48.

Englander, S. W. 2006. Hydrogen exchange and mass spectrometry: A historical perspective. *J Am Soc Mass Spectrom,* 17**,** 1481-9.

Englander, S. W., Mayne, L., Bai, Y. & Sosnick, T. R. 1997. Hydrogen exchange: the modern legacy of Linderstrom-Lang. *Protein Sci,* 6**,** 1101-9.

Eschweiler, J. D., Frank, A. T. & Ruotolo, B. T. 2017. Coming to Grips with Ambiguity: Ion Mobility-Mass Spectrometry for Protein Quaternary Structure Assignment. *J Am Soc Mass Spectrom.*

Espino, J. A., Mali, V. S. & Jones, L. M. 2015. In Cell Footprinting Coupled with Mass Spectrometry for the Structural Analysis of Proteins in Live Cells. *Anal Chem,* 87**,** 7971-8.

Ettayebi, K., Crawford, S. E., Murakami, K., Broughman, J. R., Karandikar, U., Tenge, V. R., Neill, F. H., Blutt, S. E., Zeng, X. L., Qu, L., Kou, B., Opekun, A. R., Burrin, D., Graham, D. Y., Ramani, S., Atmar, R. L. & Estes, M. K. 2016. Replication of human noroviruses in stem cell-derived human enteroids. *Science.*

# References

Fallahi, S. & Mattison, K. 2011. Evaluation of Murine Norovirus Persistence in Environments Relevant to Food Production and Processing. *Journal of Food Protection,* 74**,** 1847-1851.

Fang, J., Rand, K. D., Beuning, P. J. & Engen, J. R. 2011. False EX1 signatures caused by sample carryover during HX MS analyses. *Int J Mass Spectrom,* 302**,** 19-25.

Frotscher, E., Krainer, G., Hartmann, A., Schlierf, M. & Keller, S. 2018. Conformational Dynamics Govern the Free-Energy Landscape of a Membrane-Interacting Protein. *ACS Omega,* 3**,** 12026-12032.

Gan, J., Ben-Nissan, G., Arkind, G., Tarnavsky, M., Trudeau, D., Noda Garcia, L., Tawfik, D. S. & Sharon, M. 2017. Native mass spectrometry of recombinant proteins from crude cell lysates. *Analytical chemistry,* 89**,** 4398-4404.

Garcia-Moreno, M., Noerenberg, M., Ni, S., Jarvelin, A. I., Gonzalez-Almela, E., Lenz, C. E., Bach-Pages, M., Cox, V., Avolio, R., Davis, T., Hester, S., Sohier, T. J. M., Li, B., Heikel, G., Michlewski, G., Sanz, M. A., Carrasco, L., Ricci, E. P., Pelechano, V., Davis, I., Fischer, B., Mohammed, S. & Castello, A. 2019. System-wide Profiling of RNA-Binding Proteins Uncovers Key Regulators of Virus Infection. *Mol Cell.*

Geiger, T. & Clarke, S. 1987. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *Journal of Biological Chemistry,* 262**,** 785-794.

Glass, P. J., White, L. J., Ball, J. M., Leparc-Goffart, I., Hardy, M. E. & Estes, M. K. 2000. Norwalk virus open reading frame 3 encodes a minor structural protein. *Journal of Virology,* 74**,** 6581-6591.

Goth, M. & Pagel, K. 2017. Ion mobility-mass spectrometry as a tool to investigate protein-ligand interactions. *Anal Bioanal Chem,* 409**,** 4305-4310.

Grassi, L., Regl, C., Wildner, S., Gadermaier, G., Huber, C. G., Cabrele, C. & Schubert, M. 2017. Complete NMR Assignment of Succinimide and Its Detection and Quantification in Peptides and Intact Proteins. *Anal Chem,* 89**,** 11962-11970.

Griffith, S. C., Sawaya, M. R., Boutz, D. R., Thapar, N., Katz, J. E., Clarke, S. & Yeates, T. O. 2001. Crystal structure of a protein repair methyltransferase from Pyrococcus furiosus with its L-isoaspartyl peptide substrate. *J Mol Biol,* 313**,** 1103-16.

Guttman, M., Scian, M. & Lee, K. K. 2011. Tracking hydrogen/deuterium exchange at glycan sites in glycoproteins by mass spectrometry. *Anal Chem,* 83**,** 7492-9.

Guttman, M., Weis, D. D., Engen, J. R. & Lee, K. K. 2013. Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. *J Am Soc Mass Spectrom,* 24**,** 1906-12.

Haag, A. M. 2016. Mass Analyzers and Mass Spectrometers. *Adv Exp Med Biol,* 919**,** 157-169.

Haga, K., Fujimoto, A., Takai-Todaka, R., Miki, M., Doan, Y. H., Murakami, K., Yokoyama, M., Murata, K., Nakanishi, A. & Katayama, K. 2016. Functional receptor molecules CD300lf and CD300ld within the CD300 family enable murine noroviruses to infect cells. *Proc Natl Acad Sci U S A,* 113**,** E6248-E6255.

Hageman, T. S. & Weis, D. D. 2019a. Reliable Identification of Significant Differences in Differential Hydrogen Exchange-Mass Spectrometry Measurements Using a Hybrid Significance Testing Approach. *Anal Chem,* 91**,** 8008-8016.

Hageman, T. S. & Weis, D. D. 2019b. A Structural Variant Approach for Establishing a Detection Limit in Differential Hydrogen Exchange-Mass Spectrometry Measurements. *Anal Chem,* 91**,** 8017-8024.

Hagemans, D., Van Belzen, I. A., Moran Luengo, T. & Rudiger, S. G. 2015. A script to highlight hydrophobicity and charge on protein surfaces. *Front Mol Biosci,* 2**,** 56.

Hansman, G. S., Biertumpfel, C., Georgiev, I., Mclellan, J. S., Chen, L., Zhou, T., Katayama, K. & Kwong, P. D. 2011. 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**,** 6687-701.

Hao, P., Adav, S. S., Gallart-Palau, X. & Sze, S. K. 2017. Recent advances in mass spectrometric analysis of protein deamidation. *Mass Spectrom Rev,* 36**,** 677-692.

Hartmann, A., Krainer, G., Keller, S. & Schlierf, M. 2015. Quantification of Millisecond Protein-Folding Dynamics in Membrane-Mimetic Environments by Single-Molecule Forster Resonance Energy Transfer Spectroscopy. *Anal Chem,* 87**,** 11224-32.

Hebling, C. M., Morgan, C. R., Stafford, D. W., Jorgenson, J. W., Rand, K. D. & Engen, J. R. 2010. Conformational analysis of membrane proteins in phospholipid bilayer nanodiscs by hydrogen exchange mass spectrometry. *Anal Chem,* 82**,** 5415-9.

Hodge, E. A., Benhaim, M. A. & Lee, K. K. 2019. Bridging protein structure, dynamics, and function using hydrogen/deuterium-exchange mass spectrometry. *Protein Sci.*

Hogan, C. J., Jr., Kettleson, E. M., Ramaswami, B., Chen, D. R. & Biswas, P. 2006. Charge reduced electrospray size spectrometry of mega- and gigadalton complexes: whole viruses and virus fragments. *Anal Chem,* 78**,** 844-52.

Horne, J. E. & Radford, S. E. 2016. A growing toolbox of techniques for studying beta-barrel outer membrane protein folding and biogenesis. *Biochem Soc Trans,* 44**,** 802-9.

Houde, D., Arndt, J., Domeier, W., Berkowitz, S. & Engen, J. R. 2009. Characterization of IgG1 conformation and conformational dynamics by hydrogen/deuterium exchange mass spectrometry. *Anal Chem,* 81**,** 2644-51.

# References

Houde, D., Berkowitz, S. A. & Engen, J. R. 2011. The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. *J Pharm Sci,* 100**,** 2071-86.

Jacso, T., Bardiaux, B., Broecker, J., Fiedler, S., Baerwinkel, T., Mainz, A., Fink, U., Vargas, C., Oschkinat, H., Keller, S. & Reif, B. 2013. The mechanism of denaturation and the unfolded state of the alpha-helical membrane-associated protein Mistic. *J Am Chem Soc,* 135**,** 18884-91.

Jefferson, R. E., Min, D., Corin, K., Wang, J. Y. & Bowie, J. U. 2018. Applications of Single-Molecule Methods to Membrane Protein Folding Studies. *J Mol Biol,* 430**,** 424-437.

Jensen, P. F., Comamala, G., Trelle, M. B., Madsen, J. B., Jorgensen, T. J. & Rand, K. D. 2016. Removal of N-Linked Glycosylations at Acidic pH by PNGase A Facilitates Hydrogen/Deuterium Exchange Mass Spectrometry Analysis of N-Linked Glycoproteins. *Anal Chem,* 88**,** 12479-12488.

Jensen, P. F. & Rand, K. D. 2016. Hydrogen Exchange. *In:* WEIS, D. D. (ed.) *Hydrogen Exchange Mass Spectrometry of Proteins.*

Jiang, X., Wang, M., Wang, K. & Estes, M. K. 1993. Sequence and genomic organization of Norwalk virus. *Virology,* 195**,** 51-61.

Jung, J., Grant, T., Thomas, D. R., Diehnelt, C. W., Grigorieff, N. & Joshua-Tor, L. 2019. High-resolution cryo-EM structures of outbreak strain human norovirus shells reveal size variations. *Proc Natl Acad Sci U S A,* 116**,** 12828-12832.

Kadek, A., Mrazek, H., Halada, P., Rey, M., Schriemer, D. C. & Man, P. 2014. Aspartic protease nepenthesin-1 as a tool for digestion in hydrogen/deuterium exchange mass spectrometry. *Anal Chem,* 86**,** 4287-94.

Kamarasu, P., Hsu, H. Y. & Moore, M. D. 2018. Research Progress in Viral Inactivation Utilizing Human Norovirus Surrogates. *Frontiers in Sustainable Food Systems,* 2.

Kang, Y., Gohlke, U., Engstrom, O., Hamark, C., Scheidt, T., Kunstmann, S., Heinemann, U., Widmalm, G., Santer, M. & Barbirz, S. 2016. Bacteriophage Tailspikes and Bacterial O-Antigens as a Model System to Study Weak-Affinity Protein-Polysaccharide Interactions. *J Am Chem Soc.*

Kapikian, A. Z., Wyatt, R. G., Dolin, R., Thornhill, T. S., Kalica, A. R. & Chanock, R. M. 1972. Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. *J Virol,* 10**,** 1075-81.

Karas, M., Bachmann, D., Bahr, U. & Hillenkamp, F. 1987. Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International journal of mass spectrometry and ion processes,* 78**,** 53-68.

Katta, V. & Chait, B. T. 1991. Conformational-Changes in Proteins Probed by Hydrogen-Exchange Electrospray-Ionization Mass-Spectrometry. *Rapid Communications in Mass Spectrometry,* 5**,** 214-217.

Kavan, D. & Man, P. 2011. MSTools—Web based application for visualization and presentation of HXMS data. *International Journal of Mass Spectrometry,* 302**,** 53-58.

Kefala, G., Kwiatkowski, W., Esquivies, L., Maslennikov, I. & Choe, S. 2007. Application of Mistic to improving the expression and membrane integration of histidine kinase receptors from Escherichia coli. *J Struct Funct Genomics,* 8**,** 167-72.

Kilic, T., Koromyslova, A. & Hansman, G. S. 2019. Structural Basis for Human Norovirus Capsid Binding to Bile Acids. *J Virol,* 93.

Kilic, T., Koromyslova, A., Malak, V. & Hansman, G. S. 2018. Atomic Structure of the Murine Norovirus Protruding Domain and Soluble CD300lf Receptor Complex. *J Virol,* 92.

Konermann, L. 2017. Addressing a Common Misconception: Ammonium Acetate as Neutral pH "Buffer" for Native Electrospray Mass Spectrometry. *Journal of The American Society for Mass Spectrometry,* 28**,** 1827-1835.

Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. 2013. Unraveling the Mechanism of Electrospray Ionization. *Analytical Chemistry,* 85**,** 2-9.

Konermann, L., Pan, J. & Liu, Y. H. 2011. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem Soc Rev,* 40**,** 1224-34.

Konermann, L., Rodriguez, A. D. & Sowole, M. A. 2014. Type 1 and Type 2 scenarios in hydrogen exchange mass spectrometry studies on protein-ligand complexes. *Analyst,* 139**,** 6078-6087.

Koopmans, M., Vennema, H., Heersma, H., Van Strien, E., Van Duynhoven, Y., Brown, D., Reacher, M., Lopman, B. & European Consortium on Foodborne, V. 2003. Early identification of common-source foodborne virus outbreaks in Europe. *Emerg Infect Dis,* 9**,** 1136-42.

Koromyslova, A., Tripathi, S., Morozov, V., Schroten, H. & Hansman, G. S. 2017. Human norovirus inhibition by a human milk oligosaccharide. *Virology,* 508**,** 81-89.

Koromyslova, A. D., Leuthold, M. M., Bowler, M. W. & Hansman, G. S. 2015. The sweet quartet: Binding of fucose to the norovirus capsid. *Virology,* 483**,** 203-8.

Krainer, G., Batet, M., Dülfer, J., Anadamurugan, A., Textor, M., Frotscher, E., Hartmann, A., Uetrecht, C., Keller, S. & Schlierf, M. in preparation.

Krainer, G., Hartmann, A., Anandamurugan, A., Gracia, P., Keller, S. & Schlierf, M. 2018. Ultrafast Protein Folding in Membrane-Mimetic Environments. *J Mol Biol,* 430**,** 554-564.

Kumar, S., Prakash, S., Gupta, K., Dongre, A., Balaram, P. & Balaram, H. 2016. Unexpected functional implication of a stable succinimide in the structural stability of Methanocaldococcus jannaschii glutaminase. *Nat Commun,* 7**,** 12798.

References

Lee, J. C., Kang, S. U., Jeon, Y., Park, J. W., You, J. S., Ha, S. W., Bae, N., Lubec, G., Kwon, S. H., Lee, J. S., Cho, E. J. & Han, J. W. 2012. Protein L-isoaspartyl methyltransferase regulates p53 activity. *Nat Commun,* 3**,** 927.

Leney, A. C. & Heck, A. J. R. 2017. Native Mass Spectrometry: What is in the Name? *Journal of the American Society for Mass Spectrometry,* 28**,** 5-13.

Lermyte, F., Valkenborg, D., Loo, J. A. & Sobott, F. 2018. Radical solutions: Principles and application of electron-based dissociation in mass spectrometry-based analysis of protein structure. *Mass Spectrom Rev,* 37**,** 750-771.

Leurs, U., Mistarz, U. H. & Rand, K. D. 2015. Getting to the core of protein pharmaceuticals--Comprehensive structure analysis by mass spectrometry. *Eur J Pharm Biopharm,* 93**,** 95-109.

Lim, X. X., Chandramohan, A., Lim, X. Y., Bag, N., Sharma, K. K., Wirawan, M., Wohland, T., Lok, S. M. & Anand, G. S. 2017. Conformational changes in intact dengue virus reveal serotype-specific expansion. *Nat Commun,* 8**,** 14339.

Lindesmith, L. C., Donaldson, E. F., Beltramello, M., Pintus, S., Corti, D., Swanstrom, J., Debbink, K., Jones, T. A., Lanzavecchia, A. & Baric, R. S. 2014. Particle conformation regulates antibody access to a conserved GII.4 norovirus blockade epitope. *J Virol,* 88**,** 8826-42.

Lindesmith, L. C., Mallory, M. L., Debbink, K., Donaldson, E. F., Brewer-Jensen, P. D., Swann, E. W., Sheahan, T. P., Graham, R. L., Beltramello, M., Corti, D., Lanzavecchia, A. & Baric, R. S. 2018. Conformational Occlusion of Blockade Antibody Epitopes, a Novel Mechanism of GII.4 Human Norovirus Immune Evasion. *mSphere,* 3.

Lindsay, L., Wolter, J., De Coster, I., Van Damme, P. & Verstraeten, T. 2015. A decade of norovirus disease risk among older adults in upper-middle and high income countries: a systematic review. *Bmc Infectious Diseases,* 15.

Loo, J. A. 1995. Observation of Large Subunit Protein Complexes by Electrospray-Ionization Mass-Spectrometry. *Journal of Mass Spectrometry,* 30**,** 180-183.

Lundberg, M. E., Becker, E. C. & Choe, S. 2013. MstX and a putative potassium channel facilitate biofilm formation in Bacillus subtilis. *PLoS One,* 8**,** e60993.

Luo, Y. & Muesing, M. A. 2014. Mass spectrometry-based proteomic approaches for discovery of HIV-host interactions. *Future Virol,* 9**,** 979-992.

Makarov, A. 2000. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem,* 72**,** 1156-62.

Maleknia, S. D., Ralston, C. Y., Brenowitz, M. D., Downard, K. M. & Chance, M. R. 2001. Determination of macromolecular folding and structure by synchrotron X-ray radiolysis techniques. *Analytical Biochemistry,* 289**,** 103-115.

Mallagaray, A., Creutznacher, R., Dulfer, J., Mayer, P. H. O., Grimm, L. L., Orduna, J. M., Trabjerg, E., Stehle, T., Rand, K. D., Blaum, B. S., Uetrecht, C. & Peters, T.

2019. A post-translational modification of human Norovirus capsid protein attenuates glycan binding. *Nat Commun,* 10**,** 1320.

Mallagaray, A., Lockhauserbäumer, J., Hansman, G., Uetrecht, C. & Peters, T. 2015. Attachment of Norovirus to Histo Blood Group Antigens: A Cooperative Multistep Process. *Angewandte Chemie-International Edition,* 54**,** 12014-12019.

Mallagaray, A., Rademacher, C., Parra, F., Hansman, G. & Peters, T. 2016. Saturation transfer difference nuclear magnetic resonance titrations reveal complex multistep-binding of l-fucose to norovirus particles. *Glycobiology.*

Mallory, M. L., Lindesmith, L. C., Graham, R. L. & Baric, R. S. 2019. GII.4 Human Norovirus: Surveying the Antigenic Landscape. *Viruses,* 11.

Mamyrin, B. A., Karataev, V. I., Shmikk, D. V. & Zagulin, V. A. 1973. Mass-Reflectron a New Nonmagnetic Time-of-Flight High-Resolution Mass-Spectrometer. *Zhurnal Eksperimentalnoi I Teoreticheskoi Fiziki,* 64**,** 82-89.

Marcsisin, S. R. & Engen, J. R. 2010. Hydrogen exchange mass spectrometry: what is it and what can it tell us? *Analytical and Bioanalytical Chemistry,* 397**,** 967-972.

Marklund, E. G. & Benesch, J. L. 2019. Weighing-up protein dynamics: the combination of native mass spectrometry and molecular dynamics simulations. *Curr Opin Struct Biol,* 54**,** 50-58.

Martens, C. & Politis, A. 2020. A glimpse into the molecular mechanism of integral membrane proteins through hydrogen-deuterium exchange mass spectrometry. *Protein Sci.*

Martens, C., Shekhar, M., Borysik, A. J., Lau, A. M., Reading, E., Tajkhorshid, E., Booth, P. J. & Politis, A. 2018. Direct protein-lipid interactions shape the conformational landscape of secondary transporters. *Nat Commun,* 9**,** 4151.

Martens, C., Shekhar, M., Lau, A. M., Tajkhorshid, E. & Politis, A. 2019. Integrating hydrogen-deuterium exchange mass spectrometry with molecular dynamics simulations to probe lipid-modulated conformational changes in membrane proteins. *Nat Protoc,* 14**,** 3183-3204.

Masson, G. R., Burke, J. E., Ahn, N. G., Anand, G. S., Borchers, C., Brier, S., Bou-Assaf, G. M., Engen, J. R., Englander, S. W., Faber, J., Garlish, R., Griffin, P. R., Gross, M. L., Guttman, M., Hamuro, Y., Heck, A. J. R., Houde, D., Iacob, R. E., Jorgensen, T. J. D., Kaltashov, I. A., Klinman, J. P., Konermann, L., Man, P., Mayne, L., Pascal, B. D., Reichmann, D., Skehel, M., Snijder, J., Strutzenberg, T. S., Underbakke, E. S., Wagner, C., Wales, T. E., Walters, B. T., Weis, D. D., Wilson, D. J., Wintrode, P. L., Zhang, Z., Zheng, J., Schriemer, D. C. & Rand, K. D. 2019. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods,* 16**,** 595-602.

Mcallister, R. G., Metwally, H., Sun, Y. & Konermann, L. 2015. Release of native-like gaseous proteins from electrospray droplets via the charged residue mechanism:

# References

insights from molecular dynamics simulations. *Journal of the American Chemical Society,* 137**,** 12667-12676.

Mcfadden, P. N. & Clarke, S. 1987. Conversion of isoaspartyl peptides to normal peptides: implications for the cellular repair of damaged proteins. *Proceedings of the National Academy of Sciences,* 84**,** 2595.

Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M. I., Pragani, R., Boxer, M. B., Earl, L. A., Milne, J. L. S. & Subramaniam, S. 2016. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell,* 165**,** 1698-1707.

Merkle, P. S., Gotfryd, K., Cuendet, M. A., Leth-Espensen, K. Z., Gether, U., Loland, C. J. & Rand, K. D. 2018. Substrate-modulated unwinding of transmembrane helices in the NSS transporter LeuT. *Science Advances,* 4.

Mistarz, U. H., Bellina, B., Jensen, P. F., Brown, J. M., Barran, P. E. & Rand, K. D. 2018. UV Photodissociation Mass Spectrometry Accurately Localize Sites of Backbone Deuteration in Peptides. *Anal Chem,* 90**,** 1077-1080.

Moller, I. R., Merkle, P. S., Calugareanu, D., Comamala, G., Schmidt, S. G., Loland, C. J. & Rand, K. D. 2020. Probing the conformational impact of detergents on the integral membrane protein LeuT by global HDX-MS. *J Proteomics***,** 103845.

Moller, I. R., Slivacka, M., Hausner, J., Nielsen, A. K., Pospisilova, E., Merkle, P. S., Liskova, R., Polak, M., Loland, C. J., Kadek, A., Man, P. & Rand, K. D. 2019. Improving the Sequence Coverage of Integral Membrane Proteins during Hydrogen/Deuterium Exchange Mass Spectrometry Experiments. *Anal Chem,* 91**,** 10970-10978.

Morgan, C. R. & Engen, J. R. 2009. Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. *Curr Protoc Protein Sci,* Chapter 17**,** Unit 17 6 1-17.

Muller, M. M. 2017. Post-Translational Modifications of Protein Backbones: Unique Functions, Mechanisms, and Challenges. *Biochemistry.*

Nasir, W., Frank, M., Kunze, A., Bally, M., Parra, F., Nyholm, P. G., Hook, F. & Larson, G. 2017. Histo-Blood Group Antigen Presentation Is Critical for Binding of Norovirus VLP to Glycosphingolipids in Model Membranes. *ACS Chem Biol,* 12**,** 1288-1296.

Nelson, C. A., Wilen, C. B., Dai, Y. N., Orchard, R. C., Kim, A. S., Stegeman, R. A., Hsieh, L. L., Smith, T. J., Virgin, H. W. & Fremont, D. H. 2018. Structural basis for murine norovirus engagement of bile acids and the CD300lf receptor. *Proc Natl Acad Sci U S A,* 115**,** E9201-E9210.

Nguyen, T. T., Sabat, G. & Sussman, M. R. 2018. In vivo cross-linking supports a head-to-tail mechanism for regulation of the plant plasma membrane P-type H(+)-ATPase. *J Biol Chem,* 293**,** 17095-17106.

Noguchi, S., Miyawaki, K. & Satow, Y. 1998. Succinimide and isoaspartate residues in the crystal structures of hen egg-white lysozyme complexed with tri-N-acetylchitotriose11Edited by R. Huber. *Journal of Molecular Biology,* 278**,** 231-238.

Nolting, D., Malek, R. & Makarov, A. 2017. Ion traps in modern mass spectrometry. *Mass Spectrom Rev.*

O'brien, D. P., Hourdel, V., Chenal, A. & Brier, S. 2020. Hydrogen/Deuterium Exchange Mass Spectrometry for the Structural Analysis of Detergent-Solubilized Membrane Proteins. *Methods Mol Biol,* 2127**,** 339-358.

Olsen, J. V., Macek, B., Lange, O., Makarov, A., Horning, S. & Mann, M. 2007. Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods,* 4**,** 709-12.

Orchard, R. C., Wilen, C. B., Doench, J. G., Baldridge, M. T., Mccune, B. T., Lee, Y. C., Lee, S., Pruett-Miller, S. M., Nelson, C. A., Fremont, D. H. & Virgin, H. W. 2016. Discovery of a proteinaceous cellular receptor for a norovirus. *Science,* 353**,** 933-6.

Pagel, K., Hyung, S. J., Ruotolo, B. T. & Robinson, C. V. 2010. Alternate dissociation pathways identified in charge-reduced protein complex ions. *Anal Chem,* 82**,** 5363-72.

Pan, Y. & Konermann, L. 2010. Membrane protein structural insights from chemical labeling and mass spectrometry. *Analyst,* 135**,** 1191-200.

Parveen, N., Rimkute, I., Block, S., Rydell, G. E., Midtvedt, D., Larson, G., Hytonen, V. P., Zhdanov, V. P., Lundgren, A. & Hook, F. 2018. Membrane Deformation Induces Clustering of Norovirus Bound to Glycosphingolipids in a Supported Cell-Membrane Mimic. *Journal of Physical Chemistry Letters,* 9**,** 2278-2284.

Perry, R. H., Cooks, R. G. & Noll, R. J. 2008. Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass Spectrom Rev,* 27**,** 661-99.

Persson, B. D., Reiter, D. M., Marttila, M., Mei, Y. F., Casasnovas, J. M., Arnberg, N. & Stehle, T. 2007. Adenovirus type 11 binding alters the conformation of its receptor CD46. *Nat Struct Mol Biol,* 14**,** 164-6.

Plotnikov, N. V., Singh, S. K., Rouse, J. C. & Kumar, S. 2017. Quantifying the Risks of Asparagine Deamidation and Aspartate Isomerization in Biopharmaceuticals by Computing Reaction Free-Energy Surfaces. *J Phys Chem B,* 121**,** 719-730.

Pluskal, T., Castillo, S., Villar-Briones, A. & Oresic, M. 2010. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics,* 11**,** 395.

Pogan, R., Dulfer, J. & Uetrecht, C. 2018a. Norovirus assembly and stability. *Curr Opin Virol,* 31**,** 59-65.

Pogan, R., Schneider, C., Reimer, R., Hansman, G. & Uetrecht, C. 2018b. Norovirus-like VP1 particles exhibit isolate dependent stability profiles. *J Phys Condens Matter,* 30**,** 064006.

Politis, A. & Schmidt, C. 2018. Structural characterisation of medically relevant protein assemblies by integrating mass spectrometry with computational modelling. *J Proteomics,* 175**,** 34-41.

Prasad, B. V., Hardy, M. E., Dokland, T., Bella, J., Rossmann, M. G. & Estes, M. K. 1999. X-ray crystallographic structure of the Norwalk virus capsid. *Science,* 286**,** 287-90.

Prasad, B. V., Rothnagel, R., Jiang, X. & Estes, M. K. 1994. Three-dimensional structure of baculovirus-expressed Norwalk virus capsids. *J Virol,* 68**,** 5117-25.

Prasad, B. V. & Schmid, M. F. 2012. Principles of virus structural organization. *Adv Exp Med Biol,* 726**,** 17-47.

Preston, W. C. 1948. Some Correlating Principles of Detergent Action. *Journal of Physical and Colloid Chemistry,* 52**,** 84-97.

Psachoulia, E., Bond, P. J. & Sansom, M. S. 2006. MD simulations of Mistic: conformational stability in detergent micelles and water. *Biochemistry,* 45**,** 9053-8.

Rand, K. D. 2013. Pinpointing changes in higher-order protein structure by hydrogen/deuterium exchange coupled to electron transfer dissociation mass spectrometry. *International Journal of Mass Spectrometry,* 338**,** 2-10.

Reading, E., Hall, Z., Martens, C., Haghighi, T., Findlay, H., Ahdash, Z., Politis, A. & Booth, P. J. 2017. Interrogating Membrane Protein Conformational Dynamics within Native Lipid Compositions. *Angewandte Chemie-International Edition,* 56**,** 15654-15657.

Remmert, M., Biegert, A., Hauser, A. & Soding, J. 2011. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods,* 9**,** 173-5.

Rey, M., Man, P., Brandolin, G., Forest, E. & Pelosi, L. 2009. Recombinant immobilized rhizopuspepsin as a new tool for protein digestion in hydrogen/deuterium exchange mass spectrometry. *Rapid Commun Mass Spectrom,* 23**,** 3431-8.

Riggs, D. L., Gomez, S. V. & Julian, R. R. 2017. Sequence and Solution Effects on the Prevalence of d-Isomers Produced by Deamidation. *ACS Chem Biol,* 12**,** 2875-2882.

Robinson, A. B., Mckerrow, J. H. & Cary, P. 1970. Controlled Deamidation of Peptides and Proteins: An Experimental Hazard and a Possible Biological Timer. *Proceedings of the National Academy of Sciences,* 66**,** 753-757.

Robinson, N. E. & Robinson, A. B. 2001a. Deamidation of human proteins. *Proceedings of the National Academy of Sciences,* 98**,** 12409-12413.

Robinson, N. E. & Robinson, A. B. 2001b. Molecular clocks. *Proceedings of the National Academy of Sciences,* 98**,** 944-949.

Robinson, N. E. & Robinson, A. B. 2001c. Prediction of protein deamidation rates from primary and three-dimensional structure. *Proceedings of the National Academy of Sciences,* 98**,** 4367-4372.

Roosild, T. P., Greenwald, J., Vega, M., Castronovo, S., Riek, R. & Choe, S. 2005. NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science,* 307**,** 1317-21.

Roosild, T. P., Vega, M., Castronovo, S. & Choe, S. 2006. Characterization of the family of Mistic homologues. *BMC Struct Biol,* 6**,** 10.

Root, K., Wittwer, Y., Barylyuk, K., Anders, U. & Zenobi, R. 2017. Insight into Signal Response of Protein Ions in Native ESI-MS from the Analysis of Model Mixtures of Covalently Linked Protein Oligomers. *J Am Soc Mass Spectrom.*

Rose, R. J., Labrijn, A. F., Van Den Bremer, E. T., Loverix, S., Lasters, I., Van Berkel, P. H., Van De Winkel, J. G., Schuurman, J., Parren, P. W. & Heck, A. J. 2011. Quantitative analysis of the interaction strength and dynamics of human IgG4 half molecules by native mass spectrometry. *Structure,* 19**,** 1274-82.

Rossi, V., Gaboriaud, C., Lacroix, M., Ulrich, J., Fontecillacamps, J. C., Gagnon, J. & Arlaud, G. J. 1995. Structure of the Catalytic Region of Human-Complement Protease C(1)over-Bar-S - Study by Chemical Cross-Linking and 3-Dimensional Homology Modeling. *Biochemistry,* 34**,** 7311-7321.

Rostom, A. A. & Robinson, C. V. 1999. Detection of the intact GroEL chaperonin assembly by mass spectrometry. *Journal of the American Chemical Society,* 121**,** 4718-4719.

Ruotolo, B. T., Giles, K., Campuzano, I., Sandercock, A. M., Bateman, R. H. & Robinson, C. V. 2005. Evidence for macromolecular protein rings in the absence of bulk water. *Science,* 310**,** 1658-61.

Rydell, G. E., Svensson, L., Larson, G., Johannes, L. & Romer, W. 2013. Human GII.4 norovirus VLP induces membrane invaginations on giant unilamellar vesicles containing secretor gene dependent alpha1,2-fucosylated glycosphingolipids. *Biochim Biophys Acta,* 1828**,** 1840-5.

Sansom, M. S., Scott, K. A. & Bond, P. J. 2008. Coarse-grained simulation: a high-throughput computational approach to membrane proteins. *Biochem Soc Trans,* 36**,** 27-32.

Schroten, H., Hanisch, F. G. & Hansman, G. S. 2016. Human Norovirus Interactions with Histo-Blood Group Antigens and Human Milk Oligosaccharides. *J Virol,* 90**,** 5855-9.

Seo, J., Hoffmann, W., Warnke, S., Bowers, M. T., Pagel, K. & Von Helden, G. 2016. Retention of Native Protein Structures in the Absence of Solvent: A Coupled Ion Mobility and Spectroscopic Study. *Angew Chem Int Ed Engl,* 55**,** 14173-14176.

References

Shanker, S., Choi, J. M., Sankaran, B., Atmar, R. L., Estes, M. K. & Prasad, B. V. 2011. Structural analysis of histo-blood group antigen binding specificity in a norovirus GII.4 epidemic variant: implications for epochal evolution. *J Virol,* 85**,** 8635-45.

Sherman, M. B., Smith, H. Q. & Smith, T. J. 2020. The Dynamic Life of Virus Capsids. *Viruses,* 12.

Sherman, M. B., Williams, A. N., Smith, H. Q., Nelson, C., Wilen, C. B., Fremont, D. H., Virgin, H. W. & Smith, T. J. 2019. Bile salts alter the mouse norovirus capsid conformation; possible implications for cell attachment and immune evasion. *J Virol.*

Shoemaker, G. K., Van Duijn, E., Crawford, S. E., Uetrecht, C., Baclayon, M., Roos, W. H., Wuite, G. J., Estes, M. K., Prasad, B. V. & Heck, A. J. 2010. Norwalk virus assembly and stability monitored by mass spectrometry. *Mol Cell Proteomics,* 9**,** 1742-51.

Singh, B. K., Koromyslova, A., Hefele, L., Gurth, C. & Hansman, G. S. 2015a. Structural Evolution of the Emerging 2014-2015 GII.17 Noroviruses. *J Virol,* 90**,** 2710-5.

Singh, B. K., Leuthold, M. M. & Hansman, G. S. 2015b. Human noroviruses' fondness for histo-blood group antigens. *J Virol,* 89**,** 2024-40.

Siuzdak, G., Bothner, B., Yeager, M., Brugidou, C., Fauquet, C. M., Hoey, K. & Chang, C. M. 1996. Mass spectrometry and viral analysis. *Chem Biol,* 3**,** 45-8.

Smith, D. L., Deng, Y. & Zhang, Z. 1997. Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J Mass Spectrom,* 32**,** 135-46.

Smith, H. Q. & Smith, T. J. 2019. The Dynamic Capsid Structures of the Noroviruses. *Viruses,* 11.

Snijder, J., Uetrecht, C., Rose, R. J., Sanchez-Eugenia, R., Marti, G. A., Agirre, J., Guerin, D. M. A., Wuite, G. J. L., Heck, A. J. R. & Roos, W. H. 2013. Probing the biophysical interplay between a viral genome and its capsid. *Nature Chemistry,* 5**,** 502-509.

Snowden, J. S., Hurdiss, D. L., Adeyemi, O. O., Ranson, N. A., Herod, M. R. & Stonehouse, N. J. 2020. Dynamics in the murine norovirus capsid revealed by high-resolution cryo-EM. *PLoS Biol,* 18**,** e3000649.

Song, C., Takai-Todaka, R., Miki, M., Haga, K., Fujimoto, A., Ishiyama, R., Oikawa, K., Yokoyama, M., Miyazaki, N., Iwasaki, K., Murakami, K., Katayama, K. & Murata, K. 2020. Dynamic rotation of the protruding domain enhances the infectivity of norovirus. *PLoS Pathog,* 16**,** e1008619.

Stehle, T. & Casasnovas, J. M. 2009. Specificity switching in virus-receptor complexes. *Curr Opin Struct Biol,* 19**,** 181-8.

Stiving, A. Q., Vanaernum, Z. L., Busch, F., Harvey, S. R., Sarni, S. H. & Wysocki, V. H. 2019. Surface-Induced Dissociation: An Effective Method for Characterization of Protein Quaternary Structure. *Analytical Chemistry,* 91**,** 190-209.

Ströh, L. J. & Stehle, T. 2014. Glycan Engagement by Viruses: Receptor Switches and Specificity. *Annu Rev Virol,* 1**,** 285-306.

Tan, M., Fang, P., Chachiyo, T., Xia, M., Huang, P., Fang, Z., Jiang, W. & Jiang, X. 2008. Noroviral P particle: structure, function and applications in virus-host interaction. *Virology,* 382**,** 115-23.

Tan, M. & Jiang, X. 2005. The p domain of norovirus capsid protein forms a subviral particle that binds to histo-blood group antigen receptors. *J Virol,* 79**,** 14017-30.

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T. & Matsuo, T. 1988. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid communications in mass spectrometry,* 2**,** 151-153.

Taube, S., Kolawole, A. O., Hohne, M., Wilkinson, J. E., Handley, S. A., Perry, J. W., Thackray, L. B., Akkina, R. & Wobus, C. E. 2013. A mouse model for human norovirus. *mBio,* 4.

Taube, S., Mallagaray, A. & Peters, T. 2018. Norovirus, glycans and attachment. *Curr Opin Virol,* 31**,** 33-42.

Taube, S., Perry, J. W., Mcgreevy, E., Yetming, K., Perkins, C., Henderson, K. & Wobus, C. E. 2012. Murine noroviruses bind glycolipid and glycoprotein attachment receptors in a strain-dependent manner. *J Virol,* 86**,** 5584-93.

Taube, S., Rubin, J. R., Katpally, U., Smith, T. J., Kendall, A., Stuckey, J. A. & Wobus, C. E. 2010. High-resolution x-ray structure and functional analysis of the murine norovirus 1 capsid protein protruding domain. *J Virol,* 84**,** 5695-705.

Textor, M. 2016. *Reconstitution and Membrane Topology of Mistic from Bacillus subtilis.* PhD Thesis, University Kaiserlautern.

Thiele, C. & Huttner, W. B. 1998. The disulfide-bonded loop of chromogranins, which is essential for sorting to secretory granules, mediates homodimerization. *Journal of Biological Chemistry,* 273**,** 1223-1231.

Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J. M., Taly, J. F. & Notredame, C. 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Research,* 39**,** W13-W17.

Trabjerg, E., Nazari, Z. E. & Rand, K. D. 2018. Conformational analysis of complex protein states by hydrogen/deuterium exchange mass spectrometry (HDX-MS): Challenges and emerging solutions. *Trac-Trends in Analytical Chemistry,* 106**,** 125-138.

Tuan Zainazor, C., Hidayah, M. S., Chai, L. C., Tunung, R., Ghazali, F. M. & Son, R. 2010. The scenario of norovirus contamination in food and food handlers. *J Microbiol Biotechnol,* 20**,** 229-37.

Tubiana, T., Boulard, Y. & Bressanelli, S. 2017. Dynamics and asymmetry in the dimer of the norovirus major capsid protein. *PLoS One,* 12**,** e0182056.

References

Tyagi, S. & Lemke, E. A. 2013. Genetically encoded click chemistry for single-molecule FRET of proteins. *Methods Cell Biol,* 113**,** 169-87.

Tyanova, S., Temu, T. & Cox, J. 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc,* 11**,** 2301-2319.

Tyler-Cross, R. & Schirch, V. 1991. Effects of amino acid sequence, buffers, and ionic strength on the rate and mechanism of deamidation of asparagine residues in small peptides. *J Biol Chem,* 266**,** 22549-56.

Uetrecht, C., Rose, R. J., Van Duijn, E., Lorenzen, K. & Heck, A. J. 2010. Ion mobility mass spectrometry of proteins and protein assemblies. *Chem Soc Rev,* 39**,** 1633-55.

Uetrecht, C., Versluis, C., Watts, N. R., Wingfield, P. T., Steven, A. C. & Heck, A. J. 2008. Stability and shape of hepatitis B virus capsids in vacuo. *Angew Chem Int Ed Engl,* 47**,** 6247-51.

Van Berkel, W. J. H., Van Den Heuvel, R. H. H., Versluis, C. & Heck, A. J. R. 2000. Detection of intact megaDalton protein assemblies of vanillyl-alcohol oxidase by mass spectrometry. *Protein Science,* 9**,** 435-439.

Van De Waterbeemd, M., Llauro, A., Snijder, J., Valbuena, A., Rodriguez-Huete, A., Fuertes, M. A., De Pablo, P. J., Mateu, M. G. & Heck, A. J. 2017. Structural Analysis of a Temperature-Induced Transition in a Viral Capsid Probed by HDX-MS. *Biophys J,* 112**,** 1157-1165.

Van Den Heuvel, R. H., Van Duijn, E., Mazon, H., Synowsky, S. A., Lorenzen, K., Versluis, C., Brouns, S. J., Langridge, D., Van Der Oost, J., Hoyes, J. & Heck, A. J. 2006. Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal Chem,* 78**,** 7473-83.

Vinje, J. 2015. Advances in laboratory methods for detection and typing of norovirus. *J Clin Microbiol,* 53**,** 373-81.

Vizcaino, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q. W., Wang, R. & Hermjakob, H. 2016. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res,* 44**,** D447-56.

Wales, T. E., Eggertson, M. J. & Engen, J. R. 2013. Considerations in the analysis of hydrogen exchange mass spectrometry data. *Methods Mol Biol,* 1007**,** 263-88.

Wang, L., Pan, H. & Smith, D. L. 2002. Hydrogen Exchange-Mass Spectrometry. *Molecular & Cellular Proteomics,* 1**,** 132-138.

Wang, L. C., Krishnamurthy, S. & Anand, G. S. 2016. Hydrogen exchange mass spectrometry experimental design. *In:* WEIS, D. D. (ed.) *Hydrogen Exchange Mass Spectrometry of Proteins: Fundamentals, Methods, and Applications.* John Wiley & Sons.

Ward, A. B., Sali, A. & Wilson, I. A. 2013. Biochemistry. Integrative structural biology. *Science,* 339**,** 913-5.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., De Beer, T. a. P., Rempfer, C., Bordoli, L., Lepore, R. & Schwede, T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res,* 46**,** W296-W303.

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* 25**,** 1189-91.

Wegener, H., Mallagaray, A., Schone, T., Peters, T., Lockhauserbaumer, J., Yan, H., Uetrecht, C., Hansman, G. S. & Taube, S. 2017. Human norovirus GII.4(MI001) P dimer binds fucosylated and sialylated carbohydrates. *Glycobiology,* 27**,** 1027-1037.

Weichert, S., Koromyslova, A., Singh, B. K., Hansman, S., Jennewein, S., Schroten, H. & Hansman, G. S. 2016. Structural Basis for Norovirus Inhibition by Human Milk Oligosaccharides. *J Virol,* 90**,** 4843-8.

Weis, D. D., Engen, J. R. & Kass, I. J. 2006a. Semi-automated data processing of hydrogen exchange mass spectra using HX-Express. *J Am Soc Mass Spectrom,* 17**,** 1700-3.

Weis, D. D., Wales, T. E., Engen, J. R., Hotchko, M. & Ten Eyck, L. F. 2006b. Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis. *J Am Soc Mass Spectrom,* 17**,** 1498-509.

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D. & Hochstrasser, D. F. 1999. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol,* 112**,** 531-52.

Wobus, C. E., Karst, S. M., Thackray, L. B., Chang, K. O., Sosnovtsev, S. V., Belliot, G., Krug, A., Mackenzie, J. M., Green, K. Y. & Virgin, H. W. 2004. Replication of Norovirus in cell culture reveals a tropism for dendritic cells and macrophages. *PLoS Biol,* 2**,** e432.

Wobus, C. E., Thackray, L. B. & Virgin, H. W. 2006. Murine Norovirus: a Model System To Study Norovirus Biology and Pathogenesis. *Journal of Virology,* 80**,** 5104-5112.

Woods, R. J. *Woods Group. (2005-2020) GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA. (http://glycam.org)* [Online]. [Accessed 27.05.2020].

Xi, J. A., Min, W., Graham, D. Y. & Estes, M. K. 1992. Expression, Self-Assembly, and Antigenicity of the Norwalk Virus Capsid Protein. *Journal of Virology,* 66**,** 6527-6532.

Xi, J. N., Graham, D. Y., Wang, K. N. & Estes, M. K. 1990. Norwalk virus genome cloning and characterization. *Science,* 250**,** 1580-3.

# References

Xiao, H., Hoerner, J. K., Eyles, S. J., Dobo, A., Voigtman, E., Mel'cuk, A. I. & Kaltashov, I. A. 2005. Mapping protein energy landscapes with amide hydrogen exchange and mass spectrometry: I. A generalized model for a two-state protein and comparison with experiment. *Protein Sci,* 14**,** 543-57.

Yamashita, M. & Fenn, J. B. 1984. Electrospray Ion-Source - Another Variation on the Free-Jet Theme. *Journal of Physical Chemistry,* 88**,** 4451-4459.

Yan, Y., Wei, H., Fu, Y., Jusuf, S., Zeng, M., Ludwig, R., Krystek, S. R., Jr., Chen, G., Tao, L. & Das, T. K. 2016. 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**,** 2041-50.

Yang, M. L., Hoeppner, M., Rey, M., Kadek, A., Man, P. & Schriemer, D. C. 2015. Recombinant Nepenthesin II for Hydrogen/Deuterium Exchange Mass Spectrometry. *Analytical Chemistry,* 87**,** 6681-6687.

Yao, L., Li, F. L., Wang, L. Z., Zhai, Y. X. & Jiang, Y. H. 2014. Function of VP2 Protein in the Stability of the Secondary Structure of Virus-like Particles of Genogroup II Norovirus at Different pH Levels: Function of VP2 Protein in the Stability of NoV VLPs. *Journal of Microbiology,* 52**,** 970-975.

Yates, J. R., Ruse, C. I. & Nakorchevsky, A. 2009. Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annual Review of Biomedical Engineering,* 11**,** 49-79.

Zheng, D. P., Ando, T., Fankhauser, R. L., Beard, R. S., Glass, R. I. & Monroe, S. S. 2006. Norovirus classification and proposed strain nomenclature. *Virology,* 346**,** 312-23.

Zubarev, R. A. & Makarov, A. 2013. Orbitrap mass spectrometry. *Anal Chem,* 85**,** 5288-96.

## ACKNOWLEDGEMENT

When I first entered the field of mass spectrometry in early 2016, I could not have imagined which kind of path would lay ahead of me. I knew basically nothing about mass spectrometry and had no idea what "HDX-MS" stood for, just that it would be the main technique of my PhD project. Now, after almost 4.5 years, I have successfully established HDX-MS as a new technique in our laboratory, published in scientific journals and given talks at international conferences. What a journey.

First, I would like to thank my supervisor Dr. Charlotte Uetrecht for entrusting me with this project and giving me the opportunity to grow so much both professionally and personally. I would also like to thank my second supervisor Prof. Zoya Ignatova from the University of Hamburg for useful advice and discussions. Many thanks goes also to Prof. Hartmut Schlüter at the UKE, who gave me access to mass spectrometers and acted as the second reviewer of my thesis.

I am very glad of having worked in close collaboration with many motivated PhD students and Postdocs in the DFG research group ViroCarb. Here, special thanks go to my closest collaborators Prof. Thomas Peters, Dr. Alvaro Mallagaray and Robert Creutznacher from the University of Lübeck who not only provided numerous protein samples and complementing NMR data, but also engaged in frequent discussions and were always open for questions. I would also like to thank Prof. Stefan Taube and Jan Knickmann from the University of Lübeck for helpful discussions about murine norovirus cell biology and production of murine norovirus particles. Thanks also to Domink Pahl and Laura Soria Martinez from the University of Münster, for all the fun we had at the ViroCarb PhD symposia and while working with human papillomaviruses. Great thanks also go to Prof. Michael Schlierf and Marta Batet from the TU Dresden for the great collaboration on the Mistic project.

At the beginning of my PhD, the task of establishing HDX-MS as a new technique in our lab seemed to be a huge challenge, had it not been for the numerous people who supported me during the whole time. First, Prof. Kasper Rand and Esben Trabjerg from the University of Copenhagen must be named, where I spend two months for HDX-MS training and method development. Under your initial guidance, I transformed from a mass spectrometry newbie into a HDX-MS scientist. I am more than thankful for your still ongoing support as well as our

## DECLARATION OF AUTHORSHIP / EIDESSTATTLICHE VERSICHERUNG

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids. The submitted written document corresponds to the file on the electronic storage medium. I further declare that this thesis has not been presented previously to another examination board.

Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.

Hamburg, 17 August 2020

_____

Signature