UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Hydrogen/Deuterium Exchange Mass Spectrometry (DXMS) Analysis of the Human Carbohydrate Phosphatase, Laforin /

Permalink

https://escholarship.org/uc/item/36n7s8tw

Author

Wong, Brian Kar

Publication Date

2013

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Hydrogen/Deuterium Exchange Mass Spectrometry (DXMS) Analysis of the Human Carbohydrate Phosphatase, Laforin

A Thesis submitted in partial satisfaction of the requirements for the degree

Master of Science

in

Biology

by

Brian Kar Wong

Committee in charge:

Professor Darren Casteel, Chair
Professor Randolph Hampton, Co-Chair
Professor Lakshmi Chilukuri

2013

The Thesis of Brian Kar Wong is approved and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2013

DEDICATION

I would like dedicate this Thesis to my family:

To my father, Clinton, for teaching me the value of hard work and perseverance.

To my mother, Christina, for your constant support and love.

To my sister, Angela, for filling my life with laughter.

TABLE OF CONTENTS

Signature Page ……..... iii

Dedication …… iv

Table of Contents ……..… v

List of Figures ……..…… vi

List of Tables ……... vii

List of Supplemental Figures ……... viii

Acknowledgements ……..…… ix

Abstract ……..... x

Introduction ……...… 1

Materials and Methods ……..… 8

Results ……...……... 14

Discussion ……...……. 23

Figures ……...……... 29

Tables ……... 42

Supplemental Figures ……... 43

References ……...…...…..... 57

LIST OF FIGURES

Figure 1: Three main groups of exchangeable hydrogens on a protein …….. 29

Figure 2: pH effects on amide hydrogen/deuterium exchange…….. 30

Figure 3: Overview of a DXMS experiment ……. 31

Figure 4: Laforin schematic and proposed mechanism …….. 32

Figure 5: Representative models of glycogen and amylopectin ……. 33

Figure 6: Pepsin-digested coverage map of wild-type Laforin ……... 34

Figure 7: DXMS results for apo and glycogen-bound wild-type Laforin …….. 35

Figure 8: Wild-type Laforin DSP domain peptides that exhibited a decrease in deuterium incorporation for apo vs. glycogen-bound conditions …… 36

Figure 9: Deuterium incorporation changes in wild-type Laforin upon glucan binding ……. 37

Figure 10: Deuterium incorporation changes of W32G-Laforin in the presence of amylopectin …….. 38

Figure 11: Deuterium incorporation changes in G240S-Laforin upon glycogen binding ……. 39

Figure 12: Deuterium incorporation changes in Y294N-Laforin upon glycogen binding ……. 40

Figure 13: Differences in deuteration of DSP peptide $R^{290} – L^{296}$ between wild-type Laforin and Y294N-Laforin …… 41

LIST OF TABLES

Table 1: Biochemical and physical properties of eukaryotic glycogen, amylopectin, and Lafora Bodies …….. 42

Table 2: Average deuteration level (%) for all Laforin proteins ……. 42

LIST OF SUPPLEMENTAL FIGURES

Supplemental Figure 1: Time dependent deuteration levels for glycogen-bound (pink squares) and ligand-free (blue diamonds) wild-type Laforin peptides…… 43

ACKNOWLEDGEMENTS

Very few people had the opportunity to see Dr. Virgil Woods, Jr.'s passion and commitment to all aspects of his life. I was very fortunate to have known Dr. Woods and seen some of his many roles: doctor, researcher, leader, teacher, and mentor. His enthusiasm and drive for excellence has no doubt been passed on to his many medical students and graduate students. Although Dr. Woods will be heralded as a pioneer in the field of proteomics, I will remember him more for his lessons about life, his ability to spark curiosity, and his comedic wit. Dr. Woods' passing was much too early and disheartening to acknowledge, but I am comforted by the fact that his legacy will carry on through the numerous lives he influenced. Dr. Woods, you will be missed.

I would like to thank Dr. Darren Casteel for being my committee Chair. I would also like to thank Dr. Randolph Hampton and Dr. Lakshmi Chilukuri for being members of my committee.

This Thesis was made possible through the help of Dr. Matthew Gentry and Amanda Sherwood. I am grateful for all the Laforin purifications and the extremely helpful feedback throughout this project.

Lastly, I would like to thank the many members (past and present) of Woods Lab for their constant support and guidance. Specifically, I am indebted to Sheng Li, Tong Liu, David Lee, and Bryant Kou for the countless hours spent helping me prepare samples and process data.

ABSTRACT OF THE THESIS

Hydrogen/Deuterium Exchange Mass Spectrometry (DXMS) Analysis of the Human Carbohydrate Phosphatase, Laforin

by

Brian Kar Wong

Master of Science in Biology

University of California, San Diego, 2013

Professor Darren Casteel, Chair

Professor Randolph Hampton, Co-Chair

Lafora Disease (LD) is a fatal neurodegenerative disease that is correlated with mutation of the human phosphatase, Laforin. While Laforin has been shown to dephosphorylate phosphoglucans, little is known about the mechanism of dephosphorylation and Laforin's structure. Laforin consists of two domains: a carbohydrate binding module (CBM) and a dual specificity phosphatase (DSP) domain.

The absence of Laforin function results in hyperphosphorylated and poorly branched sugar accumulations. It is hypothesized that these insoluble sugar deposits lead to neurodegeneration and pre-mature death in LD patients.

We utilized hydrogen/deuterium exchange mass spectrometry (DXMS) to define Laforin's structural components and to probe Laforin's substrate interactions. DXMS analysis was performed on wild-type (WT) Laforin and LD mutants: W32G, G240S, and Y294N. The analysis of WT Laforin revealed the strongest substrate interaction with glycogen, when compared to interactions with the carbohydrates amylopectin and β-cyclodextrin. WT data revealed that regions of the CBM were protected from deuteration when bound to glucans. Similarly, structurally important regions of the DSP showed deuteration protection. W32G data confirmed the lack of protection from deuteration in the absence of substrate binding, whereas G240S data confirmed deuteration protection due to substrate binding. Lastly, the Y294N data revealed surprisingly strong substrate interactions in the DSP mutation region, despite weak CBM interaction. Our results confirm the role of Laforin's CBM in glucan binding and highlight the significance of structural DSP elements in the dephosphorylation of phosphoglucans.

INTRODUCTION

1.1 Hydrogen/Deuterium Exchange Mass Spectrometry

Proteins are an essential part of human life and are responsible for many integral biological processes. Despite the identification of many key proteins in the human body, the link between structure and function still remains a substantial mystery for many proteins. It is critical to understand the properties of protein dynamics, structure, and conformational change in order to create pharmaceutical drugs that target the proteins implicated in many human diseases. Although there have been advances in methods of protein analysis in the last few decades, it is extremely difficult or even impossible to create a complete protein picture from one or even a combination of the current methods. For example, techniques such as nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography can provide details about protein structure, but these methods provide little information on protein dynamics and require high quantities of pure protein or high quality crystals[1].

Peptide amide hydrogen/deuterium exchange mass spectrometry (DXMS) has recently been acknowledged as a powerful method of studying the structural dynamics of proteins. The advantages of DXMS include: elimination of the need for crystals by analysis of protein in solution, use of small quantities of protein (500 – 1000 picomoles), and compatibility with proteins that are hard to purify due to precipitation or multimer formation at high concentrations[2]. The structural dynamics of a protein, or how the domains of a protein interact, is vital to a protein's function. The study of a protein's structural dynamics is often hindered when the conformation a protein adopts under X-ray crystallography or NMR conditions differs from the protein conformation under

physiological conditions[1]. The use of DXMS can complement the structural information provided by X-ray crystallography or NMR, by providing evidence about structural dynamics and interaction surfaces[3].

Application of DXMS suggests mechanisms of protein function and the details of substrate interaction, which is essential for understanding the progression of diseases. Recent developments of higher sensitivity mass spectrometers, automated sample preparation, and improved data analysis software have made DXMS a more robust analytical method[3]. DXMS systems have been utilized to study the conformational changes of proteins demonstrated to be involved in human diseases including: viral infections[4], inflammatory diseases[5], and neurodegenerative diseases[6]. In this study, DXMS is applied to study the human phosphatase, Laforin, which is linked to the majority of cases of Lafora Disease.

1.2 Background and Theory of Hydrogen Exchange

DXMS takes advantage of a chemical phenomenon involving protein hydrogen molecules. The hydrogen bonded to the nitrogen of the protein backbone is able to constantly exchange with hydrogen in solution. When a protein is placed into an aqueous solution containing an isotope of hydrogen, such as deuterium, the backbone amide hydrogen may be replaced with the isotope of hydrogen. Since the mass of a deuteron is roughly double the mass of hydrogen, the incorporation of deuterium into a protein can be measured by mass spectrometry. Also, the ability to measure the exchange of

backbone amide hydrogens is advantageous because these hydrogens play critical roles in protein secondary structure.

There are three main groups of hydrogen on a protein that may undergo deuterium exchange. The hydrogen of side chain functional groups (-OH, -SH, -COOH, -NH_2) can undergo exchange, but the incorporated deuterium revert back to hydrogen during processing because of the rapid rate of exchange. The hydrogen attached to carbon atoms almost never exchange and can be assumed to be constant for DXMS experiments[2]. Every amino acid residue, except for Proline, has an exchangeable backbone amide hydrogen, which is monitored by DXMS (**Figure 1**).

Folded proteins under physiological conditions can have amide hydrogen/deuterium exchange rates that vary from milliseconds to months[7]. The exchange rate is influenced by solvent exposure and hydrogen bonding, which can provide valuable details about protein conformation and dynamics[8]. Since amide hydrogens are required to make physical contact with the solvent for exchange[9], regions that are dynamic and exposed to solvent will have an exchange rate on the order of seconds. Regions that are less dynamic, involved in hydrogen bonding, or buried in the hydrophobic interior of a protein may range from hours to weeks for exchange[2]. By measuring the various exchange rates of protein amide hydrogens, details about protein structure and conformation changes can be collected.

1.3 Determining Hydrogen/Deuterium Exchange by Mass Spectrometry

There are two parts to a DXMS experiment: fragmentation tuning and deuterium on-exchange. Fragmentation tuning refers to the process of selecting suitable conditions

for proteolysis of the protein, in order to generate peptide fragments for mass spectrometry detection. Among the conditions to consider are: the specific denaturant and concentration of denaturant in the "quench" solution, reduction of disulfide bonds, types of protease columns, and the flow rate over the protease column. The following represents a general outline of fragmentation tuning. Non-deuterated samples are added to "quench" buffers containing various concentrations of the denaturant, guanidinium hydrochloride. After addition of the "quench" buffer, the samples are injected onto a pepsin protease column and the cleaved peptides are separated by reverse phase liquid chromatography. In order to identify the peptides generated by the pepsin digestion condition, data acquisition proceeds with MS1 profile scan for the parent peptide ions and data dependent MS/MS scans for peptide sequence identifications of fragment ions. The collected data is inputted into the database search algorithm SEQUEST, which determines the parent peptide identities and their chromatographic retention times. The conditions which yield the best sequence coverage and density of peptides are used for deuterium on-exchange.

Deuterium on-exchange provides data on the location and rate of deuterium incorporation after exposure to D_2O buffer at physiological pH. An on-exchange reaction is initiated by placing a protein solution into an all D_2O physiological buffer. The exchange/labeling reaction is allowed to proceed for various timepoints until the reaction is "exchange quenched" by dropping the pH to ~2.5 and the temperature to 0°C. As shown in **Figure 2**, pH 2.5 is the pH where both the acid and base catalyzed exchange reaction is at a minimum[10]. Since the quench conditions stop deuterium exchange, the quenched protein can be digested and the deuterium labeled peptides measured by mass

spectrometry, as described above. The masses of the deuterated peptides are compared to the masses of the same peptides under a non-deuterated condition. By detailing the deuterium incorporation of peptides of a protein over time, regions of interest can be analyzed for protein-protein interactions, substrate binding, and/or structural characteristics. **Figure 3** presents the general workflow of a typical DXMS experiment using automated processing and data analysis software.

1.4 Lafora Disease

Lafora Disease (LD, OMIM #254780) is an autosomal recessive neurodegenerative disorder that results in severe epilepsy and death[11]. Patients with LD develop normally until they present with epileptic seizures in adolescence, followed by rapid neurodegeneration and death within 10 years of the first sign of epilepsy, which places LD in the progressive myoclonus epilepsy family[12,13]. A symptom of LD is the presence of insoluble, irregularly branched, hyperphosphorylated glycogen-like accumulations called Lafora Bodies (LBs) in the cytoplasm of cells in all tissues[14,15]. The presence of these LBs in neurons may contribute to the non-apoptotic neural cell death and seizures observed in LD patients[15,16].

Approximately 50% of LD cases are caused by mutations in the EPM2A gene, which encodes for the phosphatase Laforin[11]. Laforin has the unique ability among vertebrate phosphatases to liberate phosphates from phosphorylated carbohydrates or glucans[16]. Laforin is a bimodular protein with an N-terminal carbohydrate binding module family 20 (CBM 20) domain and a C-terminal dual specificity phosphatase (DSP) domain with the signature HCXXGXXR active site motif [17,18,19]. Dual specificity

phosphatases get their names from their ability to dephosphorylate both phosphotyrosine substrates and phosphoserine/threonine substrates[20]. Although Laforin has been shown to dephosphorylate both phosphotyrosine and phosphoserine/threonine substrates *in vitro*[16], no known proteinaceous substrates have been found. Disease-causing mutations are found on both the DSP and CBM domains (**Figure 4a**), many of which interfere with Laforin's ability to bind and dephosphorylate carbohydrates[21].

The cause of LD remained a mystery until the discovery of Laforin's ability to localize, bind, and remove phosphates from glycogen both *in vitro* and *in vivo*[22]. Glycogen is a branched polymer of glucose produced in the cytoplasm of many animal species and is used as an energy storage molecule. Glycogen is composed of glucose residues joined by α-1,4-glycosidic linkages with branches every 12 – 14 residues via α-1,6-glycosidic linkages[22]. Although similar to glycogen, LBs are poorly branched, hyperphosphorylated, and insoluble in water. Laforin is currently thought to remove phosphates that are present in glycogen as a natural error in glycogen synthesis, permitting normal glycogen branching to occur[23]. Current findings note that tissues with the highest Laforin expression are also tissues with the highest proportion of LBs in LD patients[24]. Furthermore, EPM2A knockout mice models exhibit hyperphosphorylated and decreased branching of glycogen in liver and muscle tissues[22]. These findings support the hypothesis that Laforin is necessary for the removal of aberrant phosphates during glycogen polymerization. Otherwise, hyperphosphorylated glycogen leads to poor branching and eventual formation of LBs (**Figure 4b**).

Although the complete mechanism by which Laforin interacts with and dephosphorylates glycogen remains unsolved, some insights into a possible mechanism is

provided by previous DXMS work on the Laforin functional equivalent in plants, *Arabidopsis* starch excess 4 (SEX4)[25]. Despite not being orthologues, both Laforin and SEX4 have demonstrated the ability to liberate phosphates from the glucan amylopectin, making them the only reported phosphatases with this ability[16]. Even though there are similarities between glycogen and LBs, biochemical studies show that LBs are structurally more similar to amylopectin than any other compound[15] (**Table 1**). Like glycogen, amylopectin forms α-1,4 glycosidic linkages, but branches discontinuously via α-1,6 glycosidic linkages arranged in clusters at regular intervals, rendering it water-insoluble (**Figure 5**). Amylopectin also contains more phosphate than glycogen. These characteristics show that amylopectin is extremely similar to LBs and a suitable glucan to study Laforin's mechanism. Additionally, sugar binding studies with Laforin indicate that besides amylopectin and glycogen, Laforin can also bind β-cyclodextrin[26]. Some clues into Laforin's mechanism may be obtained from analysis of Laforin's ability to bind other phosphoglucans besides glycogen.

DXMS has been widely used in order to study the binding effects of synthetic drugs[27,28], protein structure and formation[29,30], and protein dynamics[31,32]. The present study represents an attempt to study domain behavior and glucan binding effects in wild-type Laforin and LD mutants.

MATERIALS AND METHODS

2.1 Protein Expression and Purification

Wild-type Laforin in the p21a vector was used as a template for QuikChange PCR (Stratagene) in order to generate the following mutants for study: W32G, G240S, and Y294N. Wild-type Laforin and mutant proteins were expressed in *Escherichia coli* BL21 (DE3) CodonPlus cells (Stratagene) using the p21a vector featuring a C-terminal His_6 tag. The transformed cells were grown in a laboratory shaker at 37 °C in 2xYT until they reached an OD of 0.6 at 600nm. Protein expression was induced with 0.4mM IPTG overnight at room temperature and lysed by sonication in HIS buffer (50mM Tris-HCl pH 8, 300mM NaCl, 3mM TCEP and 15% (v/v) glycerol). Laforin proteins were affinity-purified using Ni^{2+}-agarose (Qiagen) followed by further purification to near homogeneity using a HiLoad 16/60 Superdex 200 size exclusion column (GE Healthcare). The protein was concentrated and subjected to DXMS studies. All purifications were performed in HIS buffer.

Wild-type Laforin and tested mutant proteins were provided by my collaborators, Amanda Sherwood and Dr. Matthew Gentry from the University of Kentucky.

2.2 Optimization of Pepsin Proteolysis Conditions

At 0°C (on ice), 32µL of 2.3mg/mL wild-type (WT) Laforin protein in HIS buffer was diluted in 64µL of the same protein buffer. This mixture was incubated on ice for one minute. After one minute, 16µL of the sample was mixed with 4µL of cold quench solution (0°C) of 5.3% formic acid (v/v), 15% glycerol (v/v), 39mM tris(2-carboxyethyl)phosphine (TCEP) and guanidine hydrochloride (GuHCl) at final

concentrations of 0, 0.5, 1.0, and 1.5M. Within one minute of mixing the sample with quench solution, the solution was transferred to dry ice. The frozen samples were stored at -80°C until transfer to the dry ice compartment of the autosampler module of the DXMS apparatus for data acquisition.

In order to maintain temperature as close to 0°C as possible, the DXMS apparatus components: valves, tubing, columns, and autosampler were kept inside a refrigerator set to 4°C. In addition to keeping the apparatus inside a refrigerator, all columns and tubing were immersed in melting ice. The frozen samples were automatically thawed by the autosampler at 4°C and 20µL of sample immediately passed over a protease column (immobilized porcine pepsin coupled to AL-20 support from PerSeptive Biosystems at 30mg/mL; 16µl column bed volume) with 0.05% trifluoroacetic acid (TFA) in water (Solvent A) at 20µL/min with 48 seconds exposure to the protease. The pepsin generated fragments were collected on a C18 trap column (Michrom MAGIC C18AQ 0.2x2) and separated by a reverse phase C18 column (Michrom, MAGIC C18AQ 0.2x50) with a linear acentonitrile gradient 8%-48% (v/v) solvent B in 30 minutes at 2µl/min. Solvent B was 80% (v/v) acetonitrile, 20% water, and 0.01% TFA. The column eluent was directly injected into the LCQ Classic Mass Spectrometer (Thermo-Finnigan Inc.) electrospray ion trap-type mass spectrometer for data acquisition in either MS1 profile mode or data-dependent MS/MS mode. For each denaturant concentration, fragmentation maps that visually overlay the pepsin cleaved peptides with the protein primary amino acid sequence were generated. It was determined that final concentrations of 0.5M GuHCl and 10mM TCEP gave the best peptide coverage and was used in all subsequent experiments.

2.3 Hydrogen/Deuterium Exchange of Protein Samples

After determining which quench condition gave the best peptide coverage, WT Laforin samples were prepared in three different states of deuteration for each deuterium exchange experiment: non-deuterated (ND), deuterated, and fully-deuterated (FD). ND samples were prepared as described in the previous section. The deuterated samples were prepared by mixing 5.3µL of 2.3mg/mL WT Laforin protein with 10.6µL of Deuterium Oxide (D_2O) buffer containing 50mM Tris-HCl (pH 7), 300mM NaCl, and 3mM TCEP on ice. The deuterated samples were incubated on ice for 10s, 30s, 100s, 300s, 1000s, 3000s, and 10000s (166.67 min) prior to the addition of quench solution. The samples were quenched with 4µL of cold quench solution (2.5M GuHCl, 39mM TCEP, 15% glycerol (v/v), and 5.3% FA), transferred to ice cold autosampler vials, frozen on dry ice, and stored at -80°C. In the initial sugar binding experiment for WT Laforin, three structurally different sugars were used: amylopectin (5 mg/mL), glycogen (5 mg/mL), and β-cyclodextrin (5 mM). Prior to preparation of deuterated samples, WT Laforin was preincubated with each sugar for 1 hour on ice. For subsequent Laforin mutants, only glycogen or amylopectin was used for sugar binding experiments. The FD samples represent "equilibrium" amount of deuterium exchange. The FD samples were prepared by incubating 5.3µL of 2.3mg/mL WT Laforin protein with 10.6µL of D_2O containing 0.8% (v/v) of formic acid overnight (16 – 18hours). FD samples were quenched with 4µL of cold quench buffer as described above. All samples for subsequent mutants were prepared as described above for WT Laforin. Data from all deuterated sample sets were acquired from a single automated run of 8 hours. Only MS1 profile data was acquired for

the deuterated sample sets. Data reduction using specialized DXMS software is described below.

For the LD mutant G240S, all samples were run on a new instrument called the OrbiTrap Elite Mass Spec (Thermo Fisher, San Jose). The samples were prepared as described above, with the exception that the total protein per sample was 10-fold more dilute to yield optimal peptide detection by the OrbiTrap Elite. The instrument was operated in the positive ESI mode with a sheath gas flow of 8units, a voltage of 4.5kV, a capillary temperature of 200°C, and an S-lens RF of 67%. Mass spectroscopy data were acquired in both MS1 profile mode and data-dependent MS1:MS2 mode. The resolution of the survey scan was set at 60,000, at m/z 400 with a target value of $1x10^6$ ions and 3 microscans. The maximum injection time for MS/MS was varied between 25 and 200 ms. Dynamic exclusion was 30 s and early expiration was disabled. The isolation window for MS/MS fragmentation was set to 2, and the five most abundant ions were selected for product ion analysis.

2.4 DXMS Data Processing

In order to identify the potential sequences of parent peptide ions from the MS/MS data collected, the program SEQUEST (Thermo Finnigan Inc.) was utilized for LCQ data and Proteome Discoverer (ThermoFisher) was used for the OrbiTrap Elite data for mutant G240S.

These potential peptide identifications were inputted into specialized DXMS software developed in collaboration with Sierra Analytics, LLC, Modesto, CA. The tentative pool of identified peptides was passed through the quality threshold set by the

DXMS data reduction software. After the initial software threshold, each peptide mass spectral signal was manually checked for accuracy and quality. For peptides that had more than one ionization charge state, the highest signal/noise charge state was used. Typically, the lowest charge state for a peptide gave the best signal. Lastly, it is necessary to correct for the "back-exchange," or the loss of deuterium on peptides, which occurs because of the presence of H_2O throughout the processing procedure. The corrections for "back-exchange," determined by the methods of Zhang and Smith, use the following equations to determine deuteration level and deuterium incorporation for each peptide[33]:

$$\text{Deuteration Level (\%)} = \frac{m(P) - m(N)}{m(E) - m(N)} \times 100\%$$

$$\text{Deuterium incorporation (number)} = \frac{m(P) - m(N)}{m(E) - m(N)} \times \text{MaxD}$$

m(P), m(N), and m(F) represent the centroid values of the partially deuterated, non-deuterated, and fully deuterated peptide, respectively. MaxD is the maximum number of deuterons that can be incorporated into the peptide. The MaxD value is determined by subtracting the number of Proline residues that do not appear in the first or second amino acid residues of the peptide plus two, from the total number of amino acid residues for the peptide of interest, which assumes that the first two amino acids of the peptide back-exchange too rapidly to retain incorporated deuterons[34].

In this Thesis, the G240S-Laforin data obtained from the OrbiTrap Elite instrument will be compared to the LCQ data of the other Laforin proteins. Despite the difference in mass spectrometers, the back-exchange among the Laforin proteins was comparable (**Table 2**). However, future experiments will include repeating the Laforin proteins analyzed using LCQ mass spectrometer on the OrbiTrap Elite instrument.

RESULTS

3.1 Wild-Type Laforin Coverage Map of Pepsin Fragmentation

The ability to localize and quantify deuterium incorporation in a protein is largely dependent on the fragmentation pattern generated by proteolysis. Therefore, we spent considerable time and effort to optimize proteolysis conditions in order to achieve maximum protein coverage. The optimal pepsin fragmentation conditions for Laforin included: 0.5M GuHCl and 10mM TCEP during quenching, 30mg/mL porcine pepsin column (16µL bed volume), and 48 s exposure to pepsin. These fragmentation conditions yielded 134 high-quality peptides that covered 100% of the 339 amino acids of recombinant wild-type (WT) Laforin with a HIS_6 tag (**Figure 6**). The same fragmentation conditions were used for subsequent Laforin mutants and complete coverage over the entire protein sequence was similarly achieved for mutants.

3.2 WT Laforin Shows Solvation of Substrate Contact Sites

The goals for this work were three-fold: 1) define structural components of WT Laforin, 2) determine what regions of Laforin interact with sugars, and 3) define differences between WT and LD mutant Laforin. DXMS was used to investigate the structural and domain properties of WT Laforin in solution. The deuteration profile of WT Laforin was then used for comparing deuteration data in the presence of binding sugars as well as analyzing subsequent mutants. The localized deuteration profile of ligand-free WT is shown in the top set of colored bars in **Figure 7**, where the first two amino acids of a peptide are unresolved because the first residue does not contain an

amide hydrogen and the first N-terminal amide hydrogen of each peptide exchanges too rapidly during processing to retain a deuteron[34].

Carbohydrate binding modules often utilize clustered aromatic residues to interact with their glucan substrate[35]. The CBM of Laforin contains two carbohydrate-binding residues (tryptophan 32 and phenylalanine 84), where mutation of these amino acids have been demonstrated to impair the ability of Laforin to bind and dephosphorylate sugars[36]. The peptide containing residue W32 showed a high level of solvent exposure, as noted by the greater than 50% deuteration by the 300 s time point. Likewise, peptide 67 – 84 showed the highest level of deuteration within the CBM domain (**Figure 7**). Furthermore peptide 67 – 84 reached the maximum deuteration (under our experimental conditions) at the earliest time point (10 s), suggesting that F84 is extremely exposed to solvent and accessible for substrate binding. The accessibility of these two regions of the CBM supports the hypothesis that the two residues are structurally positioned to be critical for Laforin's ability to bind sugars.

In addition to probing the binding residues of the CBM, the critical regions of the DSP were also analyzed for deuterium incorporation. The phosphatase region of dual specificity phosphatases contains four structural elements that confer the ability to dephosphorylate substrates: recognition domain, variable loop, PTP-loop, and D-loop. Laforin's recognition domain (residues 139 – 154) determines the depth of the active site and contributes to substrate binding[37]. The two peptides covering the recognition domain both showed deuteration values greater than 90% at the 300 s time point (**Figure 8A, B**). The peptides of the recognition domain were the only peptides in the DSP to reach maximum deuteration and showed the highest level of deuteration at the 300 s time point

(**Figure 7**). After the recognition domain comes the variable loop (residues 193 – 228), which assists in orienting the active site with the substrate[38]. The variable loop is an extended loop in nearly all DSP family members, but the recent crystal structure of SEX4 from my collaborators at Gentry lab suggests that the V-loop in glucan phosphatases contains an α-helix [39]. The first peptide of the variable loop (residues 193 – 217) showed a maximum deuteration of ~70% at the longest time point (10,000 s) (**Figure 8C**). However, the remainder of the variable loop, hypothesized to be encompassed in an α-helix, showed a lower level of deuteration, <20% deuteration at the longest time point (**Figure 7**). The PTP-loop (residues 265 – 272) consists of the HCX_2GX_2R motif required for substrate presentation and formation of a phosphoenzyme intermediate[38]. Surprisingly, the first part of the PTP-loop (residues 265 and 266) showed limited (<20%) deuteration. However, the majority of the PTP-loop (peptide 267 – 281) showed a maximal deuteration of ~60% (**Figure 8D**). The combination of the PTP-loop and the D-loop compose the active site. The D-loop (residues 231 – 240) contains the catalytic aspartate residue that makes contact with the phosphorylated substrate. Aspartate 235 of the D-loop acts as a general acid/base catalyst, allowing Cysteine 266 of the PTP-loop to nucleophilically attack the phosphorous atom of the phosphate group and dephosphorylate the substrate[40]. The majority of the D-loop (peptide 229 – 236) showed a maximum deuteration of approximately 80% at the longest time points (**Figure 8E**). The last few residues of the D-loop (peptide 237 – 251) show maximum deuteration of 60% (**Figure 8F**). The accessibility of the four proposed structural elements of the DSP suggest that these regions are indeed accessible to interact with and dephosphorylate a glucan substrate. To further probe the structural dynamics and binding characteristics of

Laforin, we studied the solvent accessibility of Laforin in the presence of three carbohydrates: amylopectin, glycogen, and β-cyclodextrin (BCD).

3.3 WT Laforin Shows Preferential Interaction with Glycogen

In order to observe the changes in substrate-bound WT Laforin, we determined the deuteration level (%) for each peptide over the nine time points tested (**Supplemental Figure 1**) and determined the maximal changes between the substrate-bound and ligand-free WT Laforin (**Figure 9**). The DXMS profile for glycogen-bound WT Laforin is shown in the bottom set of colored bars in **Figure 7**.

The CBM of Laforin covers residues 1 – 116 and contains at least two carbohydrate-binding residues (W32 and F84). Peptides containing the two residues showed high levels of solvent accessibility in ligand-free WT (**Figure 7**, top set) and were expected to show protection from deuteration upon substrate binding. Of the two binding residues (yellow arrows, **Figure 9**), only glycogen-bound WT showed a significant decrease (>10%) change for residue W32. Glycogen-bound WT showed a 13.1% decrease in deuteration for the peptide covering W32 (**Figure 9B**). Amylopectin-bound and BCD-bound WT showed lower decreases of 9.1% and 8.4%, respectively (**Figure 9A, C**). With regards to F84, amylopectin-bound and BCD-bound WT showed significant decreases in deuteration of 19.8% and 11.3%, respectively. Unfortunately, the glycogen-bound WT data lacked a peptide giving deuteration data on F84. In addition to deuteration data on residues W32 and F84, an interesting observation was the significant decrease in deuteration for peptide 98 – 111 for all three substrate-bound conditions, despite no known critical binding residue in the region. These data suggest that the

regions within the CBM that are integral for glucan binding are not well defined and that a crystal structure of Laforin is needed to further define these regions.

Among the substrate-bound conditions, only glycogen-bound WT showed significant decreases (>10%) in deuteration for all four identified regions of the DSP and the changes were the largest in magnitude. Our hypothesis was that any region within the DSP that made substantial contact with glucans would result in protection from the deuterium solvent. These four regions within other DSPs interact with the substrate; therefore, we predicted that these regions within Laforin would make contact with glucans. Glycogen-bound WT showed the greatest decreases in deuteration for the recognition domain (**Figure 9**). In addition, the two recognition domain peptides for glycogen-bound WT showed drastically lower deuteration levels for all time points tested (**Figure 8A, B**). The amylopectin-bound condition only showed a significant decrease for the second half of the variable loop (peptide 146 – 155) and the decrease was smaller than the glycogen-bound condition (35.6% vs. 32.8%). The BCD-bound condition showed no significant changes in deuteration for the recognition domain. Similarly, only glycogen-bound WT showed a significant change in deuteration for the variable loop region (**Figures 9 & 8C**). The peptide that covered the majority of the PTP loop (peptide 267 – 281) showed nearly identical decreases in deuteration for amylopectin-bound and glycogen-bound WT, 15.2% and 15.5%, respectively (**Figure 9A, B**). BCD-bound WT did not show any significant decrease in deuteration of the PTP-loop (**Figure 9C**). Both amylopectin and glycogen-bound WT showed significant decreases in deuteration for the D-loop region (**Figure 9**). These data strongly suggest that Laforin interacts more readily with glycogen than either amylopectin or BCD.

3.4 DXMS Profile for LD Mutation in CBM

After utilizing DXMS to define the glucan binding interactions of the DSP and CBM domains of WT Laforin and with different substrates, our goal shifted to defining the differences between WT Laforin and LD mutant Laforin. However, we had to verify that DXMS was capable of providing Hydrogen-Deuterium (H-D) exchange data in agreement with established glucan binding assay data and clinical patient data. Therefore, we began our LD mutant H-D exchange experiments using the CBM mutant W32G-Laforin. Mutation of the tryptophan to glycine at residue 32 in the CBM is a known LD missense mutation[36]. Glucan binding assay results from my collaborator, Amanda Sherwood, confirmed that W32G-Laforin is incapable of binding glucans (data not shown). Based on these findings, H-D exchange data from W32G-Laforin in the presence of amylopectin was expected to exhibit H-D exchange data similar to W32G-Laforin alone.

In ligand-free W32G-Laforin, the CBM segment containing the W32G mutation showed high accessibility to the solvent, as indicated by greater than 50% deuteration after 300 s in D_2O. The H-D exchange data of W32G-Laforin in the presence of amylopectin mirrored the ligand-free data, indicating that no amylopectin was bound to the CBM, supporting our hypothesis (**Figure 10B**). In addition, the DSP regions playing integral roles in dephosphorylation of substrates showed no significant changes in deuteration after incubation with amylopectin (**Figure 10A, C**). The lack of deuteration changes in the DSP confirmed the inability of W32G-Laforin to bind amylopectin. Overall, the H-D exchange results for W32G-Laforin confirmed that DXMS is consistent with data obtained from glucan binding assays and LD clinical data.

3.5 DXMS Profile of LD Mutation in DSP

After probing the deuteration profile of the LD CBM mutant W32G and the binding capabilities of WT, our goal shifted to the analysis of the LD DSP mutant G240S-Laforin. Glycine 240 is situated within the last residue of the D-loop and is positioned between the catalytic aspartate 235 and the PTP-loop. Glycogen binding data and phosphatase assay data from my collaborator, Amanda Sherwood, demonstrated that G240S-Laforin shows WT-like binding of glycogen and the ability to dephosphorylate glycogen. However, G240S-Laforin does not interact with a critical binding partner, protein targeting to glycogen (PTG), and it is this lack of binding that is hypothesized to lead to Lafora disease for patients with this mutation[36]. Thus, we predicted that G240S-Laforin should behave similarly to WT Laforin with respect to DXMS results.

The maximal changes in deuteration incorporation for the peptides of G240S-Laforin upon glycogen binding are shown in **Figure 11**. Similar to WT Laforin, glycogen-bound G240S showed protection of residue W32 from deuteration. However, G240S failed to show protection for F84 (**Figure 11,** yellow arrows). Glycogen-bound G240S showed greater decreases in deuteration for the recognition domain, compared to WT. In contrast, the peptides of the variable loop region of G240S experienced smaller decreases in deuteration in the presence of glycogen. Despite smaller decreases, the changes in deuteration were still significant (>10%) and still suggest protection due to substrate binding/interaction. Glycogen-bound G240S showed nearly identical changes in the D-Loop as glycogen-bound WT, suggesting that the mutation did not impair the ability of the D-loop to interact with the substrate. Also similar to glycogen-bound WT, G240S-Laforin showed a significant decrease in deuteration for the PTP-loop region

when bound to glycogen. The DXMS profile data for G240S Laforin confirmed the glycogen binding assay and phosphatase data by showing protection from deuteration in sites of substrate interaction. Lastly, the H-D exchange data confirmed that DXMS is a method capable of measuring the changes in deuteration due to substrate binding.

3.6 New Insights into LD DSP Mutant

The mutation of Laforin residue 294 from tyrosine to asparagine is a known LD mutation in the DSP domain. Although the mutation does not appear in the known structurally important regions of the DSP, other researchers have claimed that Y294N-Laforin shows impaired phosphatase activity[36]. My collaborator, Amanda Sherwood, collected glycogen binding assay data for Y294N-Laforin, which suggested that the mutation does not inhibit glycogen binding.

The DXMS profile of Y294N-Laforin in the presence or absence of glycogen suggested that CBM interaction with glycogen was slightly decreased (**Figure 12**). This result was somewhat surprising, given that the Y294N mutation is in the DSP domain and yet it appears to somewhat diminish glycogen interaction with the CBM. Both of the proposed binding residues (W32 and F84) showed lower changes in deuteration in the presence of glycogen. Although my collaborator's data demonstrated that Y294N-Laforin shows a stronger affinity for glycogen than WT, the DXMS data within the CBM suggested a decreased interaction with glycogen. However, surprising results within the DSP of Y294N likely account for our differences, see below.

Since other researchers have suggested that phosphatase function is impaired by the Y294N mutation, we expected H-D exchange data to show minimal interaction

between the DSP domain and glycogen. Unexpectedly, the four functional regions of the DSP all showed significant protection from deuteration when incubated with glycogen, which were similar to WT Laforin (**Figure 12**). However, glycogen-bound Y294N-Laforin showed a significant decrease in the second half of the recognition domain, whereas glycogen-bound WT showed decreases spanning the entire recognition domain. Despite the slight difference in the deuteration changes for the recognition domain, the magnitudes of decrease in deuteration for the variable loop, D-loop, and PTP-loop were comparable between Y294N and WT. Collectively, these data suggest that despite somewhat diminished interaction with the CBM domain, the DSP domain of Y294N-Laforin appeared to still interact with glycogen. Furthermore, the peptide covering the Y294N mutation showed drastic decreases in deuteration in response to the presence of glycogen, compared to no change in deuteration exhibited by WT (**Figure 13**). Although the Y294N mutation has been identified previously as a LD mutation, no extensive analysis of the mechanism of this mutant exists. The DXMS profile suggests that the Y294 residue plays a role in DSP domain function that has not yet been documented. The increased interaction with glycogen in the DSP domain, particularly around Y294N, may indicate that residue 294 is required for proper release of dephosphorylated substrates. Whether or not the region of residue 294 contains a structurally relevant element of the DSP domain still remains to be determined. Ultimately, the data suggests that there is still much about Laforin's form and function still to be elucidated. The final piece of the puzzle may ultimately rest with a complete crystal structure of Laforin.

DISCUSSION

Laforin is the only protein in the human proteome that contains a carbohydrate binding module (CBM) coupled to a dual specificity phosphatase (DSP) domain. The gene encoding Laforin was found to be mutated in Lafora disease patients in 1998, yet no structural studies have been published describing this glucan phosphatase[17]. Thus, our understanding of the structural basis for why mutations in Laforin lead to Lafora Disease (LD) is very limited. A clearer understanding of how mutations affect the structure of Laforin may lead to proposed therapies and/or treatments for LD patients.

The Laforin CBM is responsible for facilitating Laforin's binding to carbohydrates. Carbohydrate binding modules often employ clusters of aromatic residues that participate in stacking interactions with the glucan. Our W32G-Laforin results demonstrated that W32 is a critical binding residue. Furthermore, these data showed that W32G abolishes glucan interaction of both the CBM and DSP with the substrate. The wild-type Laforin (WT) data revealed the highly solvent accessible nature of W32, and evidence of glycogen binding was shown in significant decreases in solvent accessibility when incubated with glycogen (**Figure 7**). In contrast, our WT data was unable to confirm that the putative carbohydrate binding residue F84 is required for Laforin binding to glycogen (**Figure 9**), but current literature revealed that mutation of F84 abolished glycogen affinity and destroyed phosphatase activity[36]. Furthermore, F84 did show extremely high levels of solvent exposure in the ligand-free condition, >90% deuteration at the 10 s timepoint, suggesting that F84 is extremely solvent accessible for possible glucan interactions.

Interestingly, the end of the CBM domain (peptide 98 – 111), consistently showed a decrease in deuteration when incubated with glycogen (**Figures 9, 11, & 12**). Residues 98 – 111, whose function in carbohydrate binding still remains unclear, exhibited significant decreases in deuteration for WT, G240S, and Y294N-Laforin when incubated with glycogen. The decrease in H-D exchange may be attributed to substrate-induced allosteric changes or the presence of unidentified carbohydrate binding residues within this region. Previous studies identified critical residues in the CBM through mutagenesis studies of conserved LD mutations in the CBM. The region in question encompasses one conserved LD mutation, R108C, whose effects on glucan binding and Laforin function have been documented. Previous studies have concluded that R108C impairs phosphatase activity but does not affect Laforin's affinity for glycogen[41]. Our DXMS data appear to suggest that Arg 108 may actually play a role in substrate binding not seen by other methods. Additionally, it should be noted that residues 98-111 contains four aromatic residues that may be critical for substrate binding.

The regions of the DSP domain that have been implicated in dephosphorylation of phosphoglucans (recognition domain, variable loop, PTP-loop, and D-loop) all showed significant protection from solvent upon glycogen binding. These results suggest that the DSP domain makes extensive and intimate contact with the substrate. The DXMS data for WT, G240S and Y294N all indicated that the accessibility of the recognition domain and variable loop facilitates the substrate presentation to the PTP-loop, allowing the D-loop to catalyze the dephosphorylation reaction. The absence of widespread changes throughout the protein indicate that the DSP domain is in an open and active

conformation, similar to the DXMS analysis of the Laforin functional equivalent starch excess 4 (SEX4) in *Arabidopsis*[25].

The finding that WT Laforin showed stronger interaction with glycogen over other carbohydrates was not surprising, but it was a significant finding for further research. Despite common agreement that glycogen is the substrate for Laforin within the human body and that Laforin plays some role in glycogen regulation, analysis of the characteristics of amylopectin and Lafora Bodies (LBs) suggest that amylopectin is more biochemically similar to LBs than glycogen (**Table 1**). Further research suggests that amylopectin is a suitable substrate for Laforin. Gentry et. al found that Laforin is a functional equivalent of the plant protein SEX4, by demonstrating that expression of Laforin prevented the starch-excess phenotype in *Arabidopsis* plants that contained a catalytically inactivated SEX4[16]. While Laforin is capable of dephosphorylating starch *in planta* in order to prevent the starch-excess phenotype, our data suggests that Laforin prefers glycogen as a substrate.

Similarly, β-cyclodextrin (BCD) was tested due to its binding ability to the AMP-activated protein kinase targeting subunit β1 (AMPKβ1) class of CBM, which shares structural similarities to SEX4[42]. Previously, Hsu et. al discovered that BCD did interact with both the SEX4 CBM and DSP domain, but the interaction was markedly decreased compared to glycogen[25]. In contrast with these results, our Laforin data showed insignificant changes in the putative carbohydrate binding residues in the CBM, as well as lack of significant DSP solvent protection in the presence of BCD. It is quite possible that the smaller size of BCD, compared to glycogen and amylopectin, contributed to the lack of protection from solvent upon binding. Alternatively, bioinformatics data suggest

that the Laforin CBM belongs to the subfamily CBM20 and the SEX4 CBM is classified as a CBM21 member[25,43]. Thus, there are likely structural differences within the SEX4 and Laforin CBM that account for different substrate preferences.

Cumulatively, our results for WT Laforin indicated that Laforin shows stronger interaction with glycogen compared with amylopectin (**Figure 9**). This result has significant ramifications for the analysis of sugar binding and endogenous phosphate release assays. Although Laforin is a functional equivalent of SEX4 *in planta*, Laforin does show a preference towards glycogen, and this should be considered in crystallization attempts and mechanistic experiments.

The data for DSP mutant G240S-Laforin indicated that the mutation did not affect Laforin's binding of glycogen and the DSP was still able to interact with glycogen. The H-D exchange profile for this mutant proved to further validate our DXMS conclusions by serving as a positive control, similar to how W32G served as a negative control. Mutagenesis studies of Laforin concluded that the G240S mutation did not affect Laforin's affinity for glycogen but did result in reduction of phosphatase activity to 80% of WT levels[41]. Our G240S data supported previous findings that G240S could bind glycogen. The G240S H-D exchange data showed that residue W32 and the DSP regions experienced solvent protection when incubated with glycogen (**Figure 11**). The magnitudes of deuteration decrease in the variable loop and the PTP-loop for G240S were smaller than the decreases seen in WT (**Figure 9B**), and may correlate with previous findings that G240S only showed 80% of WT phosphatase activity.

Our analyses also identified the region of the DSP domain around tyrosine 294 as a possibly critical element of Laforin's phosphatase function. Previously, Fernandez-

Sanchez et. al found that the Y294N-Laforin mutant showed significantly decreased phosphatase activity and abolished glycogen binding[41]. In contrast, our DXMS data suggests that although the CBM of Y294N-Laforin showed decreased glycogen interaction, the DSP mutation has no effect on the interaction between glycogen and the structurally relevant elements of the DSP (**Figure 12**). In addition, the region of the Y294N mutation showed significant interaction with glycogen (**Figure 13**). Furthermore, my collaborator observed stronger than WT levels of glycogen binding by native polyacrylamide gel electrophoresis of Y294N-Laforin in the presence of glycogen (unpublished data). Our DXMS findings of unaffected DSP interaction and increased interaction around the Y294N mutation may partially explain (and be confirmed by) the stronger glycogen binding seen by my collaborator, Amanda Sherwood. Further research into the predicted secondary structure of the Y294N region, suggests that the region encompasses a conserved α-helix[44]. The crystal structure of SEX4 and sequence alignment with Laforin confirms that the DSP of SEX4 contains the predicted α-helix that encompasses the Y294N mutation of Laforin[39]. The conservation of this α-helix may hint to an important role of this region in the release of substrate that has not yet been documented.

In summary, we have described the structural analyses of WT, W32G, G240S, and Y294N-Laforin in complex with carbohydrates. We determined that WT Laforin preferentially interacts with glycogen when compared to BCD or amylopectin. The data from mutants W32G and G240S confirmed DXMS as a vital method for collecting structural data of proteins in the absence of substrate binding (W32G data) and interactions due to substrate binding (G240S data). Lastly, the analysis of the mutant

Y294N revealed that much is still unknown about the form and function of Laforin and dual specificity phosphatases. However, our data provides insights into a protein's structure, which may ultimately be confirmed and expanded upon by successful crystallization.

FIGURES

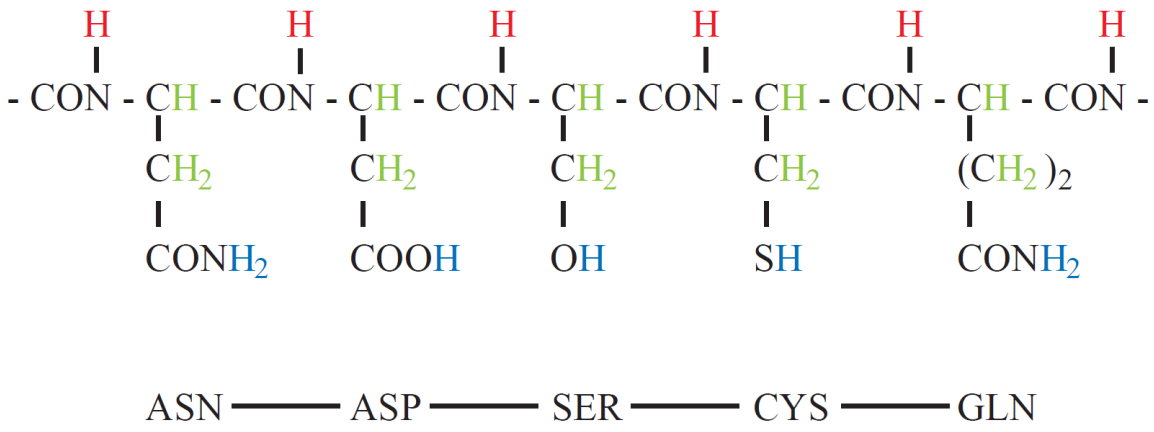


Figure 1: Three main groups of exchangeable hydrogens on a protein. Hydrogens attached to carbon atoms (green H) do not exchange. Hydrogens that are part of side change functional groups (blue H) exchange so rapidly that incorporated deuterium is not retained during processing. The amide hydrogens (red H) exchange and are measured by DXMS. (adapted from Hsu)[45].

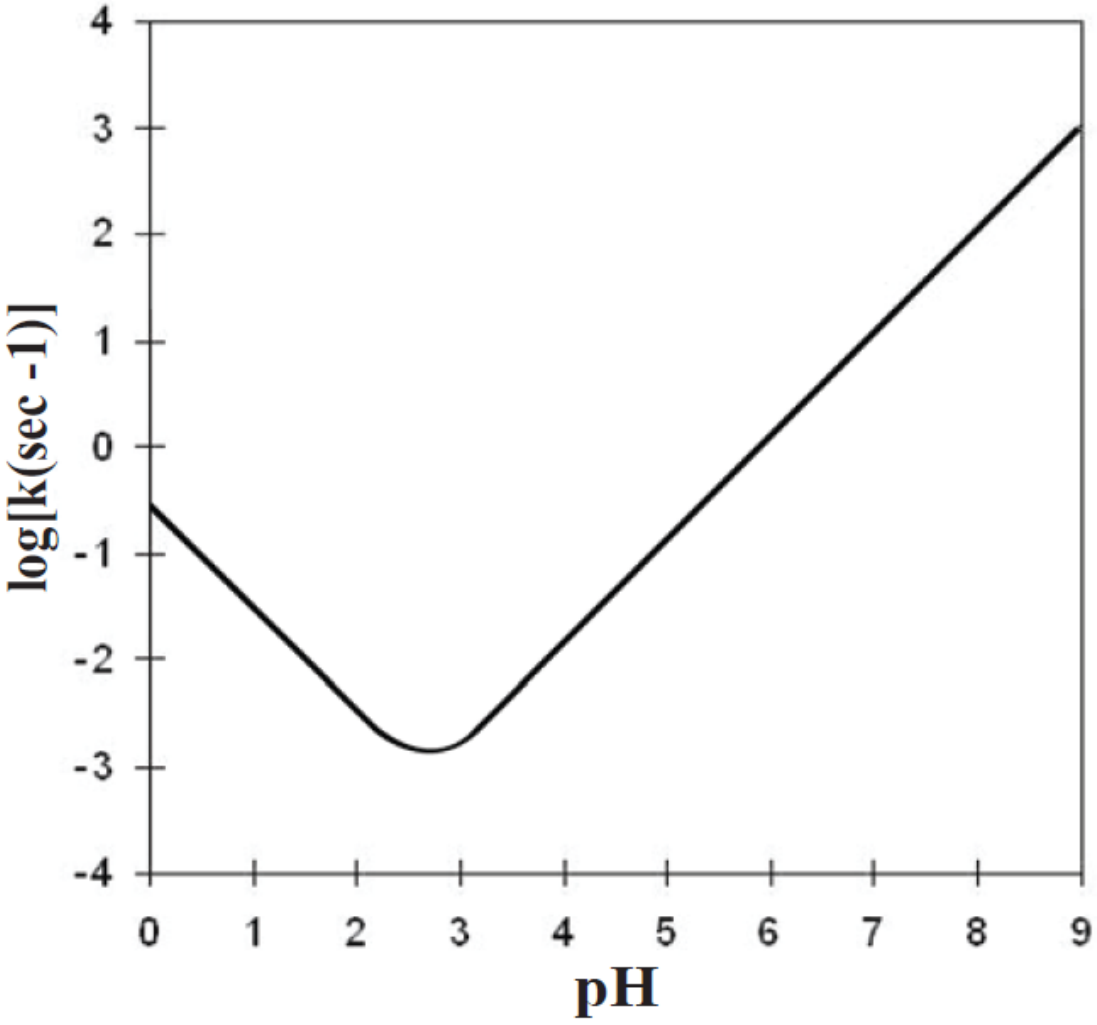


Figure 2: pH effects on amide hydrogen/deuterium exchange. The hydrogen/deuterium exchange rate constant is plotted against the pH of the solution. At pH ~2.5, hydrogen/deuterium exchange is at a minimum. (adapted from Wintrode and Tsutsui)[1].

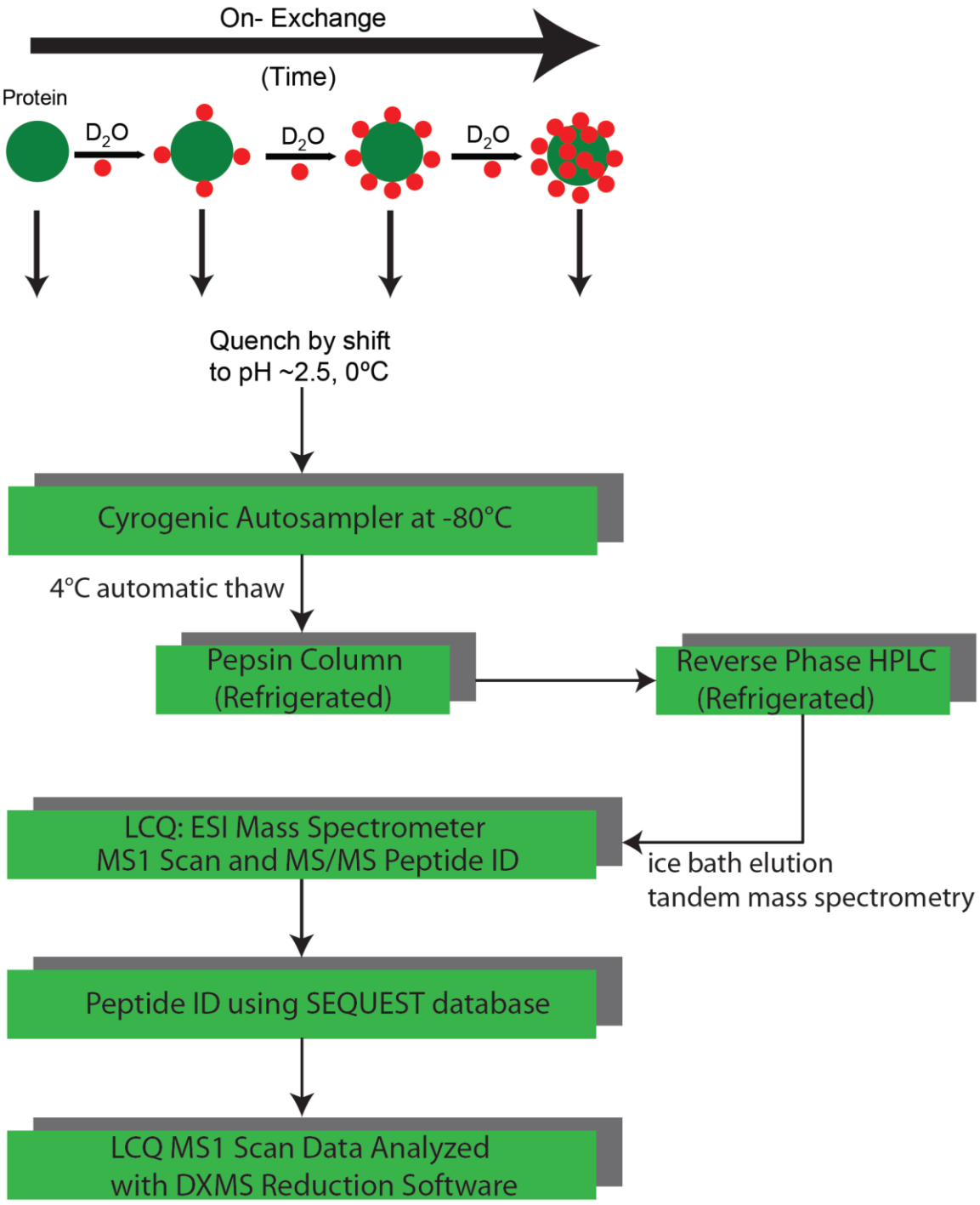


Figure 3: Overview of a DXMS experiment. After manual preparation of on-exchange samples, automated protein processing generates the data used for SEQUEST peptide identification and DXMS data analysis.

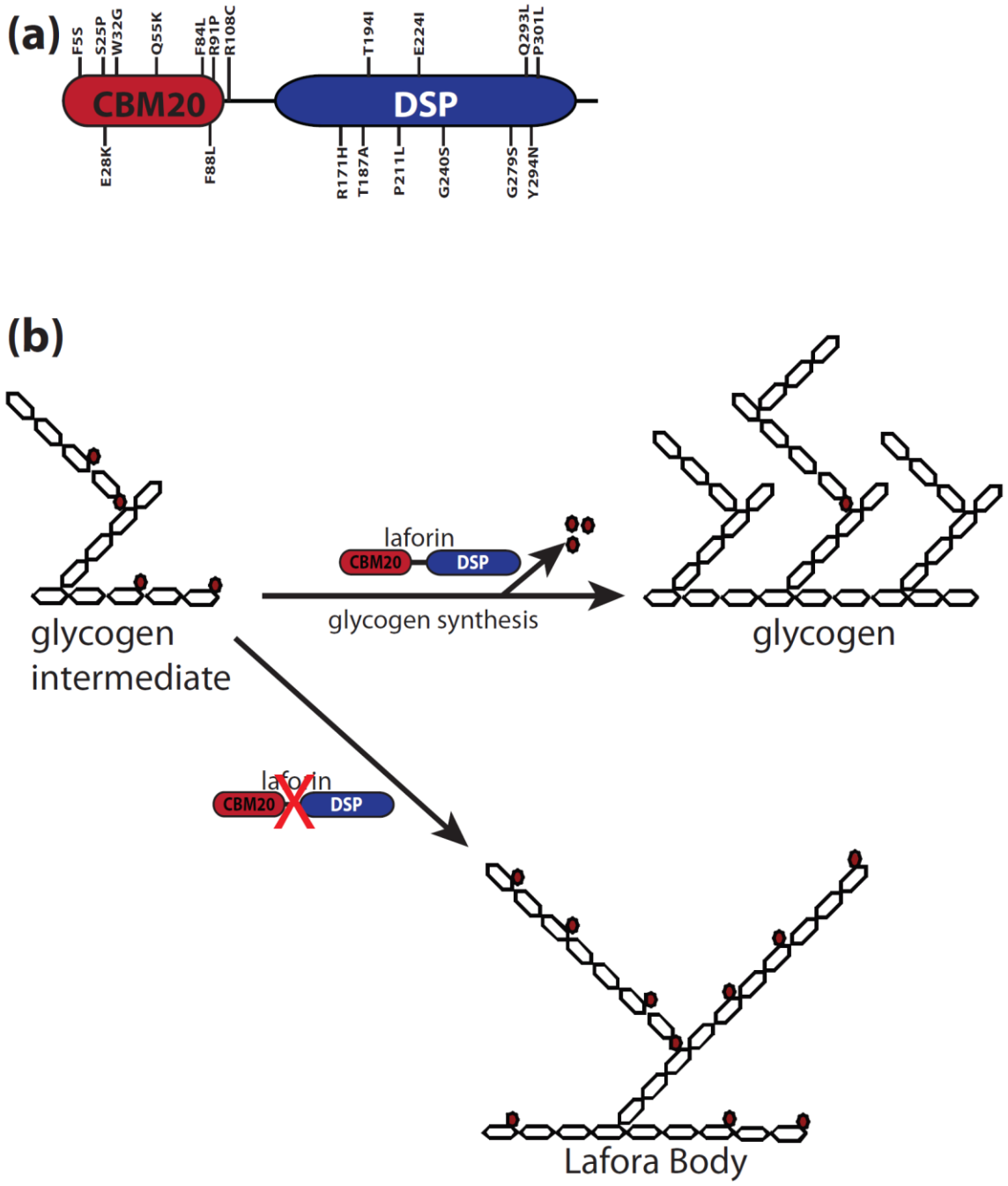


Figure 4: Laforin schematic and proposed mechanism. (a) Laforin contains a carbohydrate binding module family 20 (CBM 20) and a dual specificity phosphatase (DSP) domain. Lafora Disease missense mutations are distributed over both domains. (b) Functional Laforin removes phosphate groups from glycogen intermediates during glycogen synthesis and allows for correct glycogen production. The absence of functional Laforin results in accumulation of phosphate groups, which decreases branching and leads to formation of Lafora Bodies. (adapted from Gentry et. al)[44].

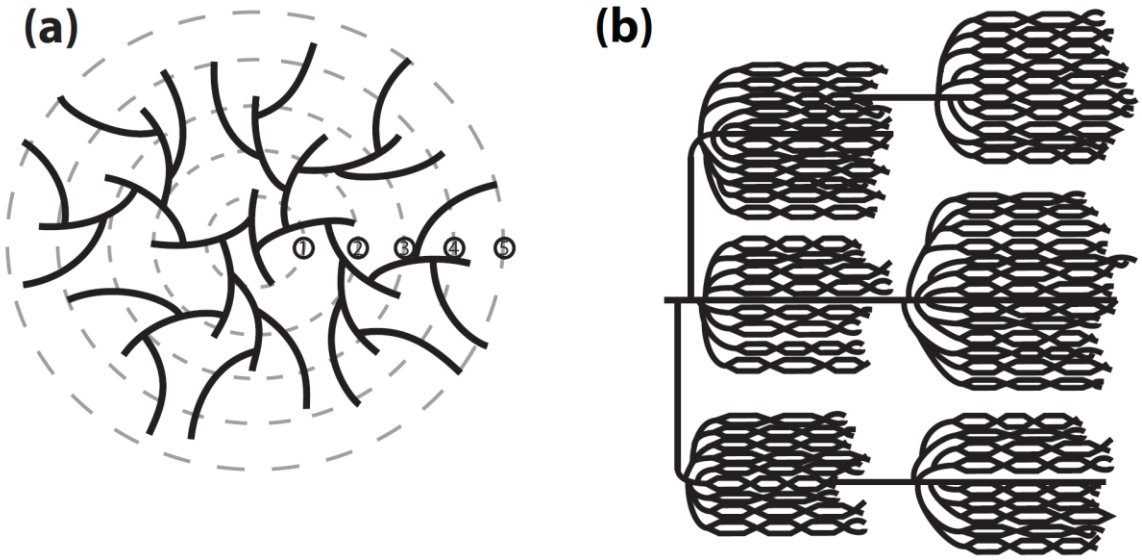


Figure 5: Representative models of glycogen and amylopectin. In each model, unbroken lines represent chains of linked glucose. (a) Glycogen synthesis consists of linking glucose molecules by α-1,4-glycosidic linkages with branches linked by α-l,6-glycosidic linkages ever 12 – 14 residues. Tiers of branching, indicated by dotted circles, are organized in a continuous manner, which renders glycogen water soluble. (b) Amylopectin is also composed of α-1,4-glycosidic linkages with branches linked by α-l,6-glycosidic linkages, but the branches are arranged in clusters at regular intervals. The intersecting unbroken lines represent how glucose chains interact within clusters. The decreased branching renders amylopectin insoluble in water. (adapted from Gentry et. al)[44].

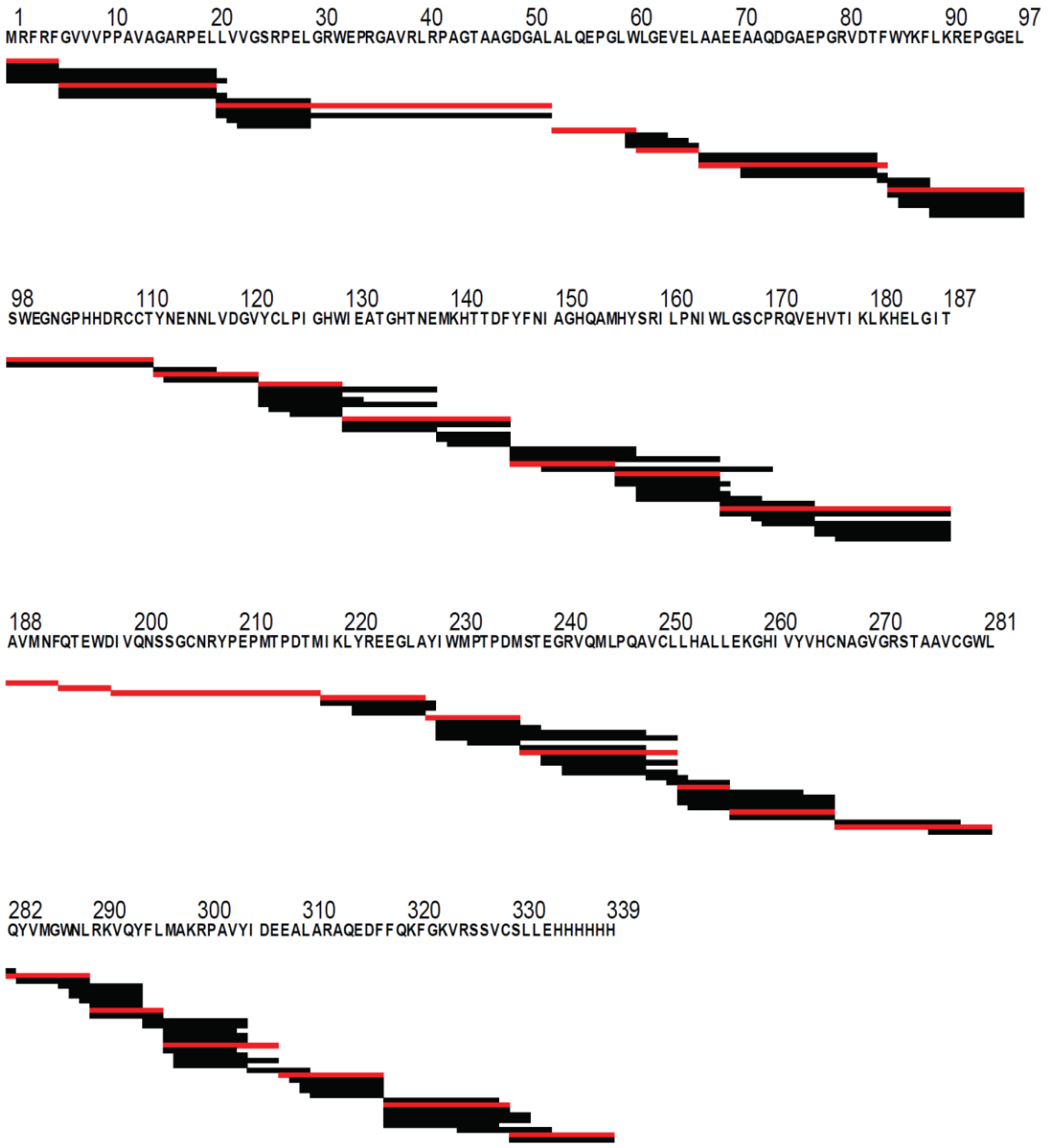


Figure 6: Pepsin-digested coverage map of wild-type Laforin. The pepsin-digested peptides that were identified from the MS/MS experiments are presented overlaid with the primary amino acid sequence and residue number. A total of 134 peptides were detected. Solid red lines indicate the peptides used for analysis, and solid black lines indicate peptides not utilized for analysis.

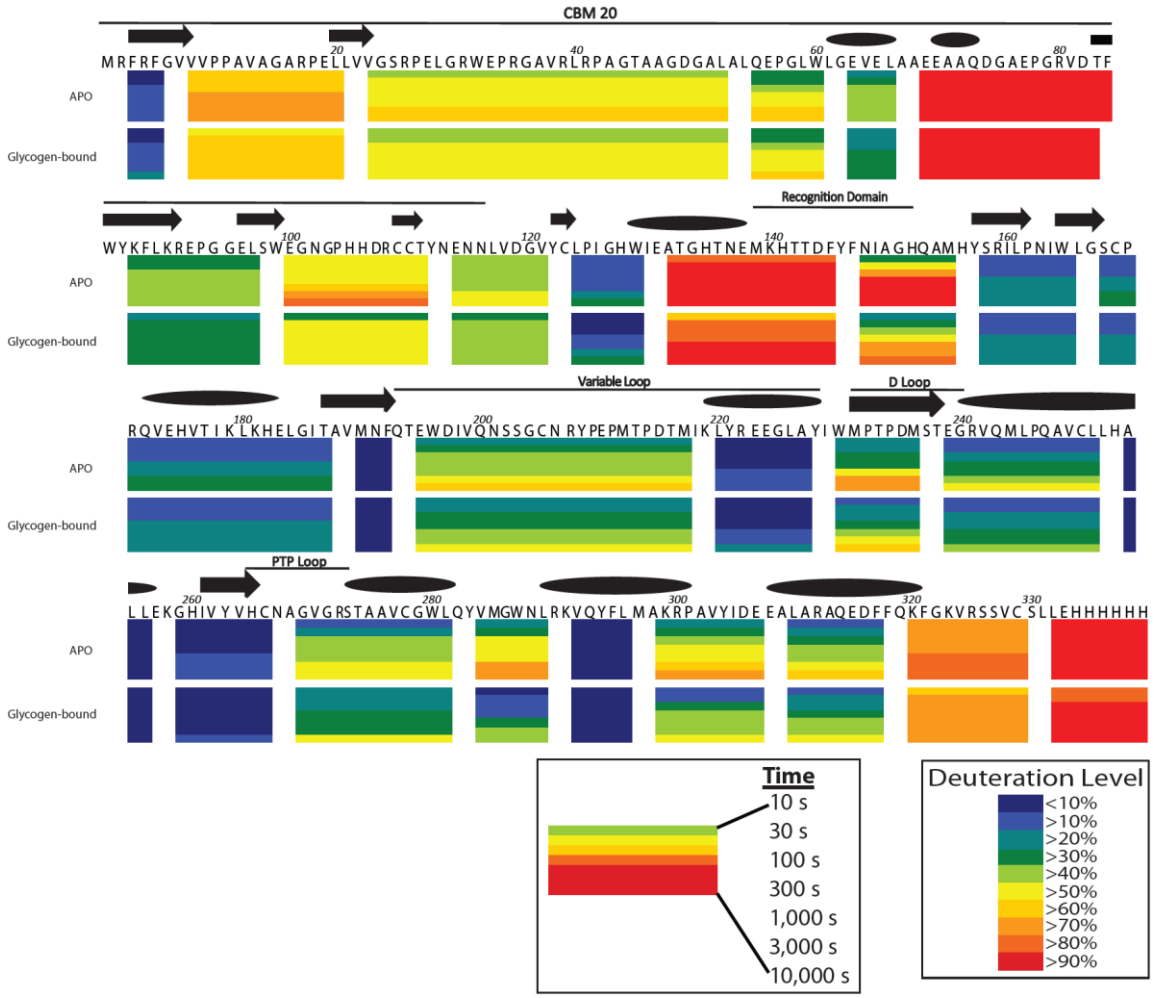


Figure 7: DXMS results for apo and glycogen-bound wild-type Laforin. The top set of colored bars corresponds to apo-WT Laforin and the bottom set corresponds to glycogen-bound WT Laforin. Each bar, consisting of seven rows, represents the peptides we analyzed from Figure 6. Each row represents the level of deuteration at one of the seven time points from 10 s to 10,000 s for the selected peptide, and the color of the row ranges from low deuteration levels in blue to highest deuteration levels in red (inset). Predicted secondary structural elements, from PROMALS[44,46], are shown above the primary sequence (α-helices as ellipses and β-sheets as arrows). Deuterons from the first two amino acids of each peptide are lacking due to the fact that the first amino acid of each peptide lacks an amide hydrogen and the amide hydrogen from the second residue exchanges too rapidly to retain deuterons during processing[34]. Regions of importance are labeled above the primary sequence.

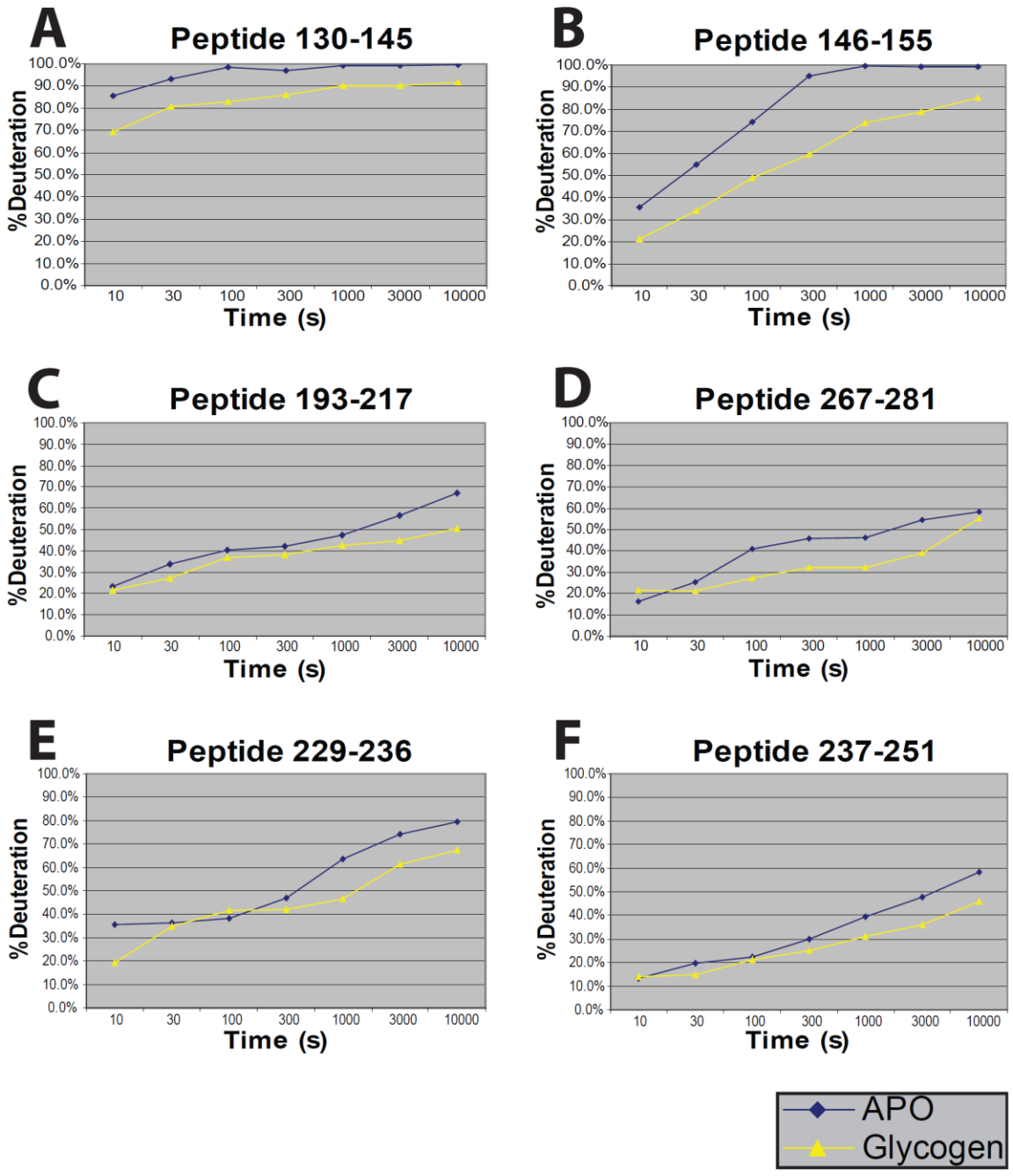


Figure 8: Wild-type Laforin DSP domain peptides that exhibited a decrease in deuterium incorporation for apo vs glycogen-bound conditions. Time dependent deuterium incorporation for (A) recognition domain peptide, $I^{130} – F^{145}$ (B) recognition domain peptide, $Y^{146} – M^{155}$ (C) variable loop peptide, $Q^{193} – M^{217}$ (D) PTP-loop peptide, $C^{266} – L^{281}$ (E) D-loop peptide, $I^{229} – M^{236}$ (F) D-loop peptide, $S^{237} – L^{251}$.

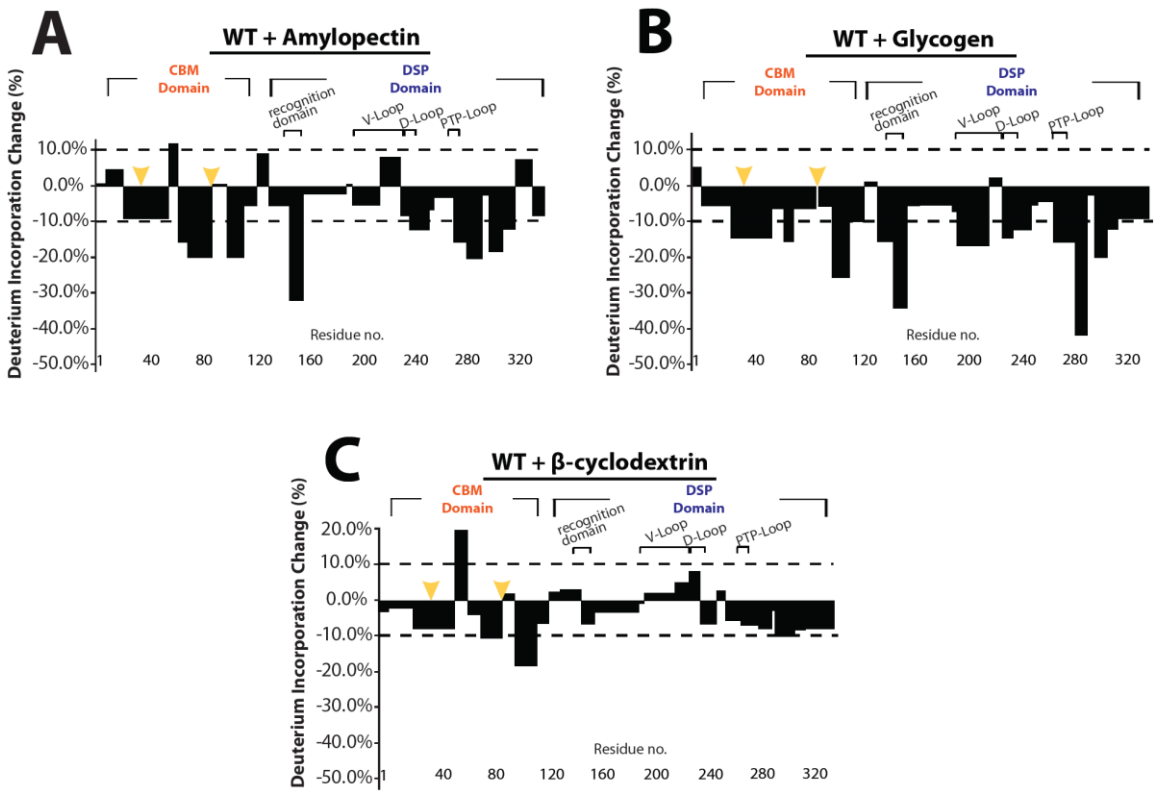


Figure 9: Deuterium incorporation changes in wild-type Laforin upon glucan binding. (A) Bar graph showing the maximum percent changes in deuteration between apo-WT Laforin and amylopectin-bound WT Laforin at any time point. The different regions of the DSP domain are noted at the top of the graph and the yellow arrows indicate the positions of residues within the CBM required for glucan binding. (B), (C) Like panel A, with the exception that the comparison is between apo-WT Laforin and WT Laforin in the presence of glycogen and β-cyclodextrin, respectively. In all panels, a positive value for change indicates increased deuteration, and a negative value signifies a decrease in deuteration after glucan binding. Changes greater than 10% (dashed lines) are considered to be significant.

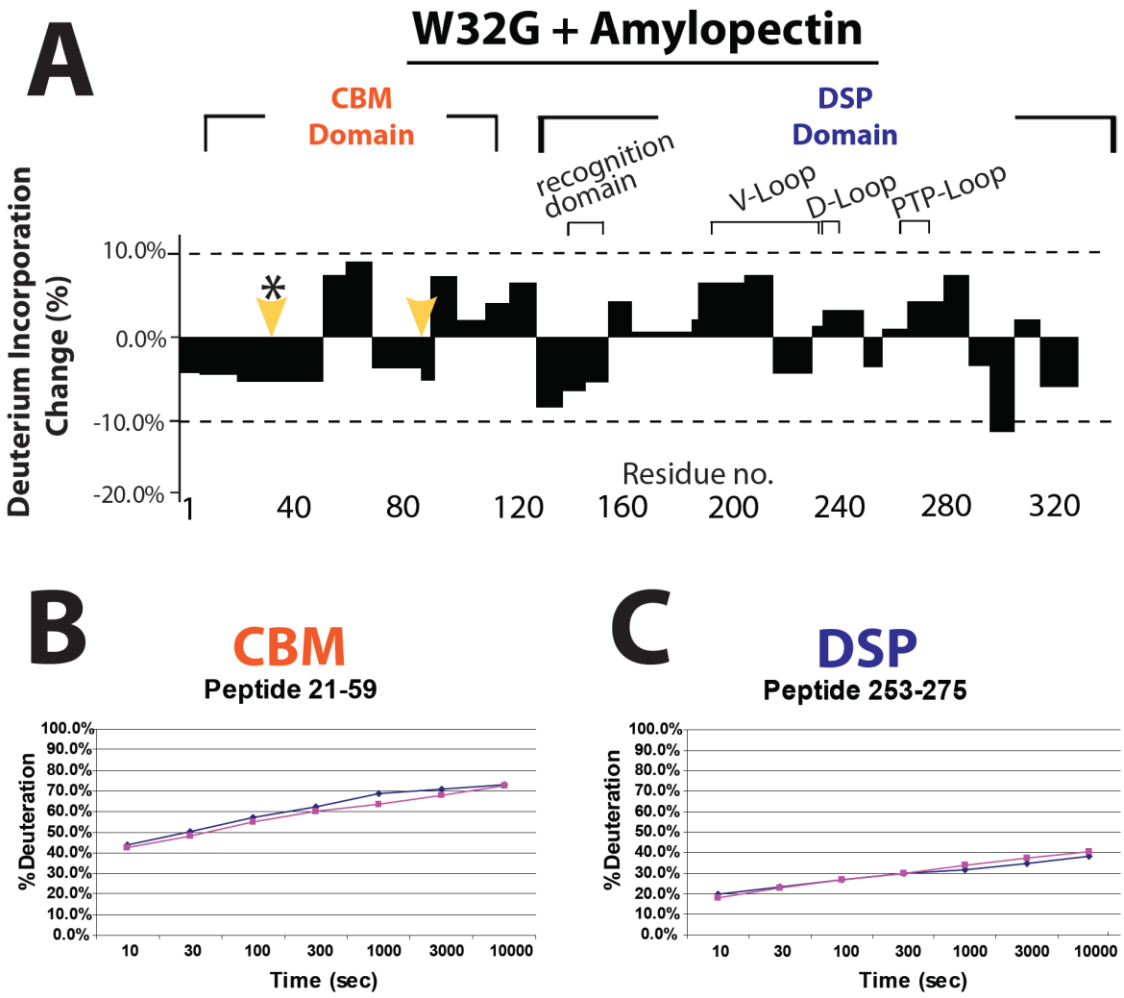


Figure 10: Deuterium incorporation changes of W32G-Laforin in the presence of amylopectin. (A) Bar graph depicting the maximal percent changes between W32G-Laforin incubated with amylopectin and W32G-Laforin alone. The different regions of the DSP domain are noted at the top of the graph and the yellow arrows indicate the positions of residues within the CBM required for glucan binding. The asterisk indicates the location of the mutation. A positive value for change indicates increased deuteration, and a negative value signifies a decrease in deuteration after glucan binding. Changes greater than 10% (dashed lines) are considered to be significant. (B), (C) Time dependent incorporation of deuterium into W32G-Laforin. Data shown as pink squares and purple diamonds represent percentages of deuterium incorporation into W32G-Laforin in the presence of amylopectin and W32G-Laforin alone, respectively. (B) CBM peptide $L^{21} – L^{59}$, which encompasses the W32G mutation. (C) DSP peptide $H^{253} – A^{275}$, which encompasses the PTP loop.

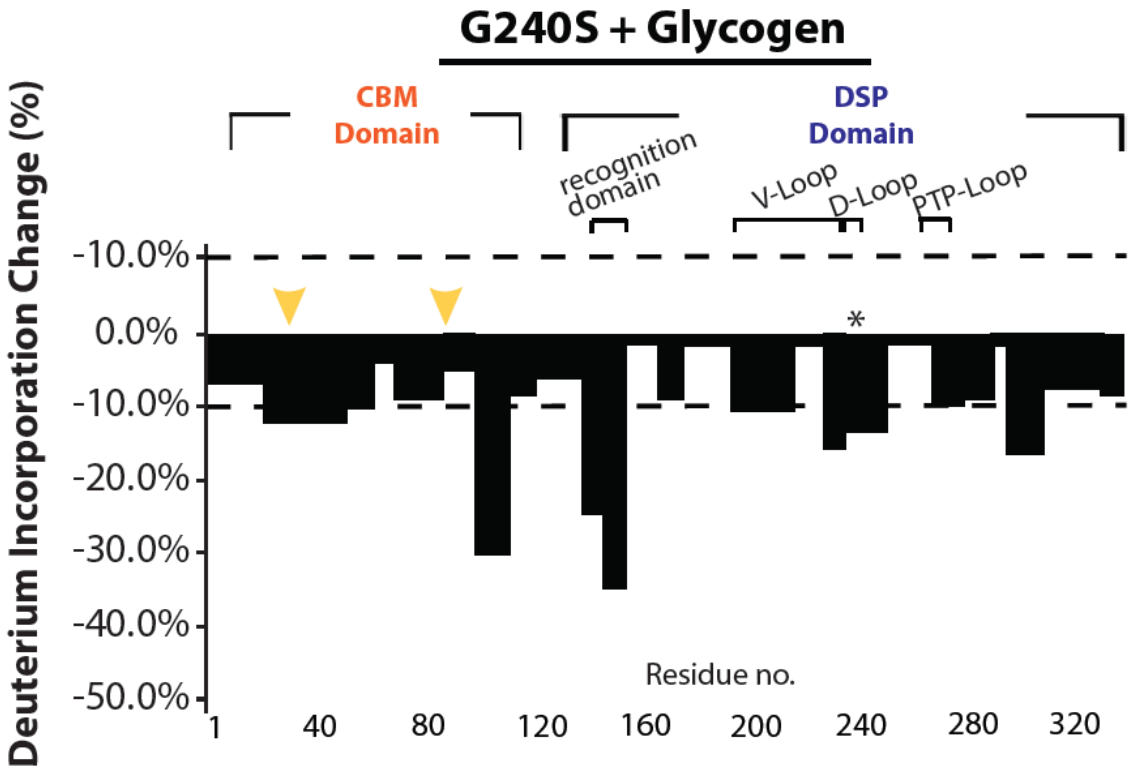


Figure 11: Deuterium incorporation changes in G240S-Laforin upon glycogen binding. (A) Bar graph showing the maximal percent decreases in deuteration between apo G240S-Laforin and glycogen-bound G240S-Laforin at any time point. The different regions of the DSP domain are noted at the top of the graph and the yellow arrows indicate the positions of residues within the CBM required for glucan binding. The asterisk indicates the location of the mutation. A positive value for change indicates increased deuteration, and a negative value signifies a decrease in deuteration after glucan binding. Changes greater than 10% (dashed lines) are considered to be significant.

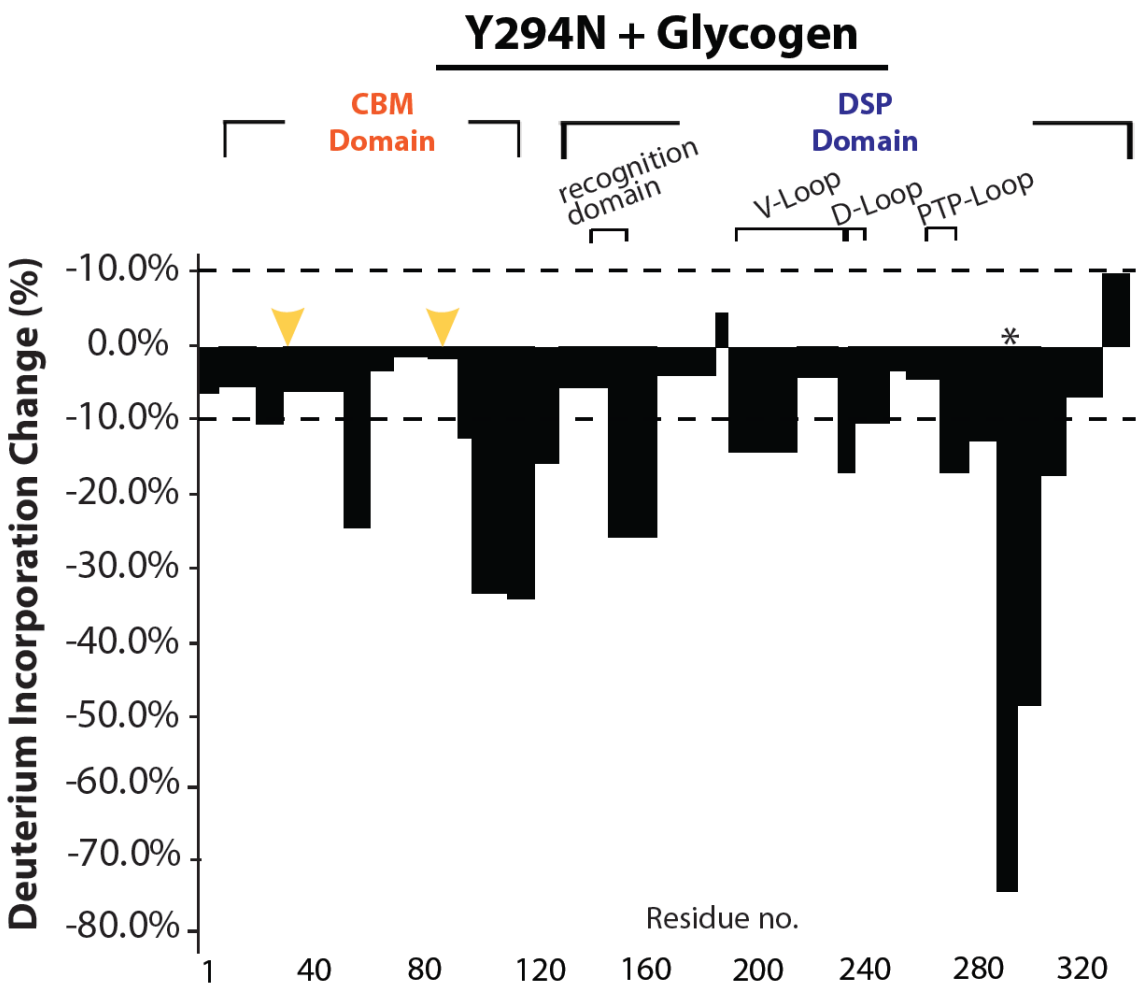


Figure 12: Deuterium incorporation changes in Y294N-Laforin upon glycogen binding. Bar graph showing the maximal percent changes in deuteration between apo Y294N-Laforin and glycogen-bound Y294N-Laforin at any time point. The different regions of the DSP domain are noted at the top of the graph and the yellow arrows indicate the positions of residues within the CBM required for glucan binding. The asterisk indicates the location of the mutation. A positive value for change indicates increased deuteration, and a negative value signifies a decrease in deuteration after glucan binding. Changes greater than 10% (dashed lines) are considered to be significant.

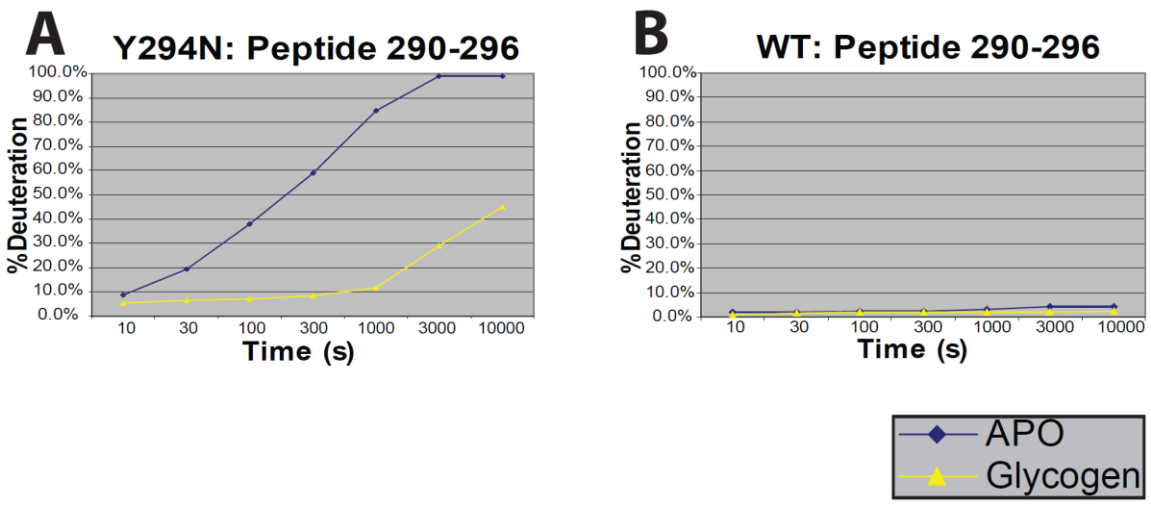


Figure 13: Differences in deuteration of DSP peptide R^{290} – L^{296} between wild-type Laforin and Y294N-Laforin. The time dependent deuterium incorporation for the same peptide is shown. Data shown as blue diamonds and yellow triangles represent percentages of deuterium incorporation into apo and glycogen-bound conditions of (A) Y294N-Laforin and (B) WT Laforin.

TABLES

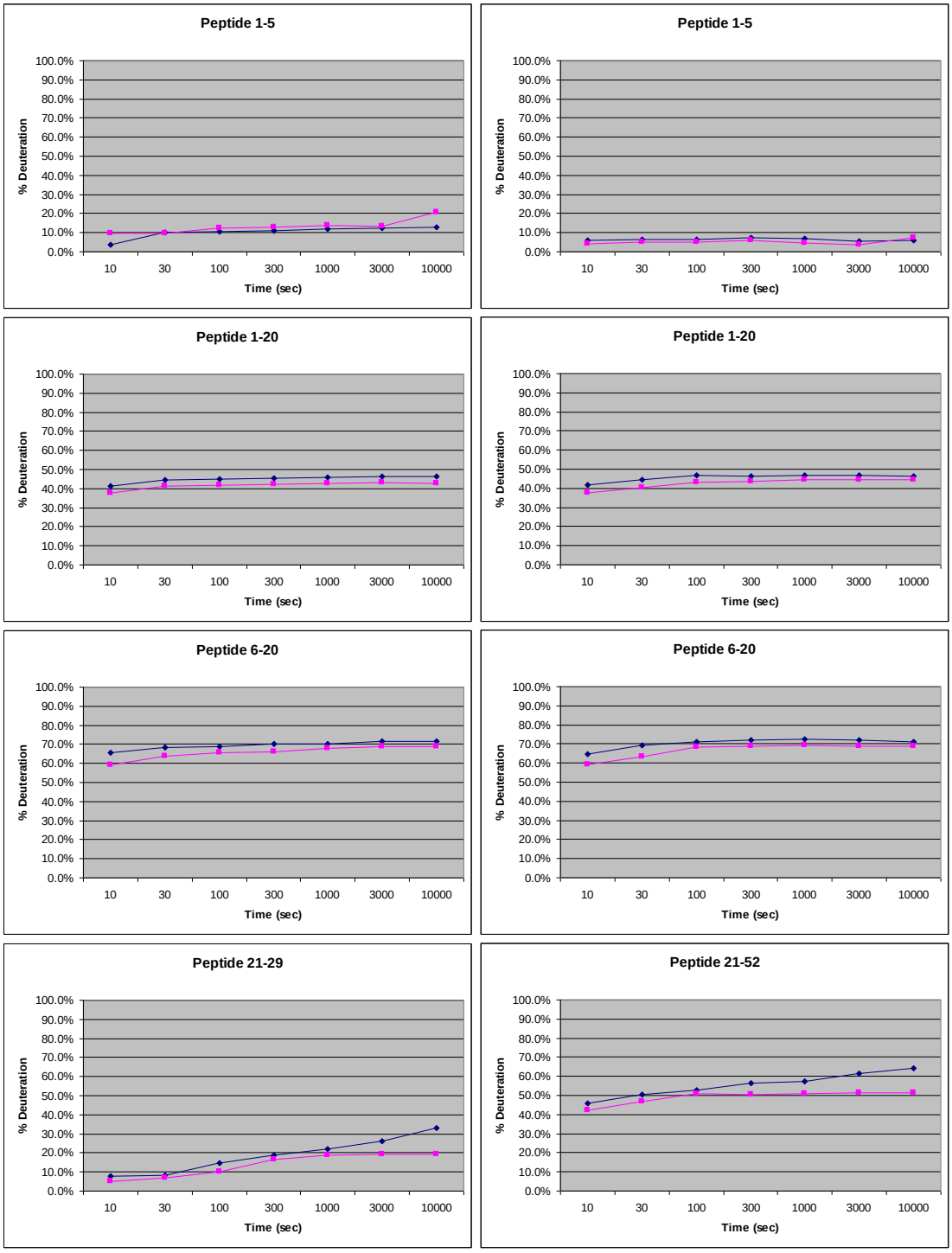
Table 1: Biochemical and physical properties of eukaryotic glycogen, amylopectin, and Lafora Bodies. Glycogen's continuous branching pattern confers water solubility. Amylopectin and Lafora Bodies both display discontinuous branching and higher phosphate content than glycogen.

	residues/ branch	branching pattern	water soluble	phosphate content
Eukaryotic Glycogen	12-14	continuous	yes	trace amount
Amylopectin	12-25	discontinuous	no	0.1 - 0.5% w/w
Lafora Body	12-30+	discontinuous	no	0.4 - 1.0% w/w

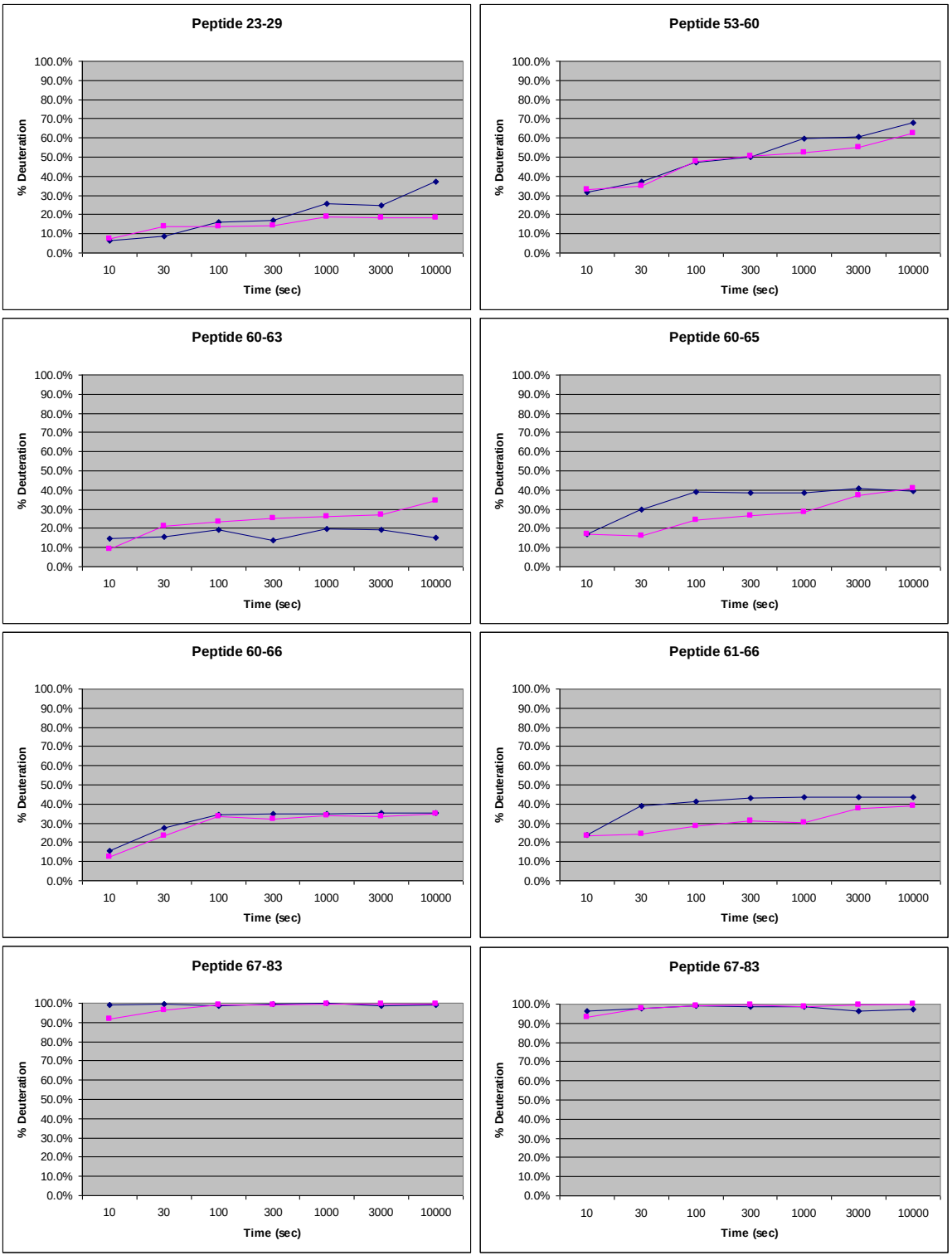
Table 2: Average deuteration level (%) for all Laforin proteins.

Protein	*Average deuteration level of peptides (%)*
Wild-type Laforin	63.35
G240S-Laforin	65.00
Y294N-Laforin	67.54
W32G-Laforin	69.67

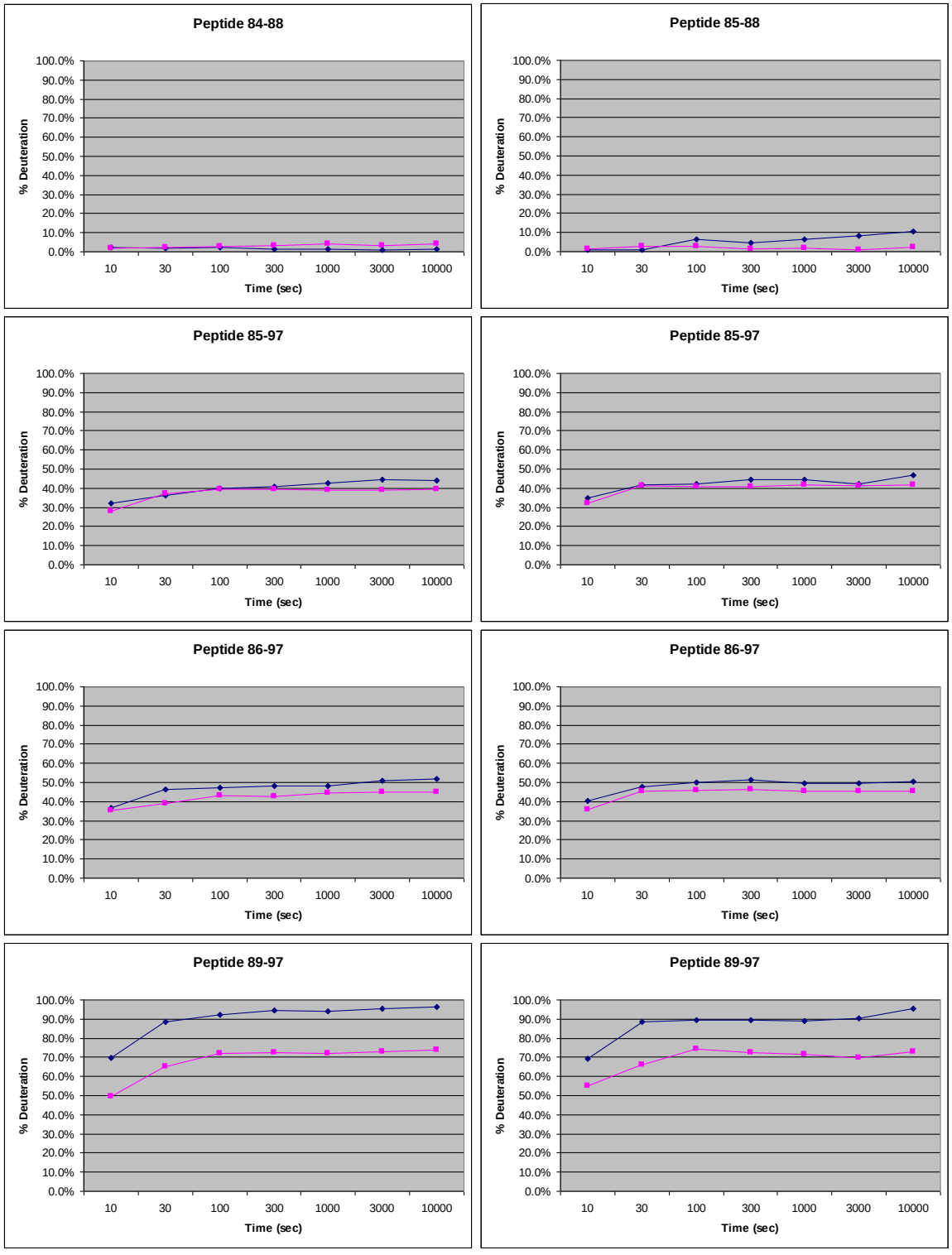
SUPPLEMENTAL FIGURES



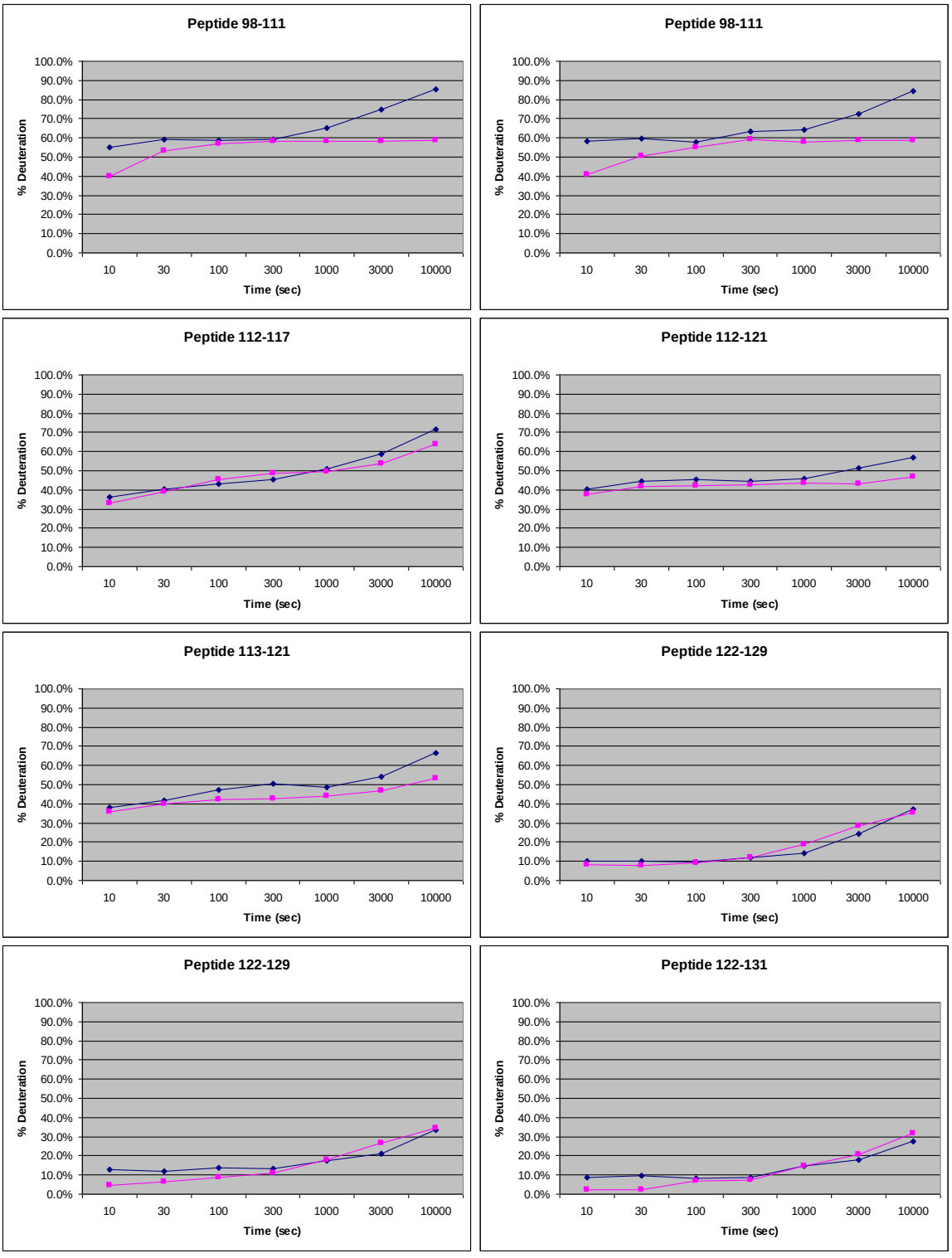
Supplemental Figure 1: Time dependent deuteration levels for glycogen-bound (pink squares) and ligand-free (blue diamonds) wild-type Laforin peptides. The deuteration levels were calculated as noted in Materials and Methods, according to the conclusions of Zhang and Smith[33].



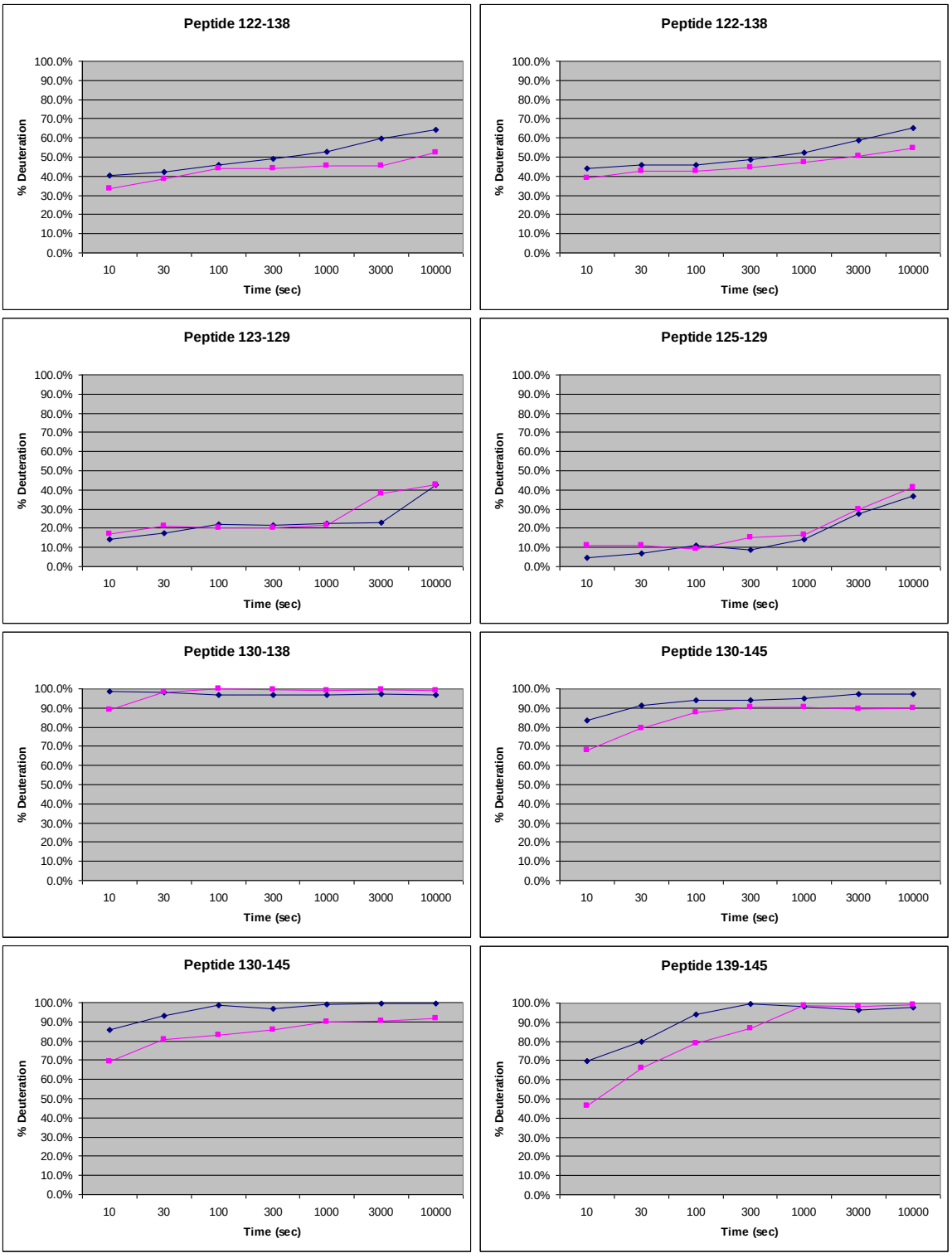
Supplemental Figure 1, Continued.



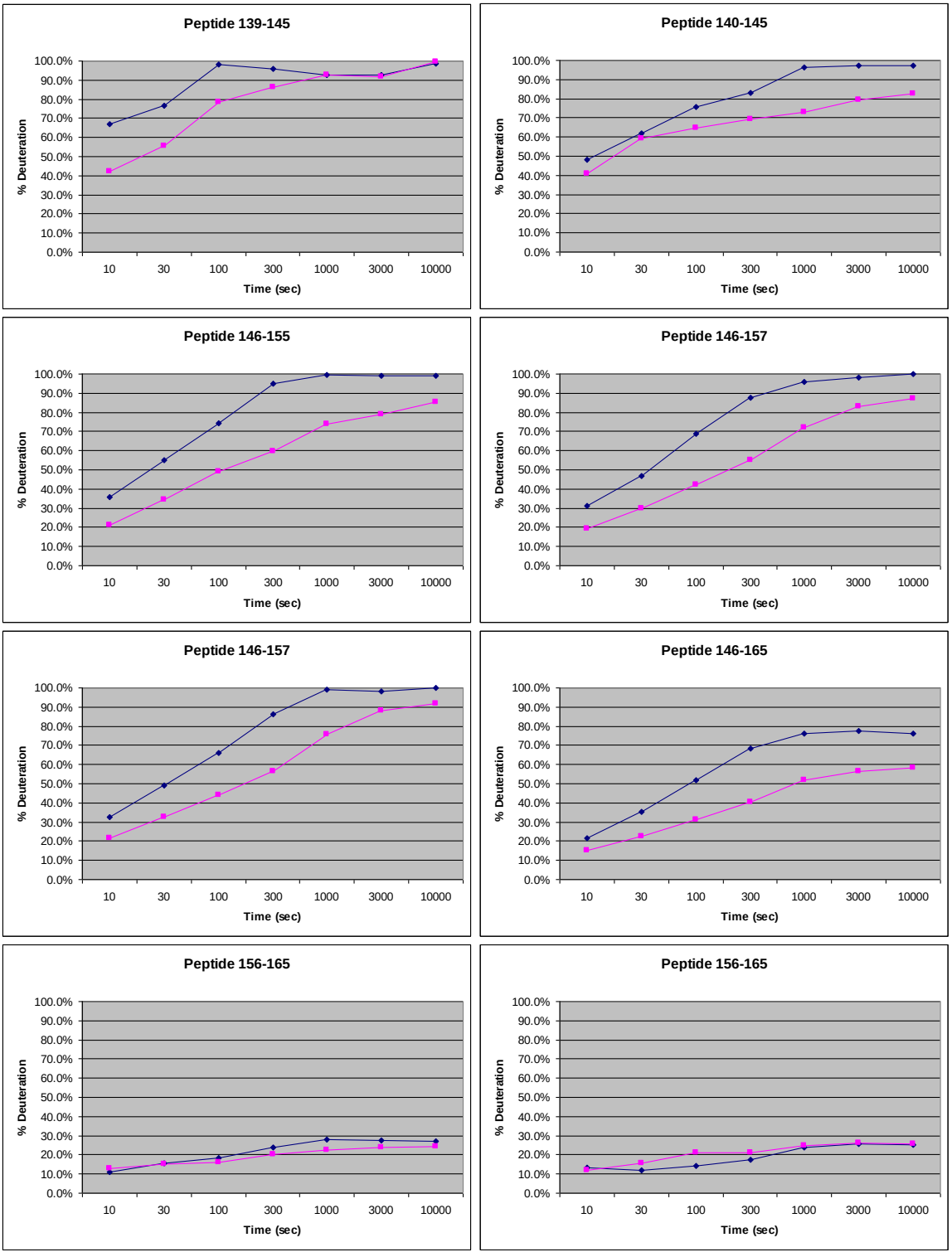
Supplemental Figure 1, Continued.



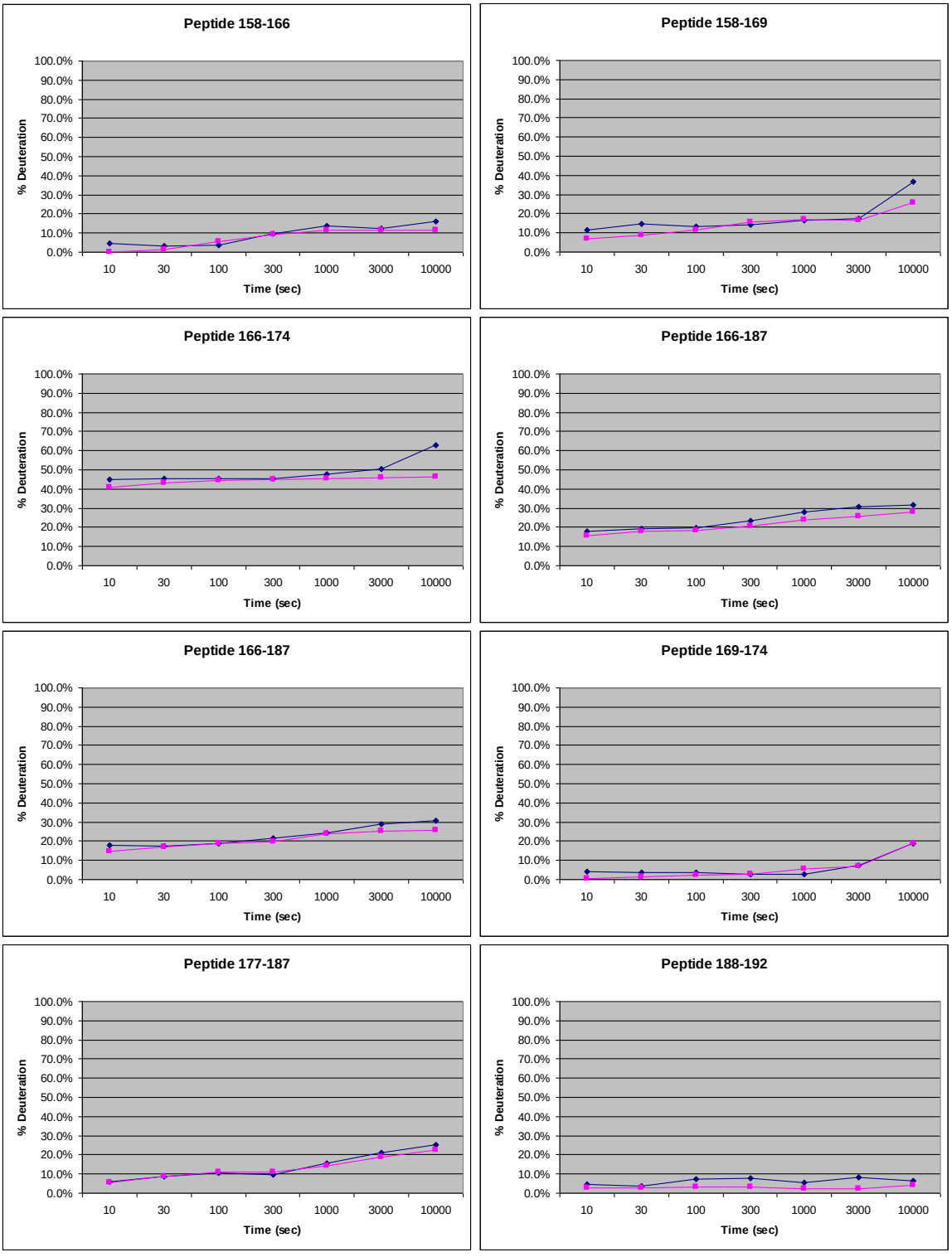
Supplemental Figure 1, Continued.



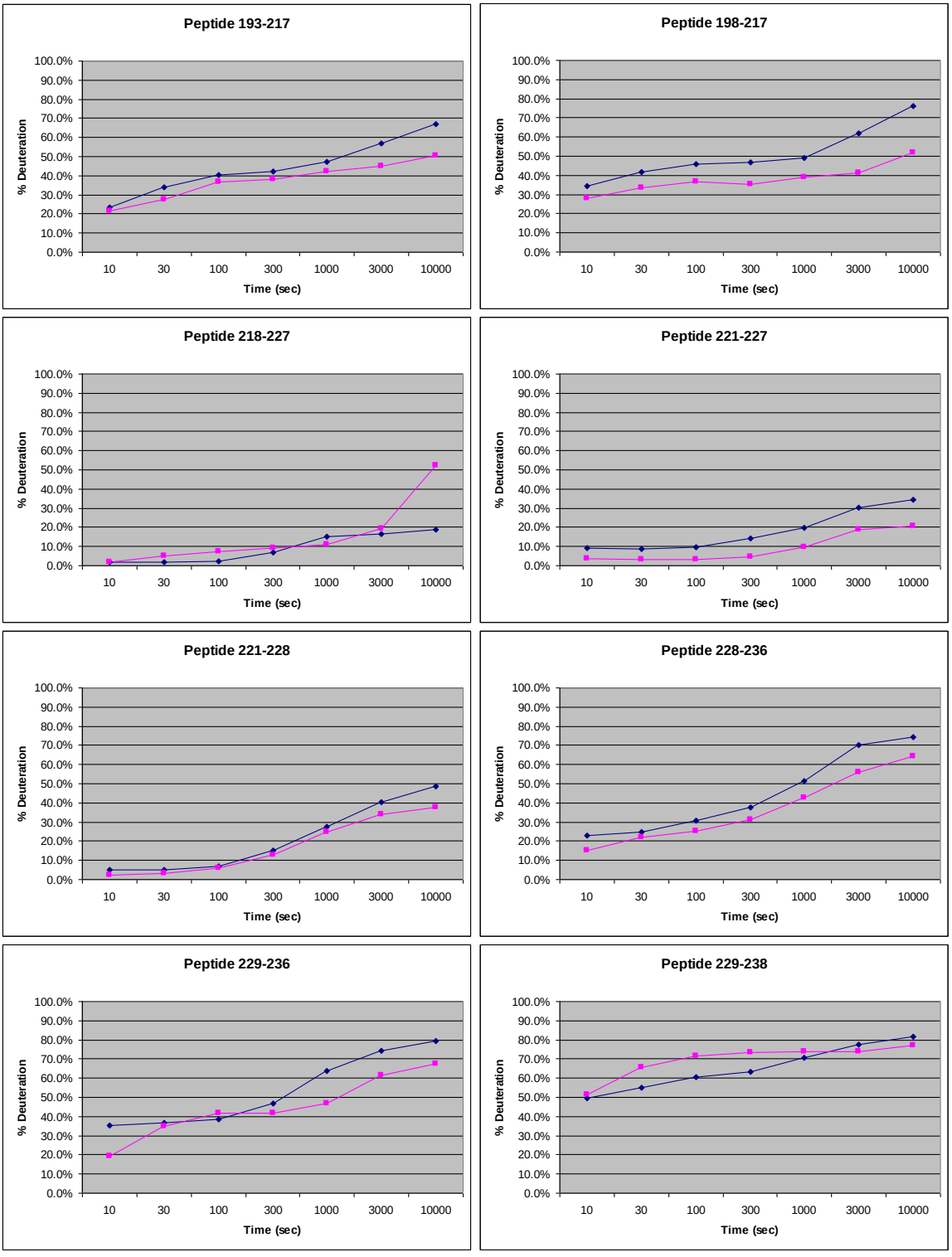
Supplemental Figure 1, Continued.



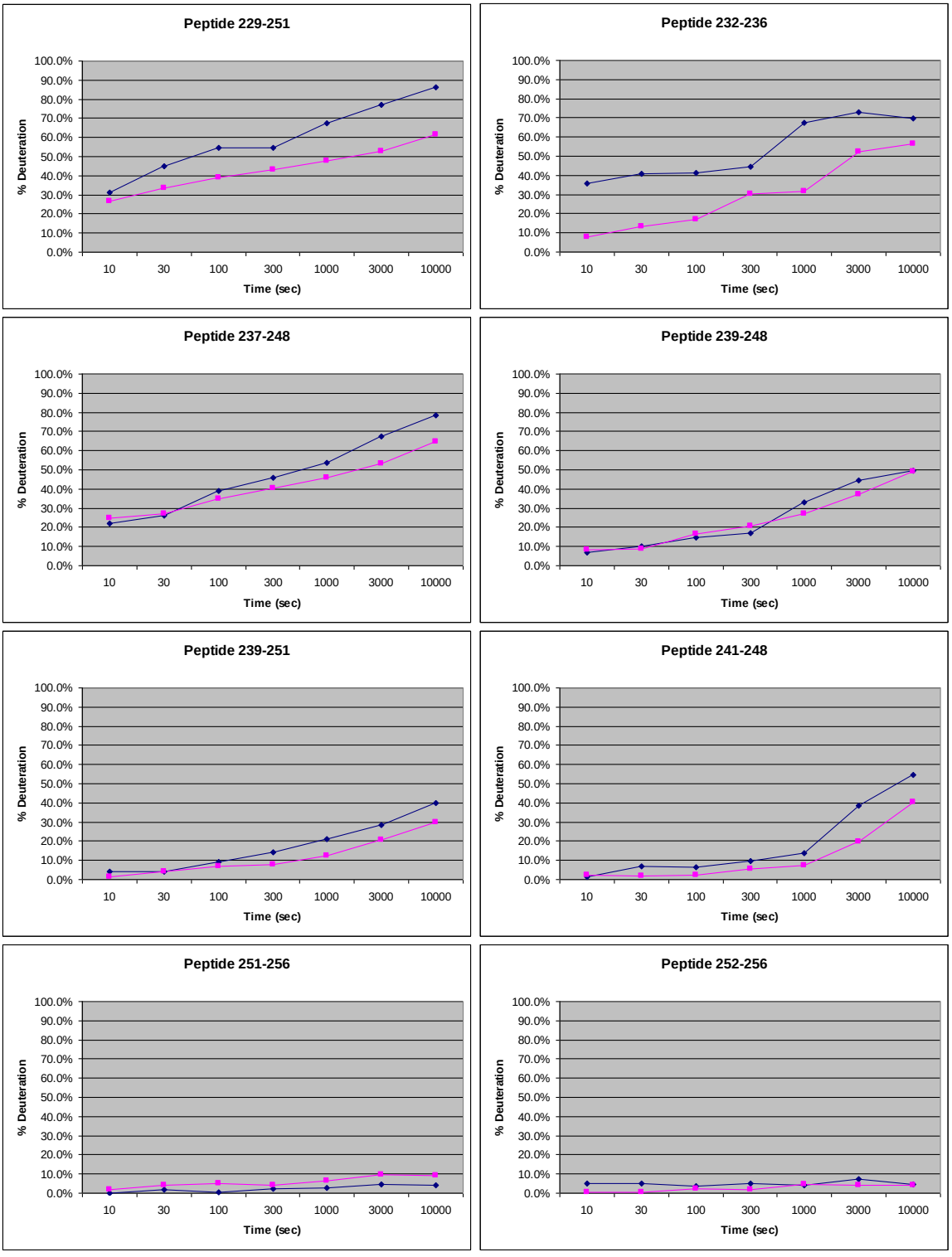
Supplemental Figure 1, Continued.



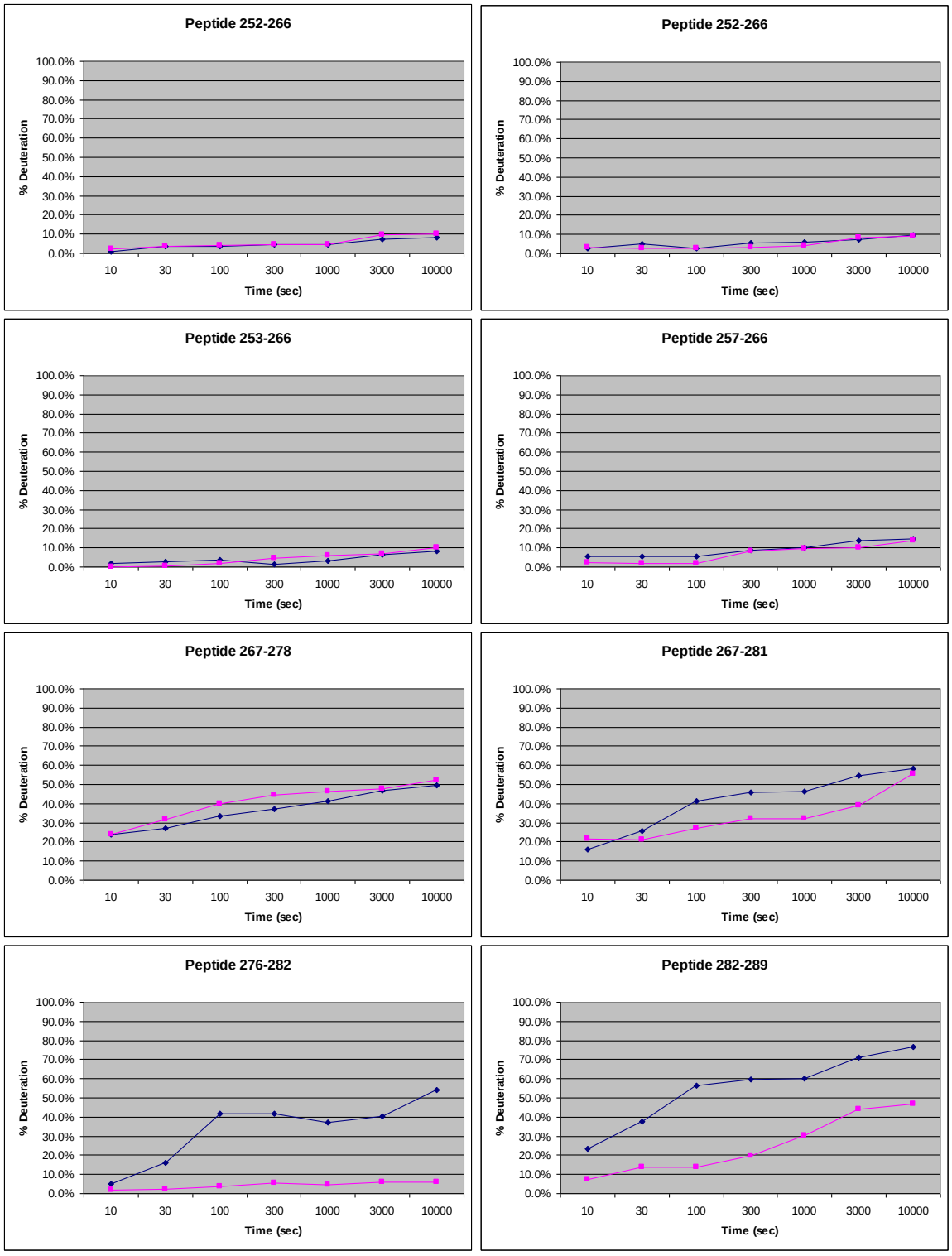
Supplemental Figure 1, Continued.



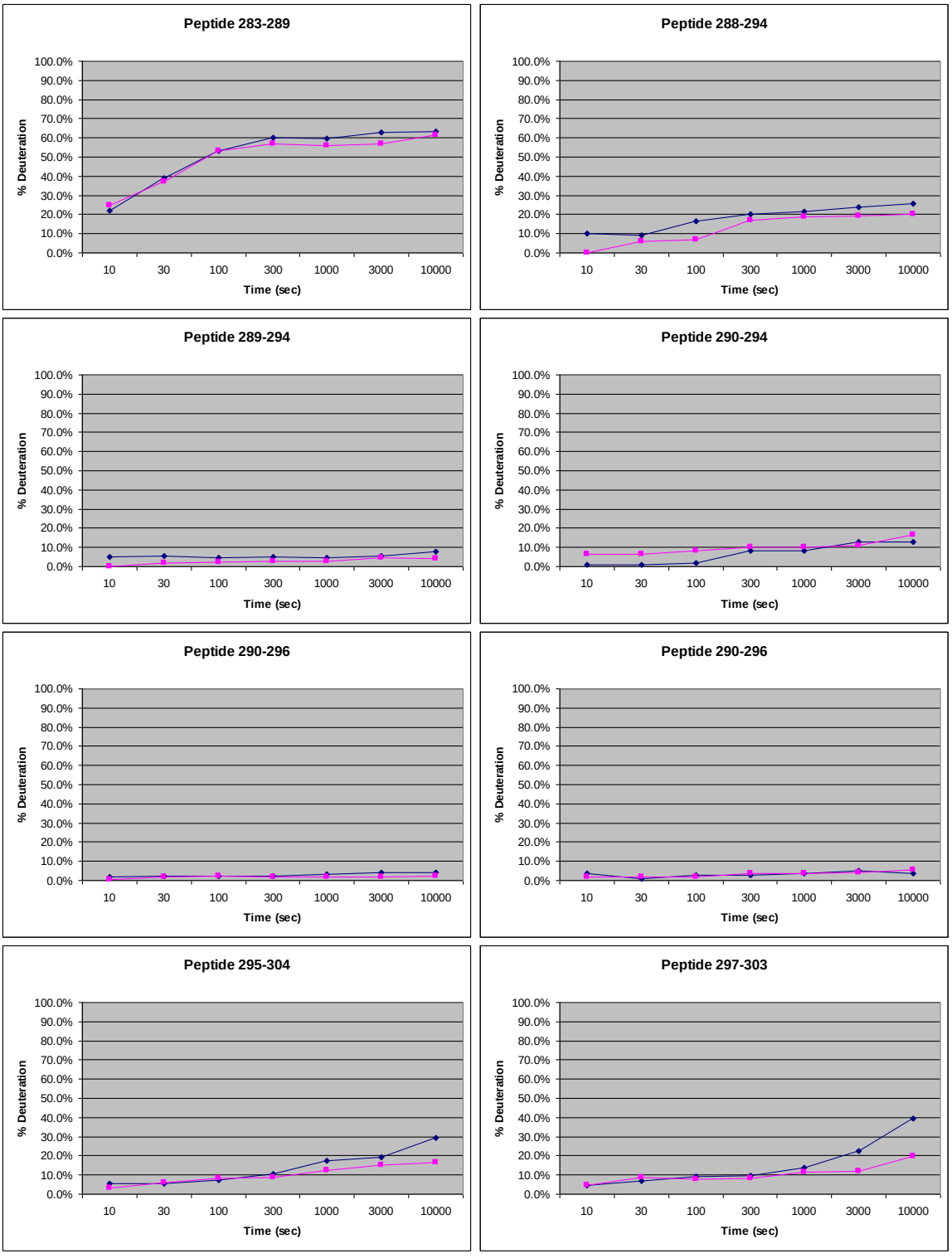
Supplemental Figure 1, Continued.



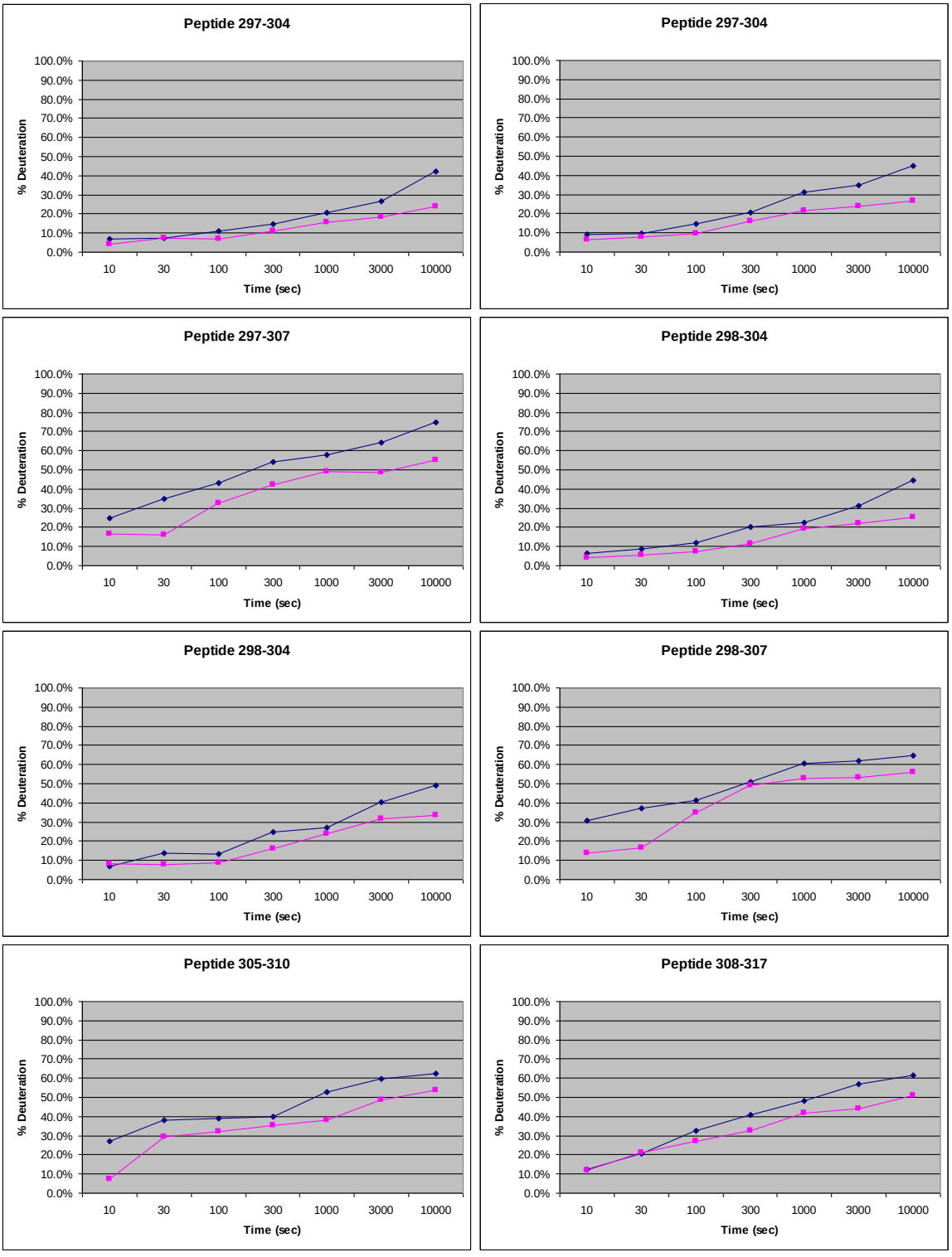
Supplemental Figure 1, Continued.



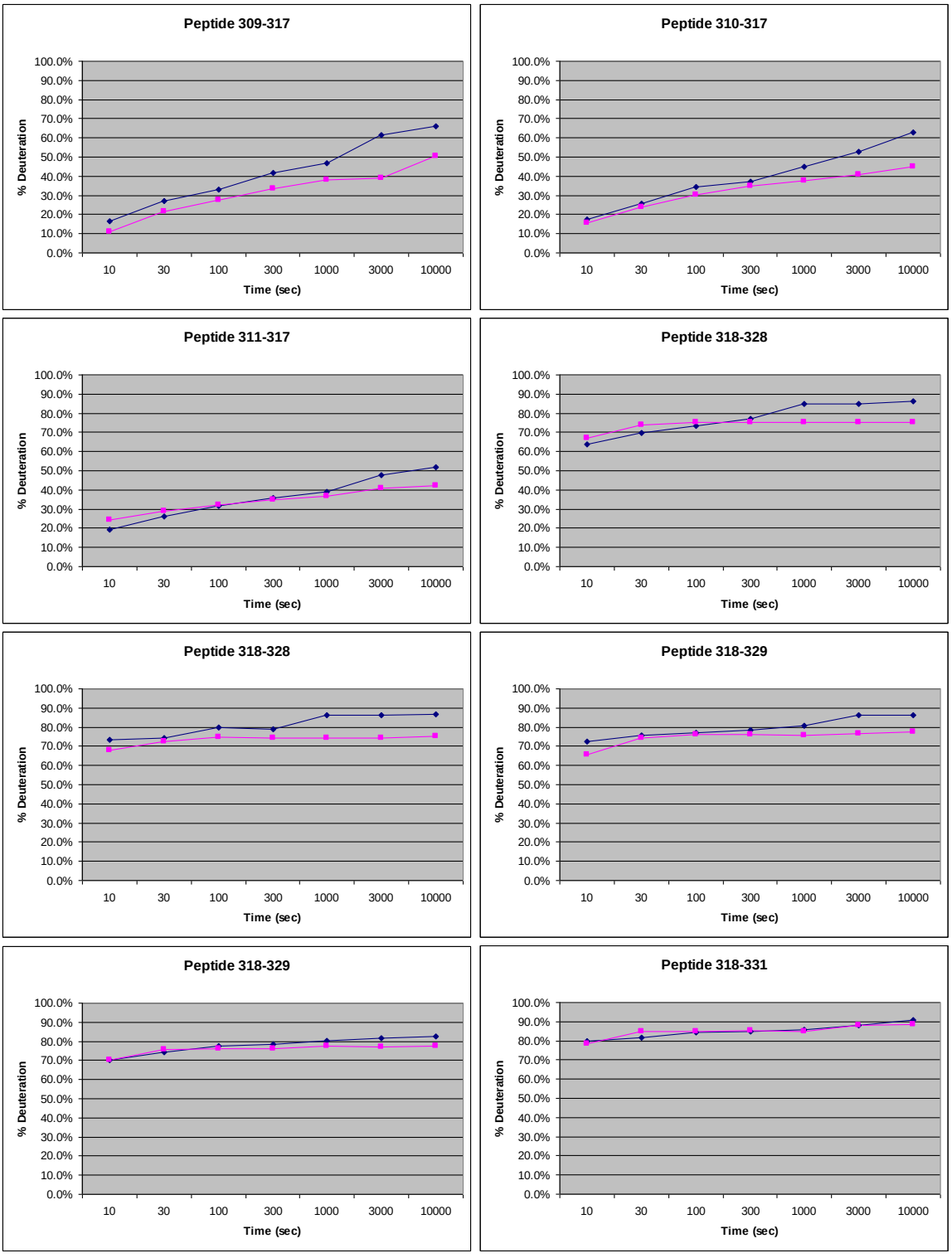
Supplemental Figure 1, Continued.



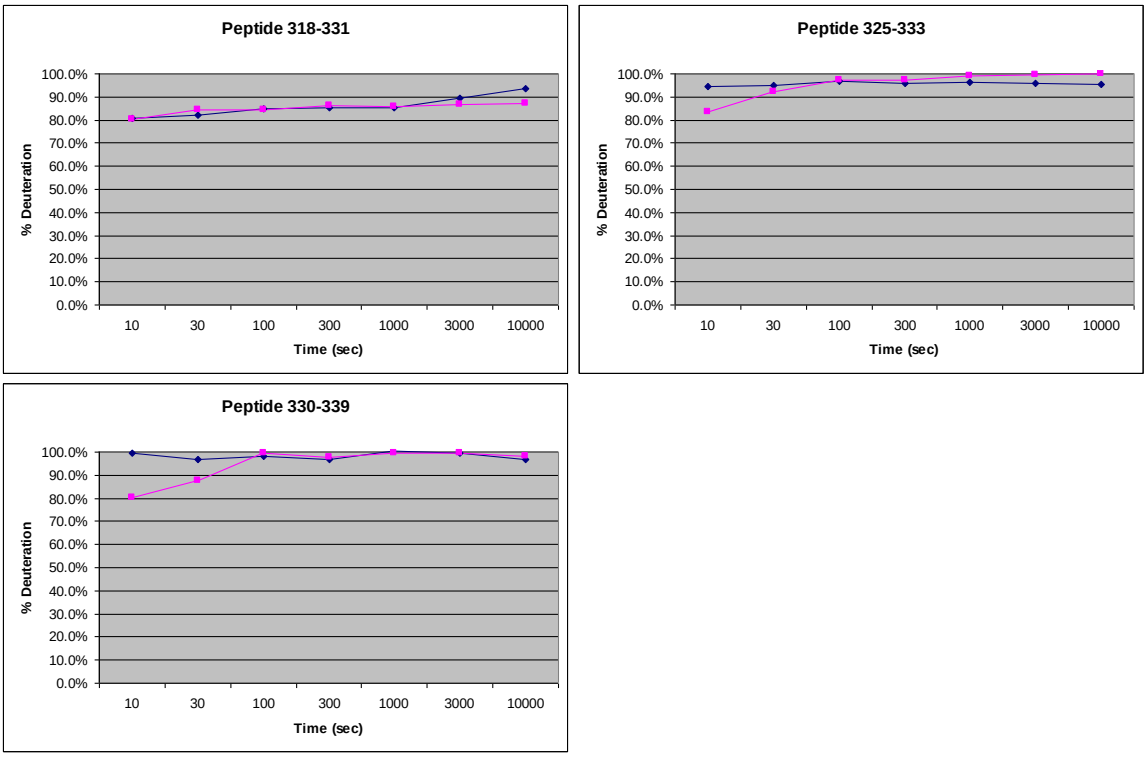
Supplemental Figure 1, Continued.



Supplemental Figure 1, Continued.



Supplemental Figure 1, Continued.



Supplemental Figure 1, Continued.

REFERENCES

1. Tsutsui Y, Wintrode PL. (2007). Hydrogen/deuterium exchange and mass spectrometry: a powerful tool for probing protein structure, dynamics and interactions. *Curr. Med. Chem*. 14:2344-2358.

2. Marcsisin SR, Engen JR. (2010). Hydrogen exchange mass spectrometry: what is it and what can it tell us? *Analytical and bioanalytical chemistry*. 397(3): 967-972.

3. Iacob RE, Engen JR. (2012). Hydrogen exchange mass spectrometry: are we out of the quicksand? *Journal of the American Society for Mass Spectrometry*. 23(6), 1003-1010.

4. Hastie KM, Liu T, Li S, King LB, Ngo N, Zandonatti MA, Woods VL Jr, de la Torre JC, Saphire EO. (2011). Crystal structure of the Lassa virus nucleoprotein-RNA complex reveals a gating mechanism for RNA binding. *Proceedings of the National Academy of Sciences of the United States of America*. 108(48), 19365-19370.

5. Cao J, Hsu Y, Li S, Woods VL, Dennis EA. (2011). Lipoprotein-associated phospholipase A(2) interacts with phospholipid vesicles via a surface-disposed hydrophobic α-helix. *Biochemistry*. 50(23), 5314-5321.

6. Damo SM, Phillips AH, Young AL, Li S, Woods VL Jr, Wemmer DE. (2010). Probing the conformation of a prion protein fibril with hydrogen exchange. *Journal of biological chemistry*. 285(42), 32303-32311.

7. Smith DL, Deng Y, Zhang Z. (1997). Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *Journal of mass spectrometry*. 32(2): 135-146.

8. Engen, JR. (2009). Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS. *Analytical chemistry.* 81(19), 7870-7875.

9. Englander SW, Mayne L, Bai Y, Sosnick TR. (1997). Hyd Hydrogen exchange: the modern legacy of Linderstrøm-Lang. *Protein Science*. 1997;6:1101-1109.

10. Englander SW, Sosnick TR, Englander JJ, Mayne L. (1996). Mechanisms and uses of hydrogen exchange. *Curr. Opin. Struct. Biol.* 6:18-23.

11. Ianzano L, Jang J, Chan EM, Zhao XC, Lohi H, Scherer SW, Minassian BA. (2005). Lafora progressive myoclonus epilepsy mutation database-EPM2A and NHLRC1 (EPM2B) genes. *Hum. Mutat.* 26:397.

12. Berkovic SF, So NK, Andermann F. (1991). Progressive myoclonus epilepsies: clinical and neurophysiological diagnosis. *Journal of clinical neurophysiology*. 8(3): 261-274.

13. Janeway R, Ravens JR, Pearce LA, Odor DL, Suzuki K. (1967). Progressive myoclonus epilepsy with Lafora inclusion bodies. I. Clinical, genetic, histopathologic, and biochemical aspects. *Arch Neurol.* 16(6):565-582.

14. Carpenter S, Karpati G. (1981). Sweat gland duct cells in Lafora disease: diagnosis by skin biopsy. *Neurology.* 31(12):1564-1568.

15. Yokoi S, Austin J, Witmer F, Sakai M. (1968). Studies in myoclonus epilepsy (Lafora body form). I. Isolation and preliminary characterization of Lafora bodies in two cases. *Arch Neurol.* 19(1):15-33.

16. Gentry MS, Dowen III RH, Worby CA, Mattoo S, Ecker JR, Dixon JE. (2007). The phosphatase laforin crosses evolutionary boundaries and links carbohydrate metabolism to neuronal disease. *J. Cell Biol.* 178:477-488.

17. Minassian BA, Lee JR, Herbrick JA, Huizenga J, Soder S, Mungall AJ, Dunham I, Gardner R, Fong CY, Carpenter S, Jardim L, Satishchandra P, Andermann E, Snead OC 3rd, Lopes-Cendes I, Tsui LC, Delgado-Escueta AV, Rouleau GA, Scherer SW. (1998). Mutations in a gene encoding a novel protein tyrosine phosphatase cause progressive myoclonus epilepsy. *Nature Genetics*. 20(2), 171-174.

18. Serratosa JM, Gómez Garre P, Gallardo ME, Anta B, de Bernabe DB, Lindhout D, Augustijn PB, Tassinari CA, Malafosse RM, Topcu M, Grid D, Dravet C, Berkovic SF, de Cordoba SR. (1999). A novel protein tyrosine phosphatase gene is mutated in progressive myoclonus epilepsy of the Lafora type (EPM2). *Human molecular genetics.* 8(2), 345-352.

19. Wang J, Stuckey JA, Wishart MJ, Dixon JE. (2002). A unique carbohydrate binding domain targets the lafora disease phosphatase to glycogen. *J Biol Chem* 277(4):2377-2380.

20. Alonso A, Rojas A, Godzik A, Mustelin T. (2004). The dual-specific protein tyrosine phosphatase family. *Topics in Curr. Genetics*. 5:333-358.

21. Ianzano L, Zhang J, Chan EM, Zhao XC, Lohi H, Scherer SW, Minassian BA. (2005). Lafora progressive myoclonus epilepsy mutation database-EPM2A and NHLRC1 (EMP2B) genes. *Human Mutation.* 26(4):397.

22. Tagliabracci VS, Turnbull J, Wang W, Girard JM, Zhao X, Skurat AV, Delgado-Escueta AV, Minassian BA, Depaoli-Roach AA, Roach PJ. (2007). Laforin is a glycogen phosphatase, deficiency of which leads to elevated phosphorylation of glycogen *in vivo*. *PNAS.* 104:19262-19266.

23. Worby CA, Gentry MS, Dixon JE. (2006). Laforin: A dual specificity phosphatase that dephosphorylates complex carbohydrates. *J. Biol. Chem.* 281(41):30412-30418.

24. Ganesh S, Agarwala KL, Amano K, Suzuki T, Delgado-Escueta AV, Yamakawa K. (2001). Regional and developmental expression of Epm2a gene and its evolutionary conservation. *Biochem Biophys Res Commun.* 283(5):1046-1053.

25. Hsu S, Kim Y, Li S, Durrant ES, Pace RM, Woods VL Jr, Gentry MS. (2009). Structural insights into glucan phosphatase dynamics using amide hydrogen-deuterium exchange mass spectrometry. *Biochemistry.* 48(41), 9891-9902.

26. Girard J, Lê KH, Lederer F. (2006). Molecular characterization of Laforin, a dual-specificity protein phosphatase implicated in Lafora disease. *Biochimie.* 88(12), 1961-1971.

27. Shi J, Koeppe JR, Komives EA, Taylor P. (2006). Ligand-induced conformational changes in the acetylcholine-binding protein analyzed by hydrogen-deuterium exchange mass spectrometry. *Journal of biological chemistry.* 281(17), 12170-12177.

28. Hamuro Y, Coales SJ, Morrow JA, Molnar KS, Tuske SJ, Southern MR, Griffin PR. (2006). Hydrogen/deuterium-exchange (H/D-Ex) of PPARgamma LBD in the presence of various modulators. *Protein science.* 15(8), 1883-1892.

29. Heegaard NH, Jørgensen TJ, Rozlosnik N, Corlin DB, Pedersen JS, Tempesta AG, Roepstorff P, Bauer R, Nissen MH. (2005). Unfolding, aggregation, and seeded amyloid formation of lysine-58-cleaved beta 2-microglobulin. *Biochemistry*. 44(11), 4397-4407.

30. Del Mar C, Greenbaum EA, Mayne L, Englander SW, Woods VL Jr. (2005). Structure and properties of alpha-synuclein and other amyloids determined at the amino acid level. *Proceedings of the National Academy of Sciences of the United States of America.* 102(43), 15477-15482.

31. Hamuro Y, Anand GS, Kim JS, Juliano C, Stranz DD, Taylor SS, Woods VL Jr. (2004). Mapping intersubunit interactions of the regulatory subunit (RIalpha) in the type I holoenzyme of protein kinase A by amide hydrogen/deuterium exchange mass spectrometry (DXMS). *Journal of Molecular Biology.* 340(5), 1185-1196.

32. Lanman J, Lam T, Barnes S, Sakalian M, Emmett MR, Marshall AG, Prevelige PE Jr. (2003). Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. *Journal of Molecular Biology.* 325(4), 759-772.

33. Zhang Z, Smith DL. (1993). Determination of amide exchange by mass spectrometry: A new tool for protein structure elucidation. *Prot. Sci.* 2:522-531.

34. Bai Y, Milne JS, Mayne L, Englander SW. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins: Structure, Function, and Genetics*.17:75-86.

35. Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochemical journal.* 382(3), 769-781.

36. Singh S, Ganesh S. (2009). Lafora progressive myoclonus epilepsy: a meta-analysis of reported mutations in the first decade following the discovery of the EPM2A and NHLRC1 genes. *Human mutation.* 30(5), 715-723.

37. Jia Z, Barford D, Flint AJ, Tonks NK. (1995). Structural basis for phosphotyrosine peptide recognition by protein tyrosine phosphatase 1B. *Science.* 268. 1754–1758.

38. Yuvaniyama J, Denu JM, Dixon JE, Saper MA. (1996). Crystal structure of the dual specificity protein phosphatase VHR. *Science* 272. 1328–1331.

39. Vander Kooi CW, Taylor AO, Pace RM, Meekins DA, Guo HF, Kim Y, Gentry MS. (2010). Structural basis for the glucan phosphatase activity of Starch Excess4. *Proceedings of the National Academy of Sciences of the United States of America.* 107(35), 15379-15384.

40. Denu J, Dixon JE. (1998). Protein tyrosine phosphatases: mechanisms of catalysis and regulation. *Curr. Opin. Chem. Biol.* 5:633-641.

41. Fernández Sánchez ME, Criado-García O, Heath KE, Garcia-Fojeda B, Medrano-Fernandez I, Gomez-Garre P, Sanz P, Serratosa JM, Rodriguez de Cordoba S. (2003). Laforin, the dual-phosphatase responsible for Lafora disease, interacts with R5 (PTG), a regulatory subunit of protein phosphatase-1 that enhances glycogen accumulation. *Human molecular genetics.* 12(23), 3161-3171.

42. Polekhina G, Gupta A, van Denderen BJ, Feil SC, Kemp BE, Stapleton D, Parker MW. (2005). Structural basis for glycogen recognition by AMP-activated protein kinase. *Structure 13*. 1453–1462.

43. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic acids research*. 37(Database issue), D233-D238.

44. Gentry MS, Dixon JE, Worby CA. (2009). Lafora disease: insights into neurodegeneration from plant metabolism. *Trends in biochemical sciences*. 34(12), 628-639.

45. Hsu S. Hydrogen/Deuterium-exchange (DXMS) Analysis of the Carbohydrate Phosphatase, Starch-excess 4. MS Thesis. University of California, San Diego. La Jolla, CA. 2008. Print.

46. Pei J, Kim B-H, Grishin N. (2008). PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res*. 36, 2295–2300.