

# Structure and Dynamics of N-Glycosylated Human Ribonuclease 1

Henry R. Kilgore,<sup>†</sup> Andrew P. Latham,<sup>†</sup> Valerie T. Ressler, Bin Zhang,<sup>\*</sup> and Ronald T. Raines<sup>\*</sup>

Cite This: *Biochemistry* 2020, 59, 3148–3156

Read Online

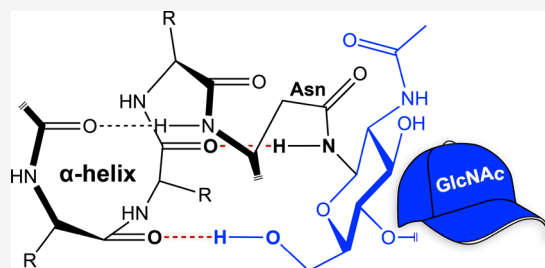
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Glycosylation is a common modification that can endow proteins with altered physical and biological properties. Ribonuclease 1 (RNase 1), which is the human homologue of the archetypal enzyme RNase A, undergoes N-linked glycosylation at asparagine residues 34, 76, and 88. We have produced the three individual glycoforms that display the core heptasaccharide,  $\text{Man}_5\text{GlcNAc}_2$ , and analyzed the structure of each glycoform by using small-angle X-ray scattering along with molecular dynamics simulations. The glycan on Asn34 is relatively compact and rigid, donates hydrogen bonds that “cap” the carbonyl groups at the C-terminus of an  $\alpha$ -helix, and enhances protein thermostability. In contrast, the glycan on Asn88 is flexible and can even enter the enzymic active site, hindering catalysis. The N-glycosylation of Asn76 has less pronounced consequences. These data highlight the diverse behaviors of  $\text{Man}_5\text{GlcNAc}_2$  pendants and provide a structural underpinning to the functional consequences of protein glycosylation.



Glycosylation is a highly prevalent co- or post-translational modification that is bestowed upon a fifth or so of proteins.<sup>1</sup> Glycosylation is known to have diverse consequences,<sup>2–8</sup> and articulating the precise effects of glycosylation on the structure and function of proteins is an important ongoing challenge.<sup>9–12</sup> We sought to extend extant knowledge in a well-defined context.

Vertebrate cells secrete endoribonucleases that enable a variety of biological functions stemming from their catalysis of RNA cleavage.<sup>13–16</sup> These small, stable, cationic proteins are known as pancreatic-type ribonucleases (ptRNases, EC 3.1.27.5). The prevalent bovine ptRNase, RNase A, has served as a model protein for much foundational research in the field of biochemistry,<sup>17–30</sup> as has its N-glycosylated variant, RNase B.<sup>31–34</sup> In humans, the most abundant and catalytically active ptRNase is RNase 1 (UniProtKB P07998).

Human RNase 1 provides a practical model for investigating structure–function relationships derived from glycosylation. The endogenous glycosylation of RNase 1 occurs on asparagine side chains within three sequons: Asn-Xaa-Ser/Thr. The three sites are glycosylated differentially according to cell and tissue type.<sup>35,36</sup> Consistent with the preference for glycosyltransferases to act upon Asn-Xaa-Thr over Asn-Xaa-Ser,<sup>37</sup> glycosylation at Asn34 is observed more frequently than at Asn76 or Asn88.<sup>38</sup> Moreover, Asn34 is in the most conserved sequon in ptRNases.<sup>31,39</sup> In previous work, we showed that glycosylation at position Asn34 leads to an 80% retention of catalytic activity and a 3 °C increase in the value of  $T_m$ , which is the temperature at the midpoint of the thermal transition between folded and unfolded states.<sup>39</sup> Monoglycosylations at residues 76 and 88 lead to enzymes with 39% and

16% of the catalytic activity, respectively, and a marginal increase in  $T_m$ .

Due to the wide variety of monosaccharide building blocks and the stochastic nature of the glycosyltransferase machinery, the multiplicity of glycoforms can be enormous.<sup>1,40–43</sup> Moreover, glycans tend to be more flexible than proteins, making their structural assessment problematic. Small-angle X-ray scattering (SAXS) can be used to probe structures in solution and thus provide insight into the conformational properties of biomolecules, especially when coupled with molecular dynamics simulations.<sup>44–47</sup> Here, we employ this synergistic strategy to reveal effects of endogenous glycans on the structure and dynamics of human RNase 1.

## MATERIALS AND METHODS

**Materials.** All chemicals and reagents were of commercial reagent grade (or better) and were used without further purification. Aqueous solutions were made with water that was generated with an arium Pro water purification system from Sartorius (Bohemia, NY) and that had a resistivity of  $\geq 18 \text{ M}\Omega \text{ cm}^{-1}$ .

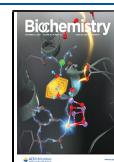
**Protein Preparation.** Unglycosylated human RNase 1 was produced in *Escherichia coli* and purified as described previously.<sup>48</sup> Human RNase 1 that is N-glycosylated on the side chain of Asn34 (NQQ), Asn76 (QNQ), or Asn88 (QQN)

**Special Issue:** The Glycoscience Issue

**Received:** March 7, 2020

**Revised:** June 16, 2020

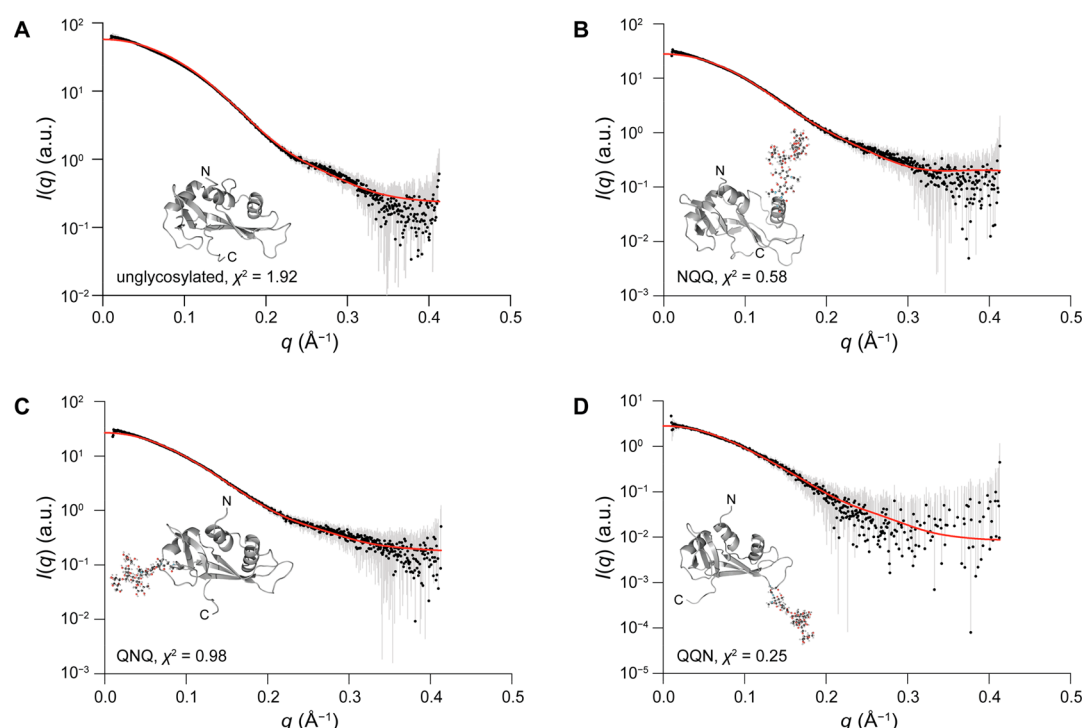
**Published:** June 16, 2020



**Table 1. Parameters Derived from Small-Angle X-ray Scattering, Dynamic Light Scattering, and Molecular Dynamics Simulation**

RNase 1	$R_{g,\text{Guinier}}$ (Å) <sup>a</sup>	$R_{g,P(r)}$ (Å) <sup>a</sup>	$D_{\text{max}}$ (Å) <sup>a</sup>	$V_p$ ( $\times 10^3$ Å <sup>3</sup> ) <sup>a</sup>	$D_{\text{exp}}$ ( $\mu\text{m}^2/\text{s}$ ) <sup>b</sup>	$R_h$ (Å) <sup>b</sup>	$R_{g,\text{sim}}$ (Å) <sup>c</sup>	$D_{\text{calc}}$ ( $\mu\text{m}^2/\text{s}$ ) <sup>c,d</sup>
unglycosylated	18.01 $\pm$ 0.27	15.80	44.44	18.8	133 $\pm$ 8	18.2 $\pm$ 0.1	15.3	115.1
NQQ glycoform	19.16 $\pm$ 0.26	16.97	45.22	20.0	112 $\pm$ 2	21.6 $\pm$ 0.3	16.0	107.9
QNQ glycoform	18.88 $\pm$ 0.15	16.89	45.85	20.3	116 $\pm$ 3	22.3 $\pm$ 0.4	16.3	106.1
QQN glycoform	20.72 $\pm$ 0.50	18.58	52.20	22.2	100 $\pm$ 2	23.8 $\pm$ 0.6	16.2	105.2

<sup>a</sup>Data from SAXS. <sup>b</sup>Data from DLS. <sup>c</sup>Data from molecular dynamics simulation. <sup>d</sup>Calculated by the method described in ref 58.



**Figure 1.** SAXS data for (A) unglycosylated RNase 1 and (B–D) its three glycoforms and molecular dynamics simulation fits to the SAXS data. Each panel displays a SAXS profile ( $\bullet \pm$  SE in gray), the simulated best-fit structure with respect to that profile (red lines), the  $\chi^2$  value between the experimental and simulated profiles, and an image of the best-fit structure from the simulation.

was produced with the MS strain of *Pichia pastoris* and purified as described previously.<sup>39</sup> Stocks of purified protein were prepared at 10 mg/mL in 10 mM BIS-TRIS buffer (pH 7.5).

**Dynamic Light Scattering.** Protein solutions were diluted to achieve concentrations of 1.0, 2.5, and 5.0 mg/mL in 10 mM BIS-TRIS buffer (pH 7.5), subjected to centrifugation at 15000g for 15 min, and filtered through a 0.22  $\mu\text{m}$  filter (GE Healthcare). These solutions were then monitored with a DynaPro Nanostar dynamic light-scattering instrument outfitted with a 658 nm laser in a 10  $\mu\text{L}$  quartz cuvette. Data were analyzed with the method of cumulants.<sup>49</sup>

**Small-Angle X-ray Scattering.** Samples were prepared for SAXS experiments with at least two concentrations to deduce potential concentration-dependent effects. Samples were prepared at concentrations of 1.0, 2.5, and 5.0 mg/mL, in 10 mM BIS-TRIS buffer (pH 7.5) containing glycerol [2% (v/v)] to attenuate radiation damage. Samples were subjected to centrifugation at 15000g for 15 min and filtered through a 0.22  $\mu\text{m}$  filter (GE Healthcare). SAXS data were collected at SIBYLS beamline 12.3.1 at the Advanced Light Source at the Lawrence Berkeley National Laboratory. An X-ray wavelength  $\lambda$  of 1.27 Å (11 keV) was used with a sample-to-detector distance of 2 m and collecting a scattering vector range of 0.009755–0.413 Å<sup>−1</sup>. Data sets were collected by performing

29 0.5-s exposures. The light path was generated by a superbend magnet to provide a flux of 1012 photons/s (1 Å wavelength) and detected with a Pilatus3 2M pixel array detector. Collected data sets were merged using the FrameSlice web server. The averaged spectra were then analyzed using PRIMUS to provide experimentally derived parameters and distance distribution functions for further analysis.<sup>50</sup>

**Mass Spectrometry.** Intact protein mass spectra were recorded on an Agilent Accurate Mass 6530 quadrupolar time-of-flight (QTOF) mass spectrometer. Samples were prepared by dilution to 10  $\mu\text{M}$  with a solution of H<sub>2</sub>O [94.9% (v/v)], acetonitrile [5% (v/v)], and formic acid [0.1% (v/v)], all LCMS grade. The resulting solution was passed through a 0.22  $\mu\text{m}$  filter prior to injection. Chromatography was performed with an Agilent biomolecule PLRP-S instrument (1000 Å, 2.1 mm  $\times$  50 mm, 50  $\mu\text{m}$ ) using an acetonitrile–water gradient. Intact mass spectra were recorded and subsequently deconvoluted.

**Simulation.** To gain molecular level insight into glycoprotein structure, we performed independent molecular dynamics simulations on the unglycosylated protein as well as each glycoform (NQQ, QNQ, and QQN). Simulations were performed on four proteins that were identical to those used experimentally, including strategic asparagine→glutamine

substitutions within sequons.<sup>39</sup> The starting model was based on chain X of Protein Data Bank entry 2q4g.<sup>51</sup> The missing C-terminal threonine residue and the N-terminal lysine residue were built into this protein structure with the program MODELER.<sup>52</sup> We used the CHARMM36 force field and set up the simulation using CHARMM-GUI, which allowed for appropriate asparagine→glutamine substitutions and glycosylation with Man<sub>5</sub>GlcNAc<sub>2</sub>.<sup>53,54</sup> Periodic boundary conditions were enforced in cubic simulation boxes with  $\geq 10$  Å between the glycosylated protein and the nearest side of the box, resulting in boxes with lengths of 58.69 Å (unglycosylated), 63.38 Å (NQQ), 75.76 Å (QNN), and 69.54 Å (QQN). Proteins were solvated with an ionic strength of 0.15 M, which also includes ions that counterbalance the charge of the protein. The structure was set up to run in the GROMACS simulation package.<sup>55,56</sup> Periods of 10 ns of NVT equilibration and 20 ns of NPT equilibration were performed before NVT production simulations of 1  $\mu$ s for the unglycosylated protein and 2  $\mu$ s for each glycoform.

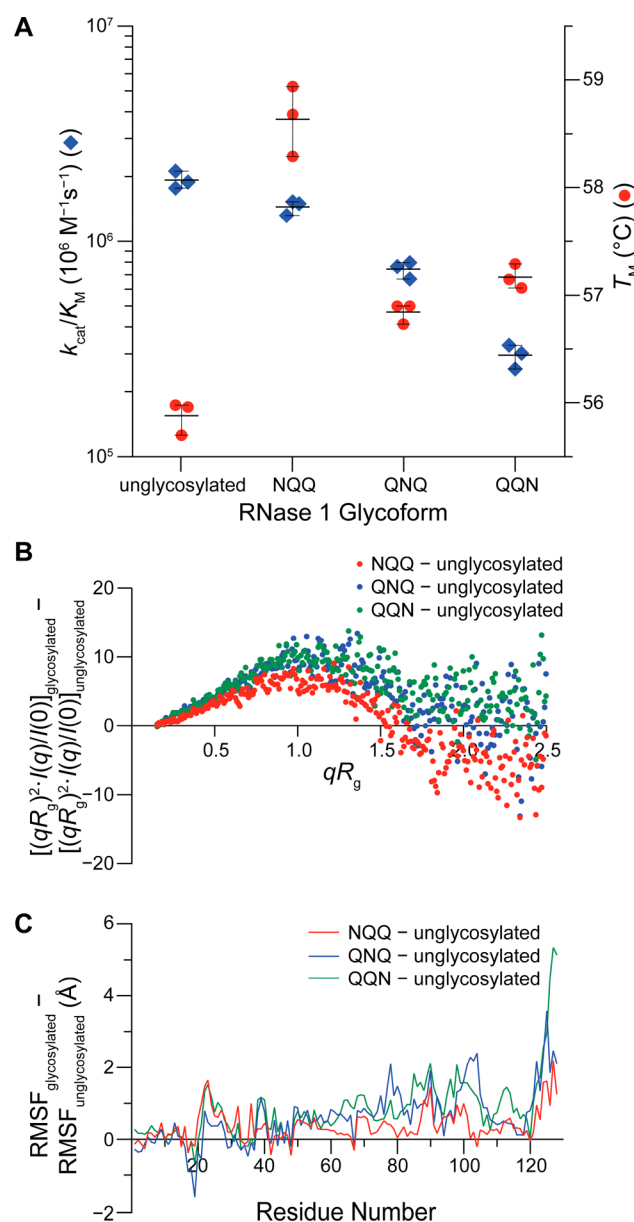
Simulated SAXS profiles were calculated and fitted with the program CRYSOLOG<sup>57</sup> to experimental profiles for both the entire trajectories and individual snapshots. Theoretical diffusion coefficients were calculated from a bead model based on the best-fit structures to SAXS data with the program HYDROPRO.<sup>58</sup> Contact maps were calculated on the basis of residues whose atoms were separated by  $<4.5$  Å with the MDAnalysis package.<sup>59,60</sup> Covariance matrices were built with the program ProDy<sup>61</sup> with protein  $\alpha$ -carbon and glycan-C1 positions, and principal component analysis was conducted by taking the eigenvalues and eigenvectors of the covariance matrix.<sup>61</sup> GROMACS analysis tools were used to calculate hydrogen bond lengths, radii of gyration, root-mean-square fluctuation (RMSF) values, and C $\alpha$  root-mean-square displacement (RMSD) values relative to the unglycosylated protein.<sup>55,56</sup>

## RESULTS AND DISCUSSION

### Biophysical Properties of RNase 1 Glycoforms.

Assessing the biophysical effects of glycosylation requires control over both the location of glycosylation and the composition of the appended glycan. As described previously,<sup>39</sup> we accessed individual glycoforms of human RNase 1 by biosynthesis using strategic substitution of asparagine residues with glutamine (which is not a substrate for oligosaccharyltransferase). We then characterized the unglycosylated protein and its three glycoforms with QTOF-LCMS. Integration of each peak in the deconvoluted mass spectra revealed that the majority of protein in solution retains a single Man<sub>5</sub>GlcNAc<sub>2</sub> glycan (Figure S1). Some additional mannosylation is evident in these preparations, mimicking biological glycosylation patterns.

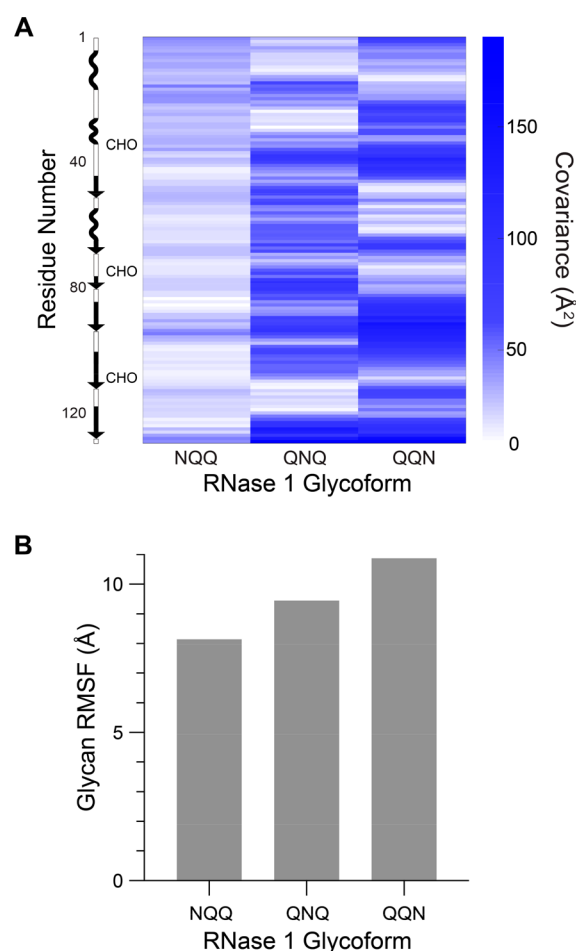
With authentically glycosylated proteins in hand, we began analyzing the structural and biophysical effects of RNase 1 glycosylation. We used dynamic light scattering (DLS) to ensure that our samples were monodisperse and to establish an effective hydrodynamic radius for the glycosylated proteins. We found there was not a significant increase in the hydrodynamic radius ( $R_h$ ) as a function of concentration (Table 1 and Figure S2). DLS measurements did, however, demonstrate a 10–30% increase in the Stokes radius upon glycosylation. The average hydrodynamic radii varied between glycosylation sites by 0.1–0.2 nm. Whereas DLS data must be interpreted in the context of simple geometric objects, the data



**Figure 2.** Relationship between the catalytic activity and protein fluctuation of RNase 1 and its three glycoforms. (A) Ribonucleolytic activity and thermostability of the enzymes. Values of  $k_{cat}/K_M$  and  $T_m$  are the mean ( $\pm$ SD) from ref 39. (B) Kratky difference plot showing the differential flexibility endowed by N-glycosylation at position 34, 76, or 88. (C) RMSF difference plot of amino acid residues in unglycosylated RNase 1 and its glycoforms.

were consistent with the proteins being monomeric and adumbrated higher information techniques that provide deeper insight into the RNase 1 glycans.

The conformational plasticity endemic to glycans often results in glycosylated proteins being refractory toward analysis by traditional X-ray crystallography. In addition, a rigid crystalline structure would not necessarily be indicative of the overall physical properties of glycoproteins. To circumvent these issues, we conducted SAXS experiments on the unglycosylated RNase 1 and its three glycoforms. We collected SAXS data at three different concentrations and merged the data sets after verifying that the radius of gyration was constant with respect to changes in concentration (Figure S3). As



**Figure 3.** Relationship between the flexibility of the glycan and protein in RNase 1 glycoforms. (A) Absolute value of protein–glycan covariance as a function of protein residue for RNase 1 glycoforms.  $\alpha$ -Helices (curves),  $\beta$ -strands (arrows), and glycosylated residues (CHO) are indicated on the ordinate. (B) Average RMSF of all glycan residues in each glycoform.

anticipated, glycosylation increases the  $R_g$  value of RNase 1 in a site-dependent manner (Table 1) and is correlated with increased hydrodynamic radii. Differences arose between the  $R_{g,Guinier}$  and  $R_{g,P(r)}$  values, but trends are consistent across all members of the data set. The variation in  $R_{g,Guinier}$  and  $R_{g,P(r)}$  values between all three glycoforms is  $\sim 5\%$ , whereas a 10% difference arose between the radii of gyration for unglycosylated and glycosylated RNase 1. Collectively, the trends in  $R_{g,Guinier}$  and  $R_{g,P(r)}$  indicate that glycosylation increases the size of RNase 1 through a glycosylation-site-specific mechanism. By all methods employed, glycosylation of RNase 1 at positions 34 and 76 leads to a more compact protein than does glycosylation at position 88.

An analysis of distance distributions extracted from SAXS data can provide insight into the geometric properties of a molecule. Specifically, distance distributions can provide a “string-length” distance,  $D_{max}$ , which is the largest length of the molecule. In addition, the distribution function can provide insight into the shape of the molecule and its flexibility. Analysis of the distance distribution functions for the proteins demonstrates their globularity (Figure S4). The perturbation of a glycan increases the  $D_{max}$  value, especially for the QQN glycoform (Table 1).

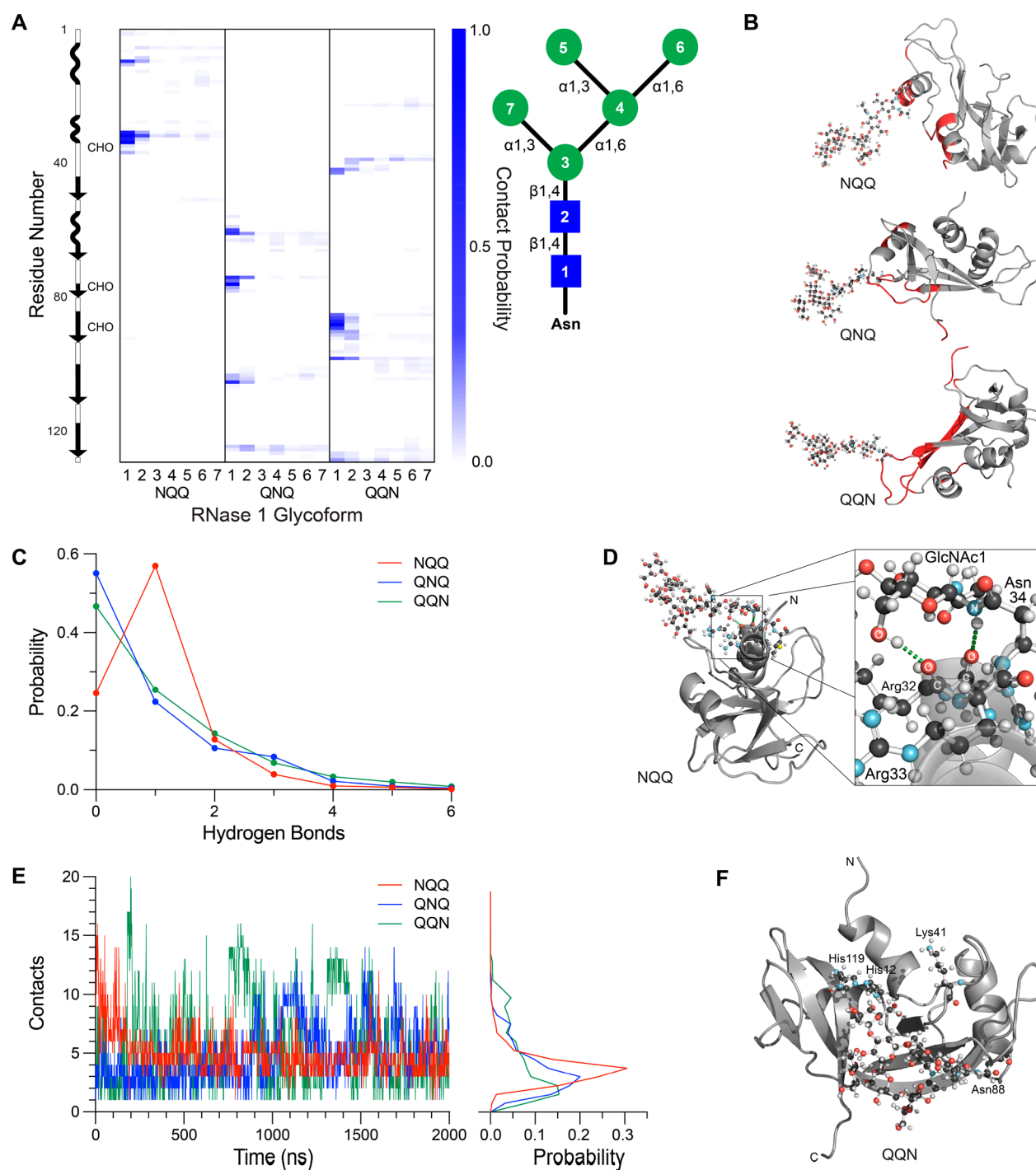
SAXS is an effective method for obtaining low-resolution structural data for biomolecules in solution. Its resolution is limited, however, by orientational averaging, resulting in a one-dimensional profile.<sup>44,45,47</sup> This limitation has led to approaches for producing structures by matching SAXS data with possible structures or by using SAXS curves to guide molecular dynamics calculations.<sup>62–64</sup> Molecular dynamics force fields continue to evolve to match experimental data more closely,<sup>65–68</sup> and recent developments expanded their ability to model glycosylated proteins.<sup>54,69–71</sup> We find that all-atom explicit solvent simulations on the unglycosylated protein and each glycoform complement the experimental results. Our modeled structures are consistent with SAXS and DLS data and suggest that the chemical environment around a glycosylation site governs the protein–glycan interactions.

To gain atomistic insight from SAXS curves, we produced an ensemble of possible structures with molecular dynamics calculations. We then computed the theoretical SAXS spectrum of each individual structure with the program CRY SOL.<sup>57</sup> The best individual fits from CRY SOL are shown in Figure 1. Because SAXS is an ensemble measurement, we also calculated the SAXS profile of the entire simulated ensemble (Figure S5) and tracked the radius of gyration throughout our simulation (Table 1 and Figure S6). Both the individual fits and the ensemble fits match the experimental profile, with single structures reaching  $\chi^2 < 2$  for each system. To verify the simulation results further, we used best-fit structures from individual SAXS fittings to produce a theoretical diffusion constant using the program HYDRO-PRO.<sup>58</sup> The trends in these modeled diffusion constants ( $D_{calc}$ ) match those in the experimental diffusion constants ( $D_{exp}$ ) (Table 1). It is worth noting that the simulated values for  $R_g$  are smaller than the experimental ones (Table 1). This discrepancy could arise from, in addition to force field inaccuracy, the heterogeneity of glycosylated states in experiments, as shown in Figure S1. In particular, besides the most prevalent glycan (which is modeled in our simulations), the presence of larger glycans could lead to the observed larger experimental values of  $R_g$ . (Simulating all possible glycoforms is, however, beyond the scope of this work.) We note too that the protein–glycan systems maintain the overall fold of the protein. The values of the  $\alpha$ -carbon RMSD relative to the crystal structure indicate that all simulations stay close to the unglycosylated structure, with an average RMSD of  $\sim 3$  Å (Figure S7).

**Effects of Glycosylation Are Site-Dependent.** All mammalian species retain at least one N-glycosylation site in their RNase 1 homologue,<sup>31</sup> suggestive of functional utility. Still, the physiological role of glycosylation in RNase 1 homologues is not clear. Previously, we demonstrated that glycosylation of Asn34 increases the thermostability of the protein but has little effect on its catalytic activity (Figure 2A).<sup>39</sup> Meanwhile, the QNQ and QQN glycoforms have unperturbed thermostability but diminished ability to catalyze RNA cleavage. We searched for the structural differences that underlie these functional consequences.

Normalized Kratky plots provide insight into the compaction, domains, and organization of a biomolecule. We plotted normalized Kratky plots from the experimental SAXS data of each glycoform alongside the unglycosylated RNase 1 to assess if glycosylation alters the fold or flexibility of the protein (Figure S8). At first glance, we observed no significant differences in the rough geometric properties of the four





**Figure 4.** Effect of the glycosylation site in RNase 1 on protein–glycan interactions during molecular dynamics simulations. (A) Protein–glycan contact probability varies according to glycosylation site.  $\alpha$ -Helices (curves),  $\beta$ -strands (arrows), and glycosylated residues (CHO) are indicated on the ordinate. The inset shows the structure of the  $\text{Man}_5\text{GlcNAc}_2$  glycan and the numbering system used on the abscissa. (B) Sites (red) at which the protein and glycan are in contact for >3% of the simulation. (C) Probability of observing a hydrogen bond between the  $\text{Man}_5\text{GlcNAc}_2$  glycan and RNase 1 in the three glycoforms. (D) Hydrogen bonds (green dashes) formed frequently between main-chain oxygens and residues of the NQQ glycoform. The side chains of residues 31–35 are shown explicitly. (E) Number of protein–glycan contacts in the three glycoforms tracked as a function of time (left) and by probability (right). (F) QQN glycoform with its glycan bound in the enzymic active site, a conformer that corresponds to the peaks in protein–glycan contacts in panel E. The side chains of active-site residues His12, Lys41, and His119 are shown explicitly.

molecules, in agreement with the subtle differences observed in their distance distribution functions (Figure S4). To probe more deeply, we plotted the difference,  $[(qR_g)^2 \times I(q)/I(0)]_{\text{glycosylated}} - [(qR_g)^2 \times I(q)/I(0)]_{\text{unglycosylated}}$  (Figure 2B). This difference shares a local maximum near  $qR_g = 1.0$ ,

followed by a decay to zero. The positive value of the  $[(qR_g)^2 \times I(q)/I(0)]_{\text{glycosylated}} - [(qR_g)^2 \times I(q)/I(0)]_{\text{unglycosylated}}$  curves between  $qR_g$  values of 0.25 and 1.75 is consistent with a less compact structure.<sup>72</sup> Comparison of the Kratky difference curves across the glycoforms suggests that the NQQ glycoform

is more compact than either the QNQ or the QQN glycoform. Collectively, the differences, though subtle, underline changes that could arise from site-specific glycosylation.

To probe protein flexibility further, we computed the RMSF as a function of residue number during the course of the molecular dynamics simulations. This metric provides a measure of the flexibility of each residue. Similar to what was observed in the Kratky difference plots derived from SAXS data, the simulations reveal that the NQQ glycoform is less flexible than either the QNQ or the QQN glycoform (Figure 2C). The degree of flexibility of a glycoform correlates inversely with its  $T_m$  value (cf. Figure 2A).

Although the RMSF results indicate that fluctuations tend to increase upon glycosylation, they cannot reveal how the glycan is causing these changes. To relate glycan motion to protein motion, we first computed the covariance map for the fluctuations of unglycosylated RNase 1 and its glycoforms (Figure S9). These maps demonstrate that fluctuations can propagate throughout the protein. This interplay between the protein and its pendant glycan is readily apparent in the protein–glycan covariance (Figure 3A), which increases with glycan RMSF (Figure 2C). To examine the protein–glycan interaction more closely, we computed the average RMSF of all glycan residues (Figure 3B). As with thermostability, we find that glycan flexibility correlates inversely with catalytic activity (Figure 2A). Relaying glycan motion to the protein could perturb its structural organization and conformational fluctuation that are crucial for function.

**Protein–Glycan Interactions.** As glycosylation-site-dependent changes in protein flexibility seem to stem from glycan mobility, we reasoned that the subtle differences between the glycoforms were a product of their local chemical environment. Certain protein conformations could produce more stable protein–glycan interactions. Accordingly, we plotted residues at which different glycan residues contacted the protein (Figure 4A,B). We find that the NQQ glycoform makes more longer-lived contacts than do the QNQ and QQN glycoforms. There are no aromatic residues near Asn34 to provide stabilizing interactions<sup>73,74</sup> with a pendant glycan. Further examination revealed, however, that the NQQ glycoform engages much more frequently in hydrogen bonding than do the other glycoforms (Figure 4C). In particular, hydrogen bonds connect the first GlcNAc residue with carbonyl groups at the C-terminus of an  $\alpha$ -helix (Figure 4D). These hydrogen bonds, which are donated largely by the 6-OH and 2-NH of GlcNAc1 and the side-chain NH of Asn34, fix the conformation of the glycan and explain its lower flexibility. In turn, the glycan appears to “cap” and thus stabilizes the  $\alpha$ -helix,<sup>75</sup> which is consistent with the enhanced thermostability of the NQQ glycoform (Figure 2A) and likely arises (in part) from a cooperative increase in the strength of hydrogen bonds<sup>76</sup> and  $n \rightarrow \pi^*$  interactions<sup>77</sup> within the main chain of the  $\alpha$ -helix. The ability of GlcNAc, which is replete with hydrogen bond donors, to be an effective C-cap for an  $\alpha$ -helix has not been described previously. Notably, asparagine alone is neither an abundant nor an effective C-cap.<sup>78,79</sup> The role for N-linked glycans that we have discovered in RNase 1 could be common, given their high frequency near the termini of  $\alpha$ -helices<sup>37</sup> and their being well-tolerated there.<sup>80</sup> Still, glycan context is important; our data with an N-glycosylated protein contrast with the observation that O-glycosylation of a C-terminal serine or threonine residue with GlcNAc slightly destabilizes the  $\alpha$ -helix of a model peptide.<sup>81</sup>

In comparison to the NQQ glycoform, the QNQ and QQN glycoforms exhibit a less-pronounced hydrogen-bonding network. Our simulations suggest that the glycan in the QNQ glycoform engages in hydrogen bonds with Gln60, Glu61, Lys62, and Ser77. Still, none of these protein–glycan hydrogen bonds are as populated as those in the NQQ glycoform. The QQN glycoform has the most flexible glycans (Figure 4E). Most importantly, several conformations showed its glycans within the active-site cleft (Figure 4F). This behavior is apparent from a transient increase in the number of contacts throughout the simulation. Steric occlusion caused by the entry of glycan into the enzymic active site is consistent with the observation that the QQN glycoform has the lowest catalytic activity of the three glycoforms (Figure 2A).<sup>39</sup>

The significance of the observed protein–glycan interactions is supported further by the principal components of the covariance matrix. The first principal component accounts for a significant amount of variance within our simulations (Figure S10), so we chose to examine that component in greater detail. For the NQQ glycoform, this motion represents the formation and breaking of the hydrogen bond between the main-chain oxygen of Arg31 and the glycan (Movie 1).

The glycan in the QNQ glycoform forms fewer hydrogen bonds with the protein, and the most significant contribution to the variance is movement of the QNQ glycan in solution (Movie 2). Similarly, the QQN glycoform is more flexible and the first principal moment shows the glycan moving toward the enzymic active site (Movie 3). Comparison of all three of these principal components supports our conclusion that protein–glycan interactions drive their correlated motions.

## CONCLUSIONS

Structural studies of glycoproteins present significant challenges. Our study provides not only structural insight into isolated glycoproteins but also a template for future studies. In particular, we demonstrated how combining SAXS and molecular dynamics simulations can elucidate key structural features of glycoproteins. This combination enabled us to discover that the N-glycosylation of human RNase 1 on Asn34 leads to a relatively inflexible glycoform. Moreover, we were able to identify features of the local chemical environment that are causative for this rigidity. In contrast, glycosylation on Asn76 and Asn88 leads to proteins and glycans that are more flexible, and a glycan on Asn88 is especially (and explicable) detrimental to enzymatic catalysis.

*In vivo*, glycoproteins do not exist in isolation. Accordingly, future work should strive to employ analyses in physiological contexts. In particular, utilizing the synergy between experiment and simulation to better understand how glycosylation modifies the binding of RNA<sup>82</sup> or the ribonuclease inhibitor protein<sup>83,84</sup> would clarify further the physiological role of RNase 1 glycosylation.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.biochem.0c00191>.

Figures S1–S10 (PDF)

Movie 1 (MP4)

Movie 2 (MP4)

Movie 3 (MP4)

## Accession Codes

Human RNase 1, P07998.

## AUTHOR INFORMATION

### Corresponding Authors

**Bin Zhang** – Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States; [orcid.org/0000-0002-3685-7503](https://orcid.org/0000-0002-3685-7503); Phone: (617) 258-0848; Email: [binz@mit.edu](mailto:binz@mit.edu)

**Ronald T. Raines** – Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States; [orcid.org/0000-0001-7164-1719](https://orcid.org/0000-0001-7164-1719); Phone: (617) 253-1470; Email: [rtraines@mit.edu](mailto:rtraines@mit.edu)

### Authors

**Henry R. Kilgore** – Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States; [orcid.org/0000-0003-4851-9656](https://orcid.org/0000-0003-4851-9656)

**Andrew P. Latham** – Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

**Valerie T. Ressler** – Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.biochem.0c00191>

### Author Contributions

<sup>†</sup>H.R.K. and A.P.L. contributed equally to this work.

### Funding

A.P.L. was supported by a National Science Foundation Graduate Research Fellowship. This work was supported by National Institutes of Health Grants R35 GM133580 and R01 CA073808.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors thank Dr. Charles Settens (Massachusetts Institute of Technology) for helpful discussions. SAXS data were collected at the Advanced Light Source (ALS), a national user facility operated by the Lawrence Berkeley National Laboratory on behalf of the U.S. Department of Energy (DOE), Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by the DOE Office of Biological and Environmental Research. Additional support comes from Grants P30 GM124169 and S10 OD018483 (National Institutes of Health).

## DEDICATION

This paper is dedicated to Professor Laura L. Kiessling on the occasion of her 60th birthday.

## ABBREVIATIONS

NQQ, N76Q/N88Q human RNase 1 with N-glycosylation on Asn34; QNQ, N34Q/N88Q human RNase 1 with N-glycosylation on Asn76; QQN, N34Q/N76Q human RNase 1 with N-glycosylation on Asn88.

## REFERENCES

- (1) Khoury, G. A., Baliban, R. C., and Floudas, C. A. (2011) Proteome-wide post-translational modification statistics: Frequency analysis and curation of the Swiss-Prot database. *Sci. Rep.* 1, 90.
- (2) Varki, A. (1993) Biological roles of oligosaccharides: All of the theories are correct. *Glycobiology* 3, 97–103.
- (3) Rudd, P. M., Joao, H. C., Coghill, E., Fiten, P., Saunders, M. R., Opdenakker, G., and Dwek, R. A. (1994) Glycoforms modify the dynamic stability and functional activity of an enzyme. *Biochemistry* 33, 17–22.
- (4) Mitra, N., Sinha, S., Ramya, T. N., and Surolia, A. (2006) N-Linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem. Sci.* 31, 156–163.
- (5) Solá, R. J., and Griebenow, K. (2009) Effects of glycosylation on the stability of protein pharmaceuticals. *J. Pharm. Sci.* 98, 1223–1245.
- (6) Russell, D., Oldham, N. J., and Davis, B. G. (2009) Site-selective chemical protein glycosylation protects from autolysis and proteolytic degradation. *Carbohydr. Res.* 344, 1508–1514.
- (7) Varki, A. (2017) Biological roles of glycans. *Glycobiology* 27, 3–49.
- (8) Amazaki, Y., Nguyen, H. M., Okamoto, R., Maki, Y., and Kajihara, Y. (2018) Effects of N-glycans on glycoprotein folding and protein dynamics. *Adv. Exp. Med. Biol.* 1104, 1–19.
- (9) Bertozzi, C. R., and Kiessling, L. L. (2001) Chemical glycobiology. *Science* 291, 2357–2364.
- (10) Kiessling, L. L., and Splain, R. A. (2010) Chemical approaches to glycobiology. *Annu. Rev. Biochem.* 79, 619–653.
- (11) Unverzagt, C., and Kajihara, Y. (2013) Chemical assembly of N-glycoproteins: A refined toolbox to address a ubiquitous posttranslational modification. *Chem. Soc. Rev.* 42, 4408–4420.
- (12) Kiessling, L. L., and Grim, J. C. (2013) Glycopolymer probes of signal transduction. *Chem. Soc. Rev.* 42, 4476–4491.
- (13) Sorrentino, S. (2010) The eight human “canonical” ribonucleases: Molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. *FEBS Lett.* 584, 2194–2200.
- (14) Koczera, P., Martin, L., Marx, G., and Schuerholz, T. (2016) The ribonuclease A superfamily in humans: Canonical RNases as the buttress of innate immunity. *Int. J. Mol. Sci.* 17, E1278.
- (15) Lu, L., Li, J., Moussaoui, M., and Boix, E. (2018) Immune modulation by human secreted RNases at the extracellular space. *Front. Immunol.* 9, 1012.
- (16) Garnett, E. R., Lomax, J. E., Mohammed, B. M., Gailani, D., Sheehan, J. P., and Raines, R. T. (2019) Phenotype of ribonuclease 1 deficiency in mice. *RNA* 25, 921–934.
- (17) Kunitz, M. (1939) Isolation from beef pancreas of a crystalline protein possessing ribonuclease activity. *Science* 90, 112–113.
- (18) Barnard, E. A. (1969) Ribonucleases. *Annu. Rev. Biochem.* 38, 677–732.
- (19) Richards, F. M., and Wyckoff, H. W. (1971) Bovine pancreatic ribonuclease. *Enzymes* 4, 647–806.
- (20) Karpeisky, M. Y., and Yakovlev, G. I. (1981) Topochemical principles of the substrate specificity of nucleases. *Sov. Sci. Rev., Sect. D* 2, 145–257.
- (21) Beintema, J. J., and Lenstra, J. A. (1982) Evolution of mammalian pancreatic ribonucleases. In *Macromolecular Sequences in Systematic and Evolutionary Biology* (Goodman, M., Ed.) pp 43–73, Springer, New York.
- (22) Blackburn, P., and Moore, S. (1982) Pancreatic ribonuclease. *Enzymes* 15, 317–433.
- (23) Eftink, M. R., and Biltonen, R. L. (1987) Pancreatic ribonuclease A: The most studied endoribonuclease. In *Hydrolytic Enzymes*, pp 333–375, Elsevier, New York.
- (24) D'Alessio, G., and Riordan, J. F., Eds. (1997) *Ribonucleases: Structures and Functions*, Academic Press, New York.
- (25) Raines, R. T. (1998) Ribonuclease A. *Chem. Rev.* 98, 1045–1066.
- (26) Kresge, N., Simoni, R. D., and Hill, R. L. (2005) The elucidation of the structure of ribonuclease by Stanford Moore and William H. Stein. *J. Biol. Chem.* 280, e47–e48.



- (27) Kresge, N.; Simoni, R. D.; and Hill, R. L. (2006) The thermodynamic hypothesis of protein folding: The work of Christian Anfinsen. *J. Biol. Chem.* 281, e11–e13.
- (28) Kresge, N.; Simoni, R. D.; and Hill, R. L. (2006) The solid phase synthesis of ribonuclease A by Robert Bruce Merrifield. *J. Biol. Chem.* 281, e21–e23.
- (29) Marshall, G. R.; Feng, J. A.; and Kuster, D. J. (2008) Back to the future: Ribonuclease A. *Biopolymers* 90, 259–277.
- (30) Cuchillo, C. M.; Nogués, N. V.; and Raines, R. T. (2011) Bovine pancreatic ribonuclease: Fifty years of the first enzymatic reaction mechanism. *Biochemistry* 50, 7835–7841.
- (31) Beintema, J. J.; Gaastra, W.; Scheffer, A. J.; and Welling, G. W. (1976) Carbohydrate in pancreatic ribonucleases. *Eur. J. Biochem.* 63, 441–448.
- (32) Joao, H. C.; Scragg, I. G.; and Dwek, R. A. (1992) Effects of glycosylation on protein conformation and amide proton exchange in RNase B. *FEBS Lett.* 307, 343–346.
- (33) Joao, H. C.; and Dwek, R. A. (1993) Effects of glycosylation on protein structure and dynamics in ribonuclease B and some of its individual glycoforms. *Eur. J. Biochem.* 218, 239–244.
- (34) Arnold, U.; and Ulbrich-Hofmann, R. (1997) Kinetic and thermodynamic thermal stabilities of ribonuclease A and ribonuclease B. *Biochemistry* 36, 2166–2172.
- (35) Beintema, J. J.; Blank, A.; Schieven, G. L.; Dekker, C. A.; Sorrentino, S.; and Libonati, M. (1988) Differences in glycosylation pattern of human secretory ribonucleases. *Biochem. J.* 255, 501–505.
- (36) Peracaula, R.; Royle, L.; Tabarés, G.; Mallorqui-Fernandez, G.; Barrabés, S.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M.; and de Llorens, R. (2003) Glycosylation of human pancreatic ribonuclease: Differences between normal and tumor states. *Glycobiology* 13, 227–244.
- (37) Petrescu, A.-J.; Milac, A.-L.; Petrescu, S. M.; Dwek, R. A.; and Wormald, M. R. (2003) Statistical analysis of the protein environment of N-glycosylation sites: Implications for occupancy, structure, and folding. *Glycobiology* 14, 103–114.
- (38) Ribó, M.; Beintema, J. J.; Ossset, M.; Fernández, E.; Bravo, J.; de Llorens, R.; and Cuchillo, C. M. (1994) Heterogeneity in the glycosylation pattern of human pancreatic ribonuclease. *Biol. Chem. Hoppe Seyler* 375, 357–363.
- (39) Ressler, V. T.; and Raines, R. T. (2019) Consequences of the endogenous N-glycosylation of human ribonuclease I. *Biochemistry* 58, 987–996.
- (40) Imperiali, B.; and Hendrickson, T. L. (1995) Asparagine-linked glycosylation: Specificity and function of oligosaccharyl transferase. *Bioorg. Med. Chem.* 3, 1565–1578.
- (41) Yang, Y.; Liu, F.; Franc, V.; Halim, L. A.; Schellekens, H.; and Heck, A. J. R. (2016) Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity. *Nat. Commun.* 7, 13397.
- (42) Wu, D.; Struwe, W. B.; Harvey, D. J.; Ferguson, M. A. J.; and Robinson, C. V. (2018) N-Glycan microheterogeneity regulates interactions of plasma proteins. *Proc. Natl. Acad. Sci. U. S. A.* 115, 8763–8768.
- (43) Fisher, P.; Thomas-Oates, J.; Wood, A. J.; and Ungar, D. (2019) The N-glycosylation processing potential of the mammalian Golgi apparatus. *Front. Cell Dev. Biol.* 7, 157.
- (44) Koch, M. H. J.; Vachette, P.; and Svergun, D. I. (2003) Small-angle scattering: A view on the properties, structures, and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* 36, 147–227.
- (45) Putnam, C. D.; Hammel, M.; Hura, G. L.; and Tainer, J. A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations, and assemblies in solution. *Q. Rev. Biophys.* 40, 191–285.
- (46) Zheng, W.; and Best, R. B. (2018) An extended Guinier analysis for intrinsically disordered proteins. *J. Mol. Biol.* 430, 2540–2553.
- (47) Brosey, C. A.; and Tainer, J. A. (2019) Evolving SAXS versatility: Solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology. *Curr. Opin. Struct. Biol.* 58, 197–213.
- (48) Thomas, S. P.; Kim, E.; Kim, J.-S.; and Raines, R. T. (2016) Knockout of the ribonuclease inhibitor gene leaves human cells vulnerable to secretory ribonucleases. *Biochemistry* 55, 6359–6362.
- (49) Koppel, D. E. (1972) Revisiting the method of cumulants for the analysis of dynamic light-scattering data. *J. Chem. Phys.* 57, 4814–4820.
- (50) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; and Svergun, D. I. (2003) PRIMUS—a Windows-PC based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* 36, 1277–1282.
- (51) Johnson, R. J.; McCoy, J. G.; Bingman, C. A.; Phillips, G. N., Jr.; and Raines, R. T. (2007) Inhibition of human pancreatic ribonuclease by the human ribonuclease inhibitor protein. *J. Mol. Biol.* 368, 434–449.
- (52) Fiser, A.; Do, R. K.; and Sali, A. (2000) Modeling of loops in protein structures. *Protein Sci.* 9, 1753–1773.
- (53) Jo, S.; Kim, T.; Iyer, V. G.; and Im, W. (2008) CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29, 1859–1865.
- (54) Park, S. J.; Lee, J.; Qi, Y.; Kern, N. R.; Lee, H. S.; Jo, S.; Joung, I.; Joo, K.; Lee, J.; and Im, W. (2019) CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology* 29, 320–331.
- (55) Berendsen, H. J. C.; van der Spoel, D.; and van Drunen, R. (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91, 43–56.
- (56) Abraham, M. J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J. C.; Hess, B.; and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2, 19–25.
- (57) Svergun, D. I.; Barberato, C.; and Koch, M. H. J. (1995) CRYSOLE—A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* 28, 768–773.
- (58) Ortega, A.; Amorós, D.; and García de la Torre, J. (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* 101, 892–898.
- (59) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; and Beckstein, O. (2011) MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* 32, 2319–2327.
- (60) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domanski, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.; and Beckstein, O. (2016) MDAnalysis: A Python package for the rapid analysis of molecular dynamics simulations. In *Proceedings of the 15th Python in Science Conference* (Benthall, S., and Rostrup, S., Eds.) pp 98–105, SCIPY, Austin, TX.
- (61) Bakan, A.; Dutta, A.; Mao, W.; Liu, Y.; Chennubhotla, C.; Lezon, T. R.; and Bahar, I. (2014) Evol and ProDy for bridging protein sequence evolution and structural dynamics. *Bioinformatics* 30, 2681–2683.
- (62) Yang, S.; Park, S.; Makowski, L.; and Roux, B. (2009) A rapid coarse residue-based computational method for X-ray solution scattering characterization of protein folds and multiple conformational states of large protein complexes. *Biophys. J.* 96, 4449–4463.
- (63) Rózycki, B.; Kim, Y. C.; and Hummer, G. (2011) SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. *Structure* 19, 109–116.
- (64) Latham, A. P.; and Zhang, B. (2019) Improving coarse-grained protein force fields with small-angle X-ray scattering data. *J. Phys. Chem. B* 123, 1026–1034.
- (65) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; and Woods, R. J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.* 26, 1668–1688.



- (66) Robustelli, P., Piana, S., and Shaw, D. E. (2018) Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. U. S. A.* 115, E4758–E4766.
- (67) Wu, H., Wolynes, P. G., and Papoian, G. A. (2018) AWSEM-IDP: A coarse-grained force field for intrinsically disordered proteins. *J. Phys. Chem. B* 122, 11115–11125.
- (68) Latham, A. P., and Zhang, B. (2020) Maximum entropy optimized force field for intrinsically disordered proteins. *J. Chem. Theory Comput.* 16, 773–781.
- (69) Jo, S., Song, K. C., Desaire, H., MacKerell, A. D., Jr, and Im, W. (2011) Glycan Reader: Automated sugar identification and simulation preparation for carbohydrates and glycoproteins. *J. Comput. Chem.* 32, 3135–3141.
- (70) Lee, S., Tran, A., Allsopp, M., Lim, J. B., Henin, J., and Klauda, J. B. (2014) CHARMM36 united atom chain model for lipids and surfactants. *J. Phys. Chem. B* 118, 547–556.
- (71) Park, S. J., Lee, J., Patel, D. S., Ma, H., Lee, H. S., Jo, S., and Im, W. (2017) Glycan Reader is improved to recognize most sugar types and chemical modifications in the Protein Data Bank. *Bioinformatics* 33, 3051–3057.
- (72) Rambo, R. P., and Tainer, J. A. (2011) Characterizing flexible and intrinsically unstructured macromolecules by SAS using the Porod-Debye law. *Biopolymers* 95, 559–571.
- (73) Hudson, K. L., Bartlett, G. J., Diehl, R. C., Agirre, J., Gallagher, T., Kiessling, L. L., and Woolfson, D. N. (2015) Carbohydrate–aromatic interactions in proteins. *J. Am. Chem. Soc.* 137, 15152–15160.
- (74) Hsu, C.-H., Park, S., Mortenson, D. E., Foley, B. L., Wang, X., Woods, R. J., Case, D. A., Powers, E. T., Wong, C.-H., Dyson, H. J., and Kelly, J. W. (2016) The dependence of carbohydrate–aromatic interaction strengths on the structure of the carbohydrate. *J. Am. Chem. Soc.* 138, 7636–7648.
- (75) Presta, L. G., and Rose, G. D. (1988) Helix signals in proteins. *Science* 240, 1632–1641.
- (76) Pauling, L., Corey, R. B., and Branson, H. R. (1951) The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.* 37, 205–211.
- (77) Newberry, R. W., and Raines, R. T. (2017) The  $n \rightarrow \pi^*$  interaction. *Acc. Chem. Res.* 50, 1838–1846.
- (78) Richardson, J. S., and Richardson, D. C. (1988) Amino acid preferences for specific locations at the ends of  $\alpha$  helices. *Science* 240, 1648–1652.
- (79) Doig, A. J., and Baldwin, R. L. (1995) N- and C-capping preferences for all 20 amino acids in  $\alpha$ -helical peptides. *Protein Sci.* 4, 1325–1336.
- (80) Chen, M. M., Bartlett, A. I., Nerenberg, P. S., Friel, C. T., Hackenberger, C. P. R., Stultz, C. M., Radford, S. E., and Imperiali, B. (2010) Perturbing the folding energy landscape of the bacterial immunity protein Im7 by site-specific N-linked glycosylation. *Proc. Natl. Acad. Sci. U. S. A.* 107, 22528–22533.
- (81) Elbaum, M. B., and Zondlo, N. J. (2014) OGlcNAcylation and phosphorylation have similar structure effects in  $\alpha$ -helices: Post-translational modifications as inducible start and stop signals in  $\alpha$ -helices, with greater structural effects on threonine modification. *Biochemistry* 53, 2242–2260.
- (82) Nogué, M. V., Moussaoui, M., Boix, E., Vilanova, M., Ribó, M., and Cuchillo, C. M. (1998) The contribution of noncatalytic phosphate-binding subsites to the mechanism of bovine pancreatic ribonuclease A. *Cell. Mol. Life Sci.* 54, 766–774.
- (83) Dickson, K. A., Haigis, M. C., and Raines, R. T. (2005) Ribonuclease inhibitor: Structure and function. *Prog. Nucleic Acid Res. Mol. Biol.* 80, 349–374.
- (84) Lomax, J. E., Eller, C. H., and Raines, R. T. (2012) Rational design and evaluation of mammalian ribonuclease cytotoxins. *Methods Enzymol.* 502, 273–290.