Structural characterization of the membrane-associating FATC domain of ataxia telangiectasia mutated by NMR and MD simulations

Munirah S. Abd Rahim[1], Yevhen K. Cherniavskyi[2], D. Peter Tieleman[2], Sonja A. Dames[1,3$]

From the [1]Chair of Biomolecular NMR Spectroscopy, Department of Chemistry, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany, [2]Department of Biological Sciences and Centre for Molecular Simulation, University of Calgary, Calgary, AB, Canada, [3]Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

[$]To whom correspondence may be addressed: Chair of Biomolecular NMR Spectroscopy, Department of Chemistry, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany, Tel.: +49-89-35831-7103, Fax: +49-89-3229-4002, E-mail: (sonja.dames@tum.de)

Running title: structure of membrane-associating ATM FATC

Supplementary results

Structure of micelle-associated hATMfatc

In 11 of the 20 lowest energy structures the C-terminal region forms an α-helix that includes residues G95/3051-W99/3055, in 7 of the 20 a 3^{10} helix is formed by residues P94/3050-A98/3054, and in 2 of the 20 only a turn-like structure is formed.

Backbone dynamics of micelle-associated hATMfatc

The association of the FATC domain of human ATM fused to GB1 (hATMfatc-gb1ent) with the significantly larger DPC micelles should modulate the dynamic properties. Fig. 1E and SI Fig. S2A shows the $\{^1H\}$-^{15}N-NOE data for the hATMfatc part (residues 68-100) in the presence of 150 mM and 200 mM DPC at 298 K. The respective spectra are shown in SI Fig. S1. Consistent with the transition from a rather unstructured, dynamic state in the free from to a more ordered one associated with micelles (1) the $\{^1H\}$-^{15}N-NOE values at both DPC concentrations are generally in the range typical for a mostly folded protein (≈ 0.6-0.8). At 150 mM DPC, the conditions used to derive structural restraints, the average value for residues L70/3026 to W99/3055 is 0.53 ± 0.02. The values for the helical regions are slightly higher: 0.60 ± 0.01 for G73/3029-A83/3039, 0.58 ± 0.01 for K87/3043-R91/3047, and 0.56 ± 0.02 for G95/3051-W99/3055. For comparison, the average value for the well-structured GB1 tag (residues 2-56) is 0.64 ± 0.01 (SI Fig. S2B). Since at higher DPC concentrations the

binding equilibrium is expected to be shifted more towards the bound state, we also recorded ^{15}N-relaxation and structural data at 200 mM DPC. However, the spectral quality was overall lower due to stronger line broadening, which in SI Fig. S1 is illustrated by spectra superpositions for the {^{1}H}-^{15}N-NOE data for samples with 150 and 200 mM DPC. Thus, the micelle-associated hATMfatc structure was calculated using the structural data from the 150 mM DPC samples recorded at 298 K. The {^{1}H}-^{15}N-NOE values at 200 mM at 298 K are very similar to that at 150 mM (Fig. 1E and SI Fig. S2A). The average ^{15}N-T_1 and -T_2 values of the GB1 tag (SI Fig. S2B, 517.6 ± 0.4 ms and 60.1 ± 0.1 ms at 200 mM DPC at 298 K) that does not interact with DPC micelles are in a range typical for an about 6 kDa protein flexibly linked to another object and similar to those observed for the GB1tag fused to the micelle-associated FATC domain of the kinase DNA-PKcs (1,2). The association of the hATMFatc part (~4 kDa) with the ~20 kDa DPC micelle results in average values for the residue range L70/3026 to W99/3055 of 767.6 ± 8.6 ms for ^{15}N-T_1 and 26.5 ± 0.8 ms for ^{15}N-T_2 (SI Fig. S2A, B) at 200 mM DPC and 298 K. The ^{15}N-T_2 values are lower than those of the micelle-associated DNA-PKcs FATC domain (average value ~40 ms (1)). This can be explained by an exchange contribution from the relative movement of the helices with respect to each other due to the dynamic linkages (Fig. 1C). The low {^{1}H}-^{15}N-NOE values for residues in these linkers also indicate these regions are dynamic. D85/3041 in the linker between α1 and α2 shows a value of 0.32 ± 0.10 and L92 in the linker between α2 and α3 of 0.42 ± 0.08 (all at 150 mM DPC, 298 K, Fig. 1E, SI Fig. S2).

Dynamic information from the MD simulations of micelle-associated hATMfatc

The top left panel of SI Fig. S4A shows the RMSF values for run 1 calculated by fitting to the starting structure. These global RMSF values for run1 are overall high and vary significantly as a function of the residue sequence position. The latter is however not observed for runs 2 and 3 (SI Fig. S4A, middle and bottom left panels). If the contribution of the relative movement of the helices with respect to each other is reduced by fitting only to the residue range 72 to 92 or 86 to 98 and to multiple structures that were obtained as average from 5 ns intervals, local RMSF values are obtained, which are overall much smaller for all runs (SI Fig. S4A middle and right column). That the relative orientation of the helices with respect to each other varies in all runs, mainly due to the dynamic linkage between the helices 1 and 2, is also illustrated by a superposition of representative structures of the top 10 clusters, which correspond to 69.3 % of the total population, from a clustering analysis based

on the relative RMSD (SI Fig. S4B). Superimposing α1, the other two helices do not superimpose as well. The backbone and side chain RMSD values as a function of the sequence are plotted in SI Fig. S4C. Due to the mentioned dynamic properties, these are also overall high.

Supplementary figure legends

Fig. S1: Analysis of changes in the backbone dynamics of hATMfatc-gb1ent upon interaction with DPC micelles ((A) 150 mM, (B) 200 mM) based on $\{^1H\}$-^{15}N-NOE data. Positive peaks are colored black and red, negative peaks in blue and yellow. Each plot shows a superposition of the spectrum without (reference spectrum) and with NOE-effect (NOE spectrum). The plots to the right are the same as the ones to left but show additionally the spectrum of the GB1 tag followed by a thrombin and factor Xa site (= GB1xa) in green on top to facilitate the identification of the peaks corresponding to the FATC part (no green peaks on top). The unassigned data for the free form and a partially assigned version of the one at 150 mM have been published (1).

Fig. S2: Backbone dynamics of hATMfatc-gb1ent based on ^{15}N-relaxation data. For 200 mM DPC ^{15}N-T_1 (top panel), ^{15}N-T_2 (middle panel) and $\{^1H\}$-^{15}N NOE values (bottom panel, also shown in Fig. 1E, here reproduced for better comparison with ^{15}N-T_1 and $-T_2$) have been plotted. For 150 mM DPC only $\{^1H\}$-^{15}N NOE values (bottom panel) have been plotted. (A) show the data for the hATMfatc part alone. The secondary structure content and the sequence are displayed at the top. In addition sequence stretches adopting a helical structure have been shaded grey. (B) show the data for the full 100-residue long hATMfatc-gb1ent fusion protein. Indicated by the negative $\{^1H\}$-^{15}N NOE values and high ^{15}N-T_2 values, the linker region between GB1 and hATMfatc shows strongly increased backbone dynamics and thus flexibly links the FATC to the GB1 part.

Fig. S3: Secondary structure content of micelle-immersed hATMfatc as a function of the simulation time for the three independent MD runs starting each from one of the three lowest energy NMR structures. The color coding is given to the lower right.

Fig. S4: (A) Plots of the root mean square fluctuations (RMSF) as a function of the residue sequence positions. The left plot shows the RMSF fluctuations calculated by fitting the whole hATMfatc domain to the respective starting structure. For the right two plots the contribution from movements of the helices with respect to each other has been reduced by calculating the RMSF for two sections of the protein (residues 72-92 or 86-98) and using multiple structures for the progressive fit that were obtained by splitting the last 500 ns of the simulation into 5 ns intervals, calculating average RMSF values for each interval with the reference structure taken as the structure at the beginning of each interval. (B) Ribbon representation in the same color coding as in Fig. 3 showing a superposition of 10 structures taken from the middle of each of the top 10 best clusters from a clustering analysis based on the relative RMSD of all three trajectories combined. (C) Plots of the backbone and side chain RMSD values as a function of the residue sequence position that were calculated by fitting the whole hATMfatc domain to the respective starting structure. (D) To estimate if there are interhelical contacts below about 0.5 nm that could give rise to NOEs, interresidue distances were calculated between the centers of mass for pairs of residues that belong to different helices (excluding adjacent residue pairs).

Fig. S5: (A) Information about the movement of the 3 helices with respect to each other from an analysis of the variation of the angle between helix 1 and 2 as well as the one between helix 2 and 3 for all three runs. The picture displaying the helix angle definition and the data for run 1 have also been shown in Fig. 3C and have been reproduced here for better comparison with runs 2 and 3. The analysis of the angles between the helices starts from 800 ns, and not 300 ns because the relative orientation of the helices changes slowly compared to the local structure of the protein. Thus we let the system equilibrate for a longer period of time to have less bias from the initial structure. (B) Superpositions of the ^{1}H-^{15}N HSQC spectra of hATMfatc-gb1xa in the presence of DPC micelles and 1 to 4 mM paramagnetically tagged 5- (left) or 16- (right) doxyl stearic acid (5-/16-SASL). The hATMfatc signals are labeled by the residue sequence position in hATMfatc-gb1ent/human ATM and the one letter amino acid code. The color coding is indicated at the top.

Fig. S6: Diagrams of the paramagnetic relaxation enhancement (PRE) effects due to the presence of 1 or 2 mM 5-SASL as a function of the residue sequence position for the full hATMfatc-gb1ent. To better compare the PRE effects to the average chemical shift changes in the second plot, 1-PRE (= 1- I(x mM SASL)/I (0 mM SASL)) was plotted. Accordingly, the larger the PRE effect, the higher the 1-PRE value. As expected the GB1 tag that does not directly interact with DPC micelles shows only minor PRE effects or chemical shift changes.

Fig. S7: Diagrams of the paramagnetic relaxation enhancement (PRE) effects due to the presence of 1 or 2 mM 16-SASL as a function of the residue sequence position for the full hATMfatc-gb1ent. To better compare the PRE effects to the average chemical shift changes in the second plot, 1-PRE (= 1- I(x mM SASL)/I (0 mM SASL)) was plotted. Accordingly, the larger the PRE effect, the higher the 1-PRE value. As expected the GB1 tag that does not directly interact with DPC micelles shows only minor PRE effects or chemical shift changes.

Fig. S8: Diagrams of the average distance of the center of mass (COM) of each residue of hATMfatc to that of the DPC micelle for all three MD runs. The averaging is performed over the last microsecond of the run. The dotted line corresponds to the average distance of the phosphor atom of the DPC head group to the micelle COM (1.914 nm). The data for run 1 have also been shown in Fig. 4C and have been reproduced here for better comparison with runs 2 and 3.

Fig. S9: Diagrams of the percentage of the solvent accessible surface area (SAS) of each residue of hATMfatc covered by DPC from the MD simulations for all three independent runs. Residues with shorter distances to the micelle COM in SI Fig. S8 show usually a higher percentage in the corresponding diagrams here. The data for run 1 have also been shown in Fig. 4D and have been reproduced here for better comparison with runs 2 and 3.

Fig. S10: Superposition of the ^{1}H-^{15}N HSQC spectra of the wild type, the mutant, and the GB1 tag followed by a thrombin and factor Xa site (= GB1xa) for each mutant. Peaks with a green peak on top belong to the GB1 tag. Since the GB1 tag does not interact with membrane mimetics its signals do not significantly shift (3). The peaks of the wild type spectrum are labeled by the 1-letter amino acid code and the sequence position based. Small letter indicate side chain signals. The left side shows the spectra superpositions in the absence of DPC, the right side in the presence of DPC.

Fig. S11: (A)-(B) Superpositions of the ^{1}H-^{15}N HSQC spectra of hATMfatc-gb1ent-F93/3049A-W96/3052 and F93/3049A in the presence of stepwise increasing concentrations of DPC. (B) The pictures show superposition of the ^{1}H-^{15}N HSQC spectra of the single mutant hATMfatc-gb1ent-F93/3049A in the absence and presence of DPC micelles (50 mM DPC, upper left), Dihep-PC micelles (50 mM DihepPC, upper right), DMPC/Dihep-PC bicelles (q = 0.2, [DMPC] = 0.04 M, and [Dihep-PC] = 0.20 M, cL 15%, lower left) or DMPC liposomes (< 50 mM DMPC, lower right). The spectrum of the free form is always shown in black and the one with the respective membrane mimetic in red. A green plus in the upper left of each plot indicates strong spectral changes and thus interactions and a red plus no significant changes and thus no significant interactions. To better identify the signals of the ATM FATC part, the spectrum of the GB1 tag followed by a thrombin and factor Xa site (= GB1xa) is additionally shown in green on top. Accordingly peaks with a green peak on top belong to the GB1 tag. Since the GB1 tag does not interact with membrane mimetics its signals do not significantly shift (3). In the picture of the spectra in the absence and presence of DPC, hATMfatc signals in the presence of DPC (red) are labeled by the 1-letter amino acid code and the sequence position in hATMfatc-gb1ent/human ATM based on similarity to the wild type spectrum. Small letters indicate side chain signals. A question mark behind the label indicates tentative assignments. Note that some red peaks that appear at new positions could not be assigned in this way. Note that free hATMfatc was not assigned since the signals of many backbone amide groups were not detectable.

Fig. S12: (A) Superpositions of the ^{1}H-^{15}N HSQC spectra of hATMfatc-gb1ent-W96/3052A in the presence of stepwise increasing concentrations of DPC. (B) The pictures show superposition of the ^{1}H-^{15}N HSQC spectra of the single mutant hATMfatc-gb1ent-W96/3052A in the absence and presence of DPC micelles (50 mM DPC, upper left), Dihep-PC micelles (50 mM DihepPC, upper right), DMPC/Dihep-PC bicelles (q = 0.2, [DMPC] = 0.04 M, and [Dihep-PC] = 0.20 M, cL 15%, lower left) or DMPC liposomes (< 50 mM DMPC, lower right). The spectrum of the free form is always shown in black and the one with the respective membrane mimetic in red. A green plus in the upper left of each plot indicates strong spectral changes and thus interactions and a red plus no significant changes and thus no significant interactions. To better identify the signals of the ATM FATC part, the spectrum of the GB1 tag followed by a thrombin and factor Xa site (= GB1xa) is additionally shown in green on top. Accordingly peaks with a green peak on top belong to the GB1 tag. Since the GB1 tag does not interact with membrane mimetics its signals do not significantly shift (3). In the picture of the spectra in the absence and presence of DPC, hATMfatc signals are labeled by the 1-letter amino acid code and the sequence position based on similarity to the wild type spectrum. Small letters indicate side chain signals. A question mark behind the label indicates tentative assignments. Note that some red peaks that appear at new positions could not be assigned in this way.

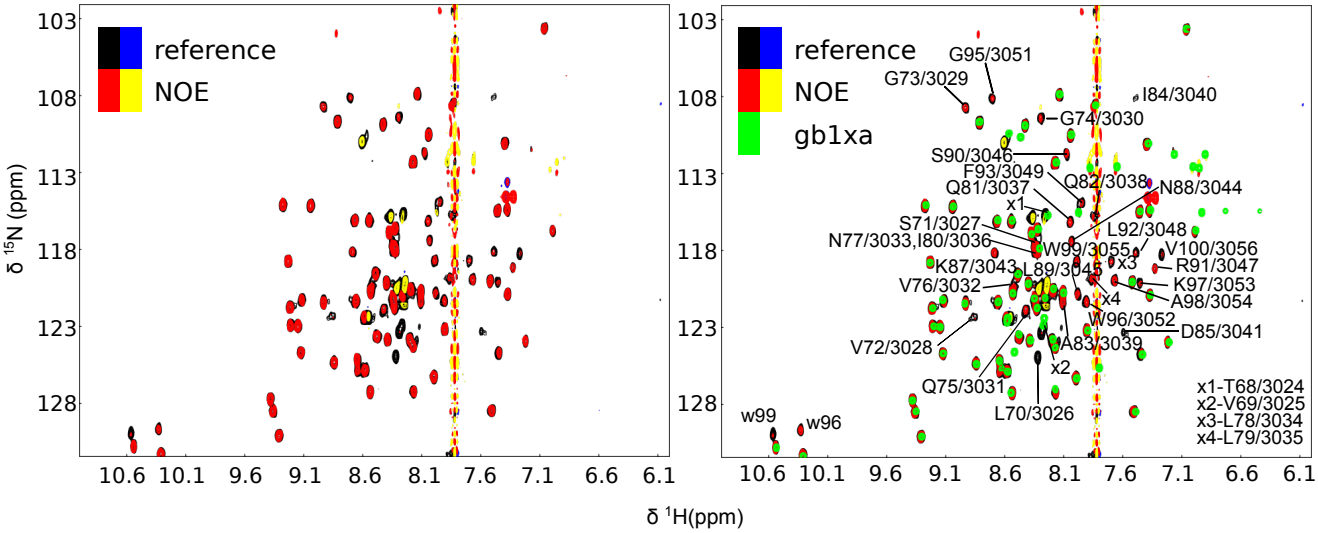
Supplementary References

1. Sommer, L. A. M., Schaad, M., and Dames, S. A. (2013) NMR- and Circular Dichroism-monitored Lipid Binding Studies Suggest a General Role for the FATC Domain as Membrane Anchor of Phosphatidylinositol 3-Kinase-related Kinases (PIKK). *Journal of Biological Chemistry* **288**, 20046-20063
2. Walsh, J. D., Meier, K., Ishima, R., and Gronenborn, A. M. (2010) NMR studies on domain diffusion and alignment in modular GB1 repeats. *Biophysical journal* **99**, 2636-2646

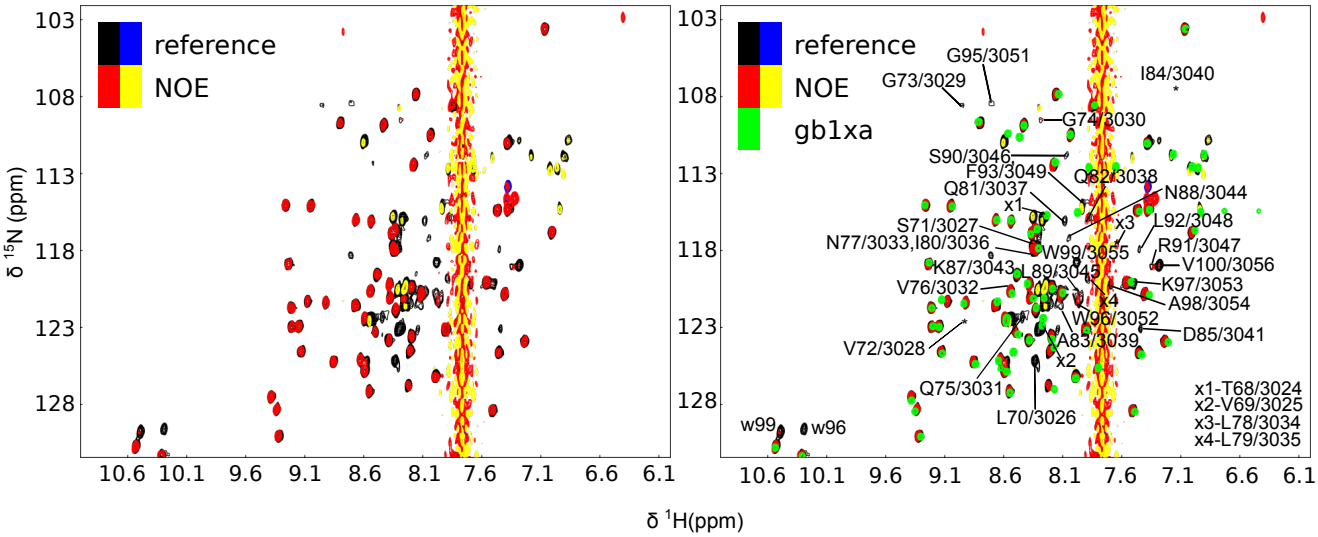
3. Sommer, L. A. M., Meier, M. A., and Dames, S. A. (2012) A fast and simple method for probing the interaction of peptides and proteins with lipids and membrane-mimetics using GB1 fusion proteins and NMR spectroscopy. *Protein science : a publication of the Protein Society* **21**, 1566-1570

Abd Rahim et al.
SI Fig. S1

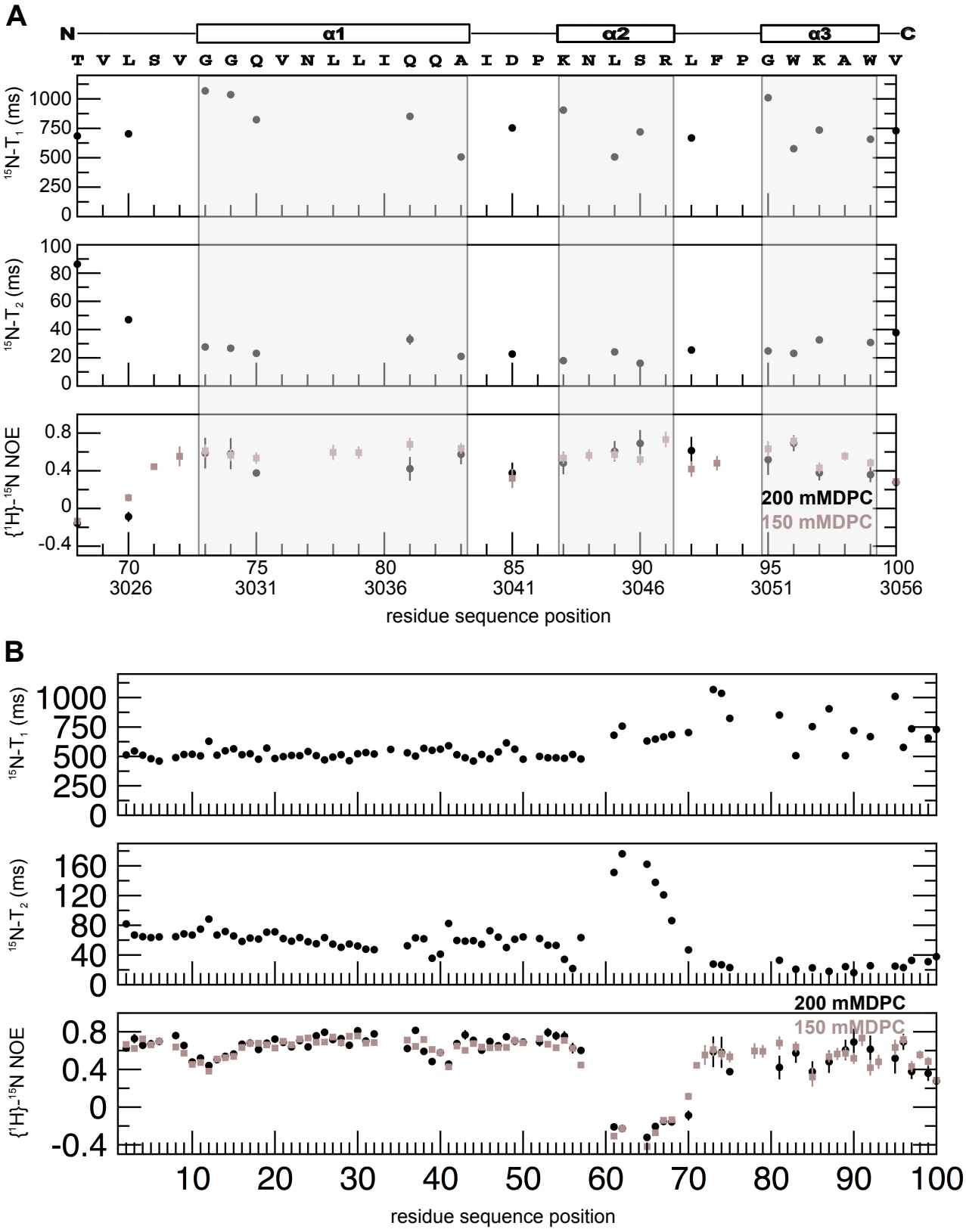
A hATMfatc-gb1ent with 150 mM DPC



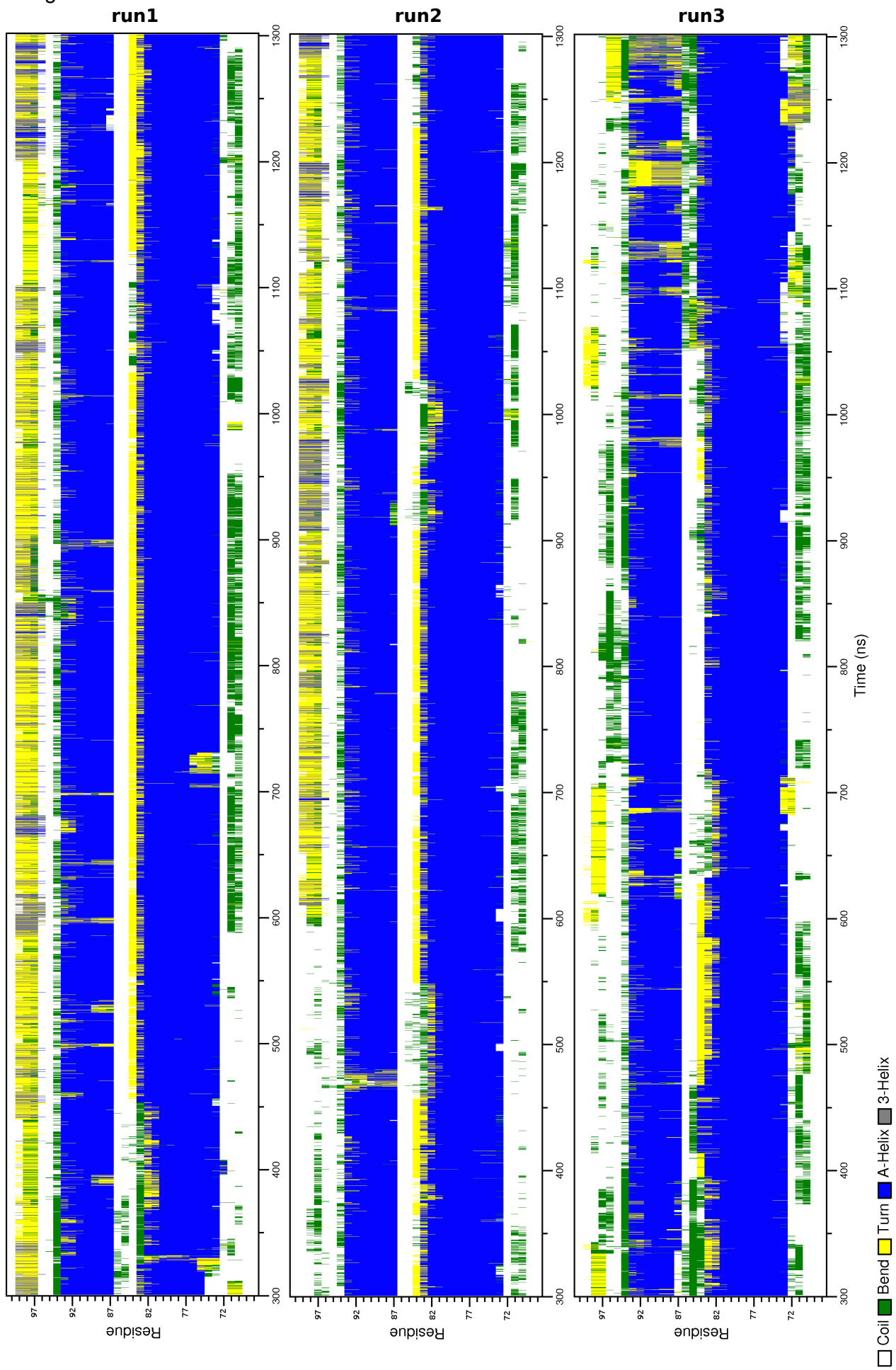
B hATMfatc-gb1ent with 200 mM DPC



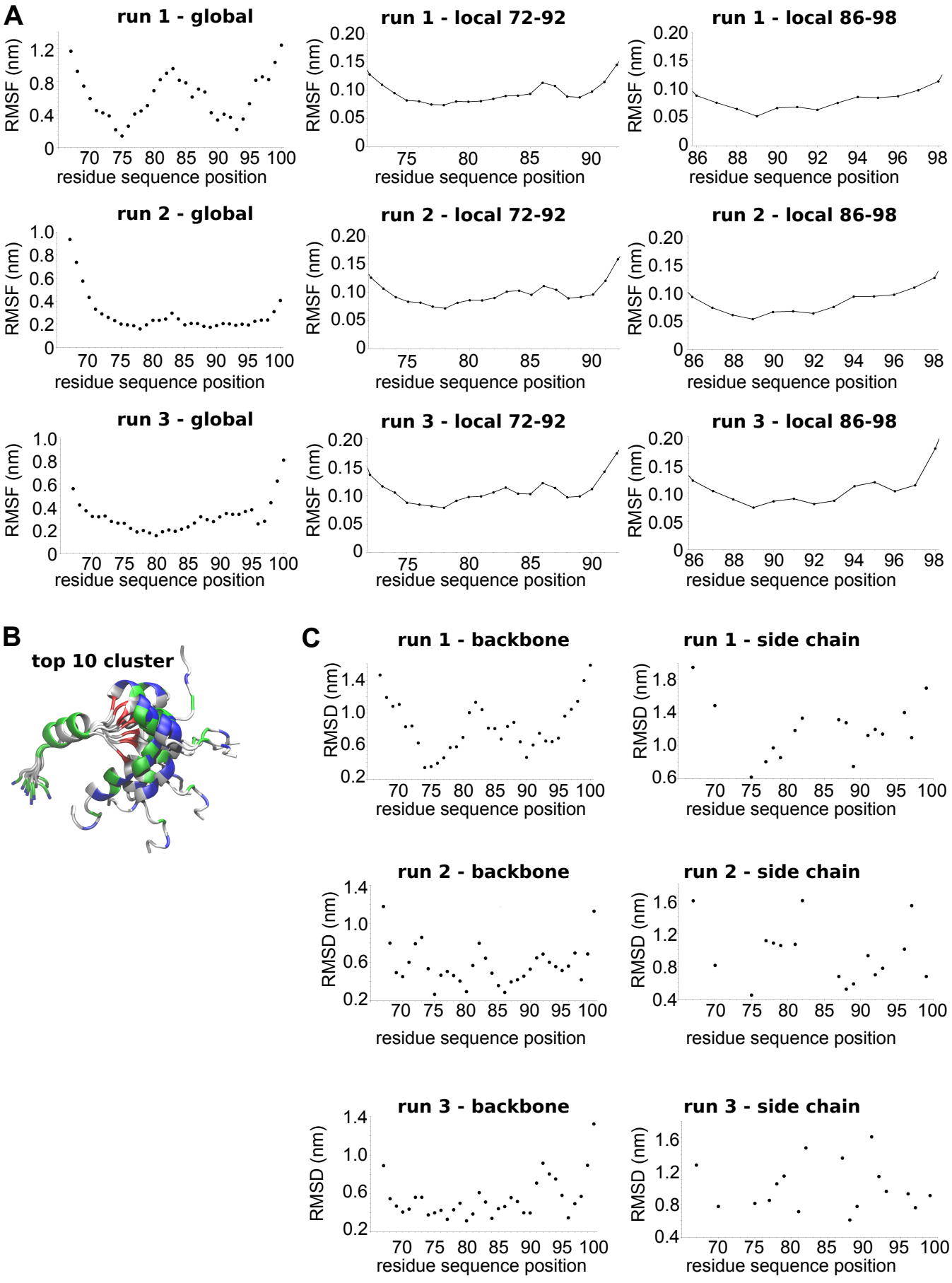
Abd Rahim et al.
SI Fig. S2



Abd Rahim et al.
SI Fig. S3

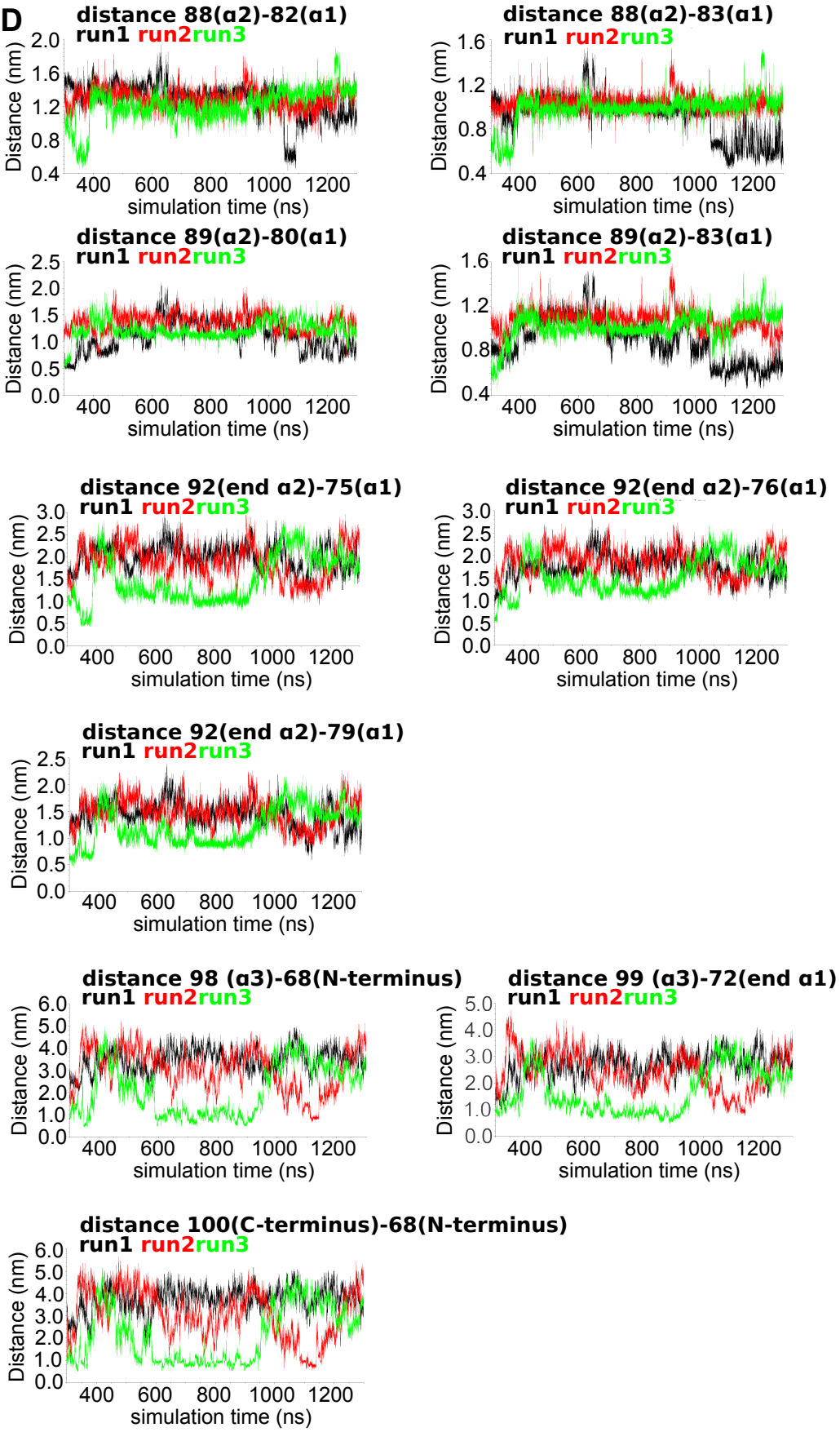


Abd Rahim et al.
SI Fig. S4

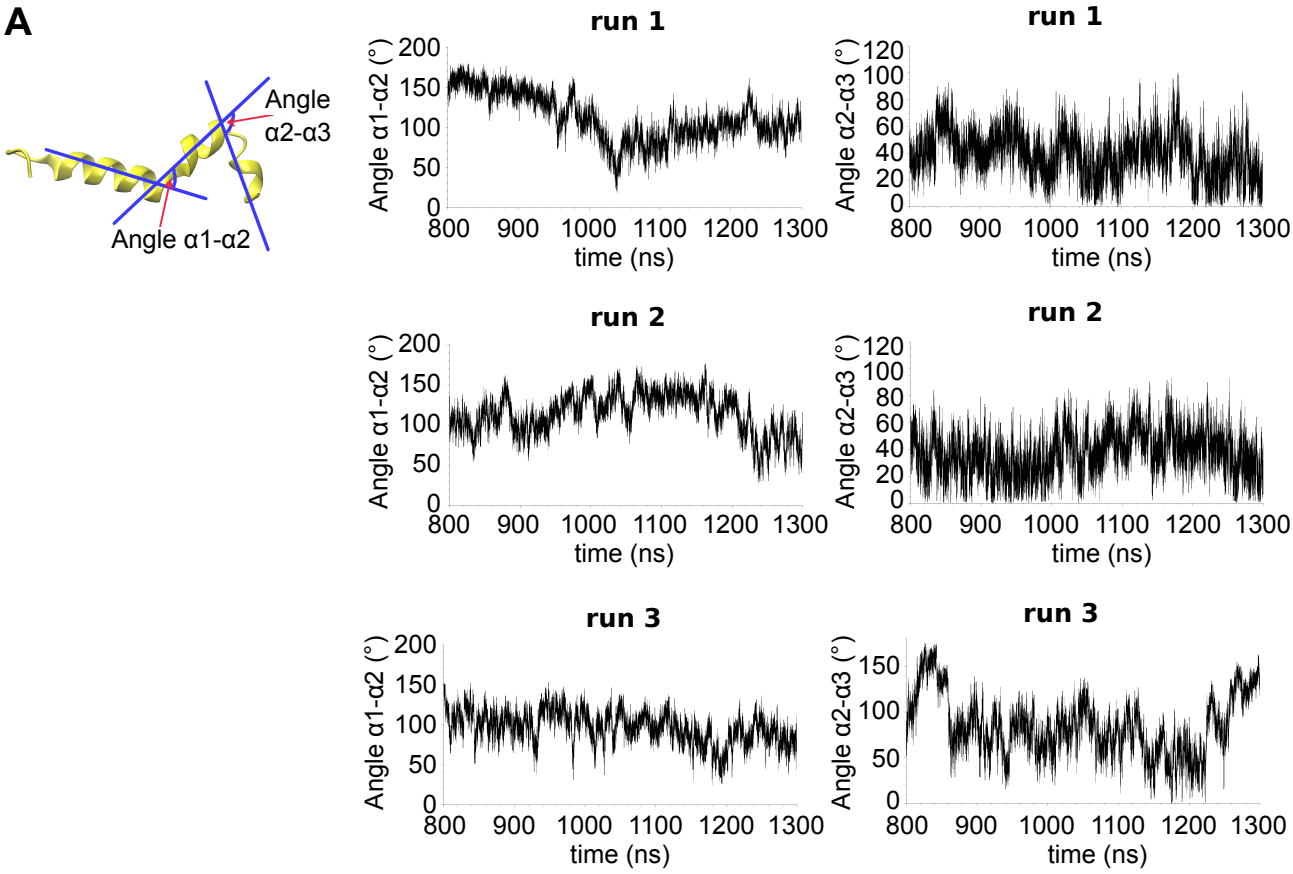


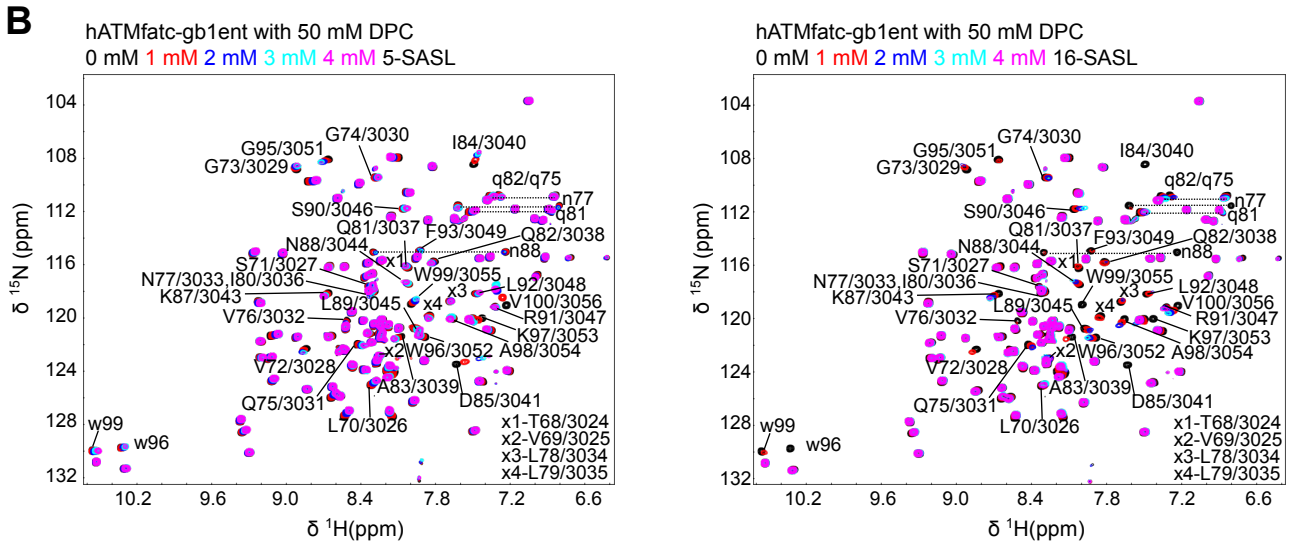
Abd Rahim et al.
SI Fig. S4-part 2

D

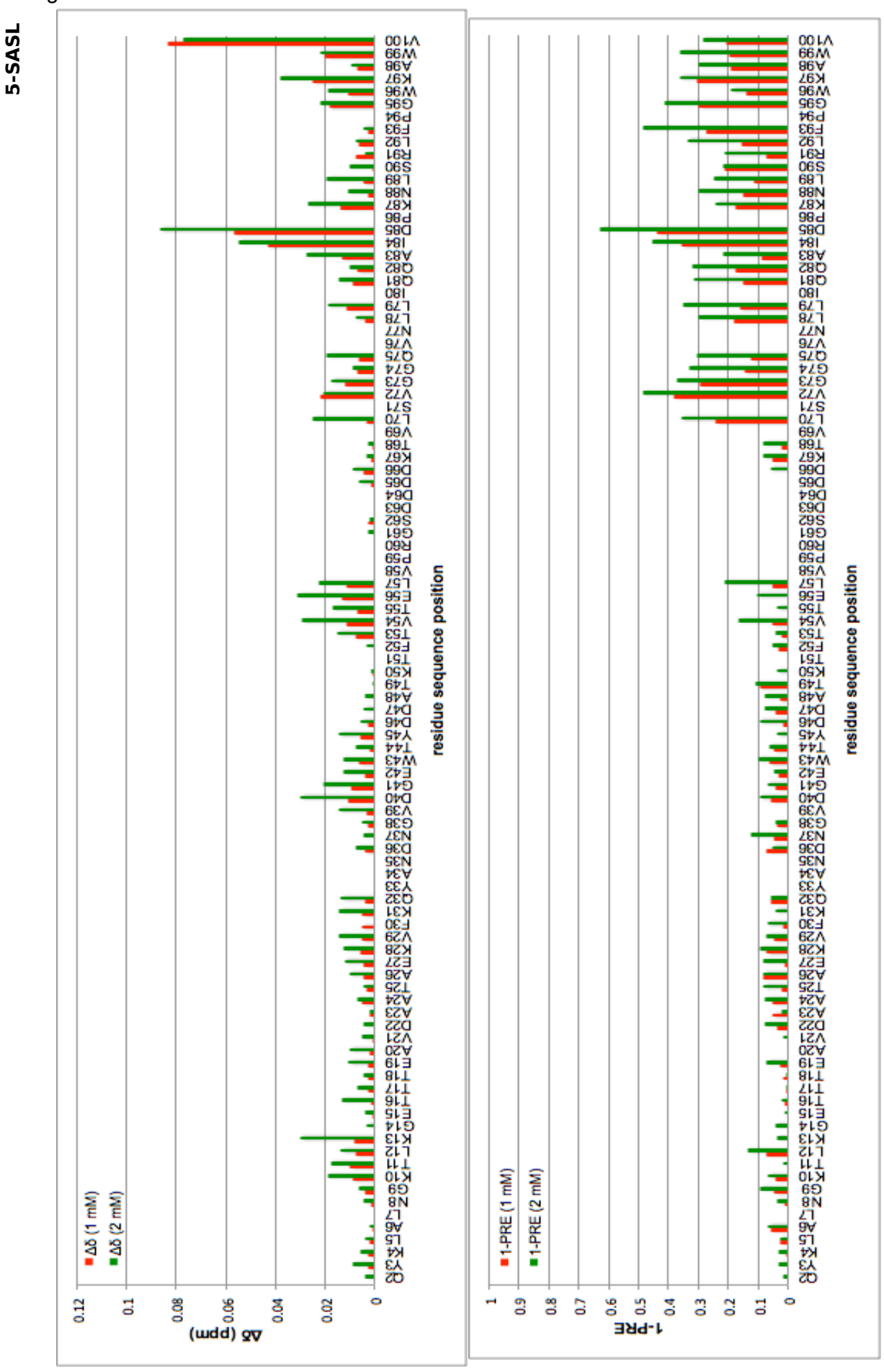


Abd Rahim et al.
SI Fig. S5

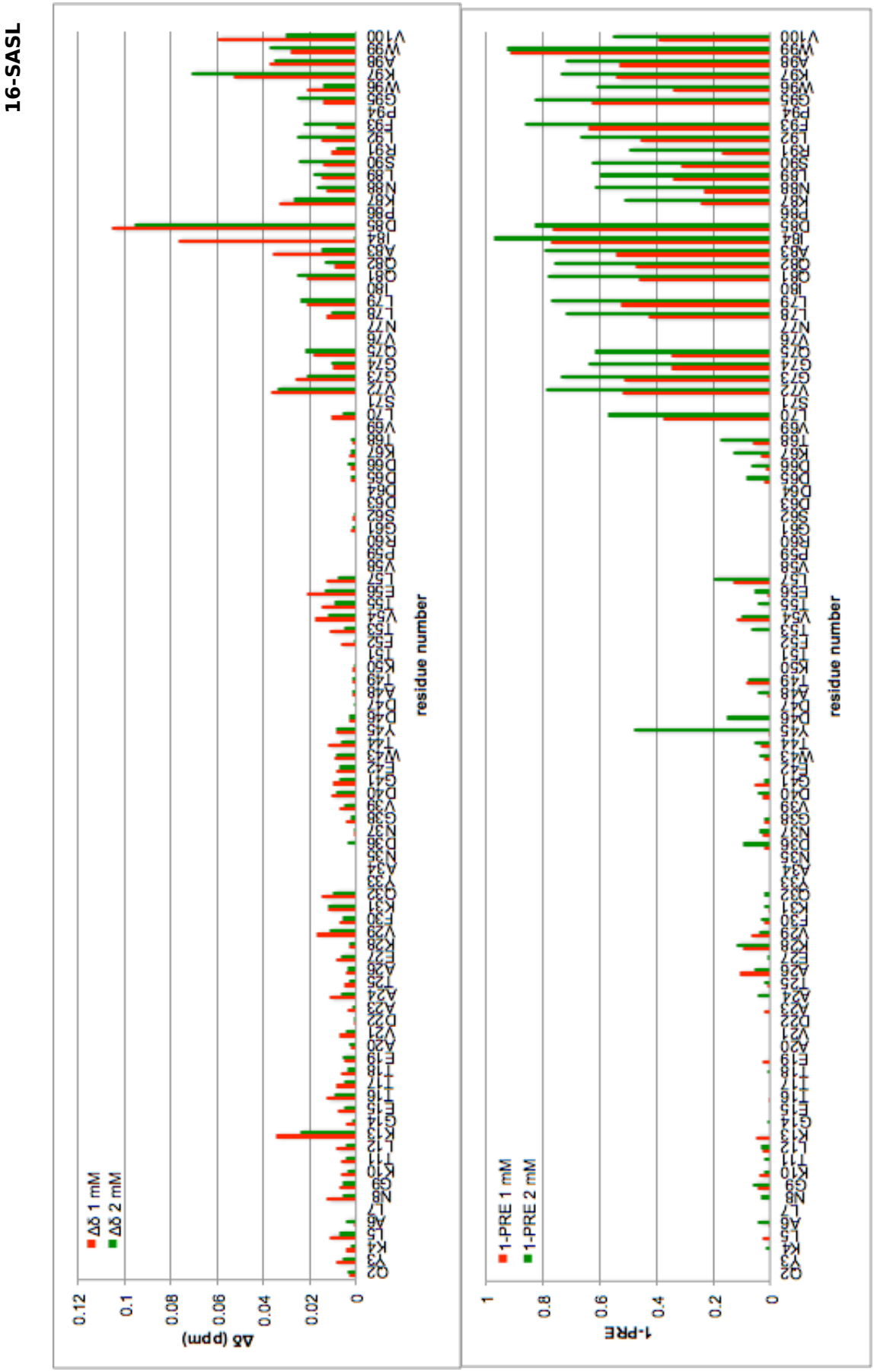




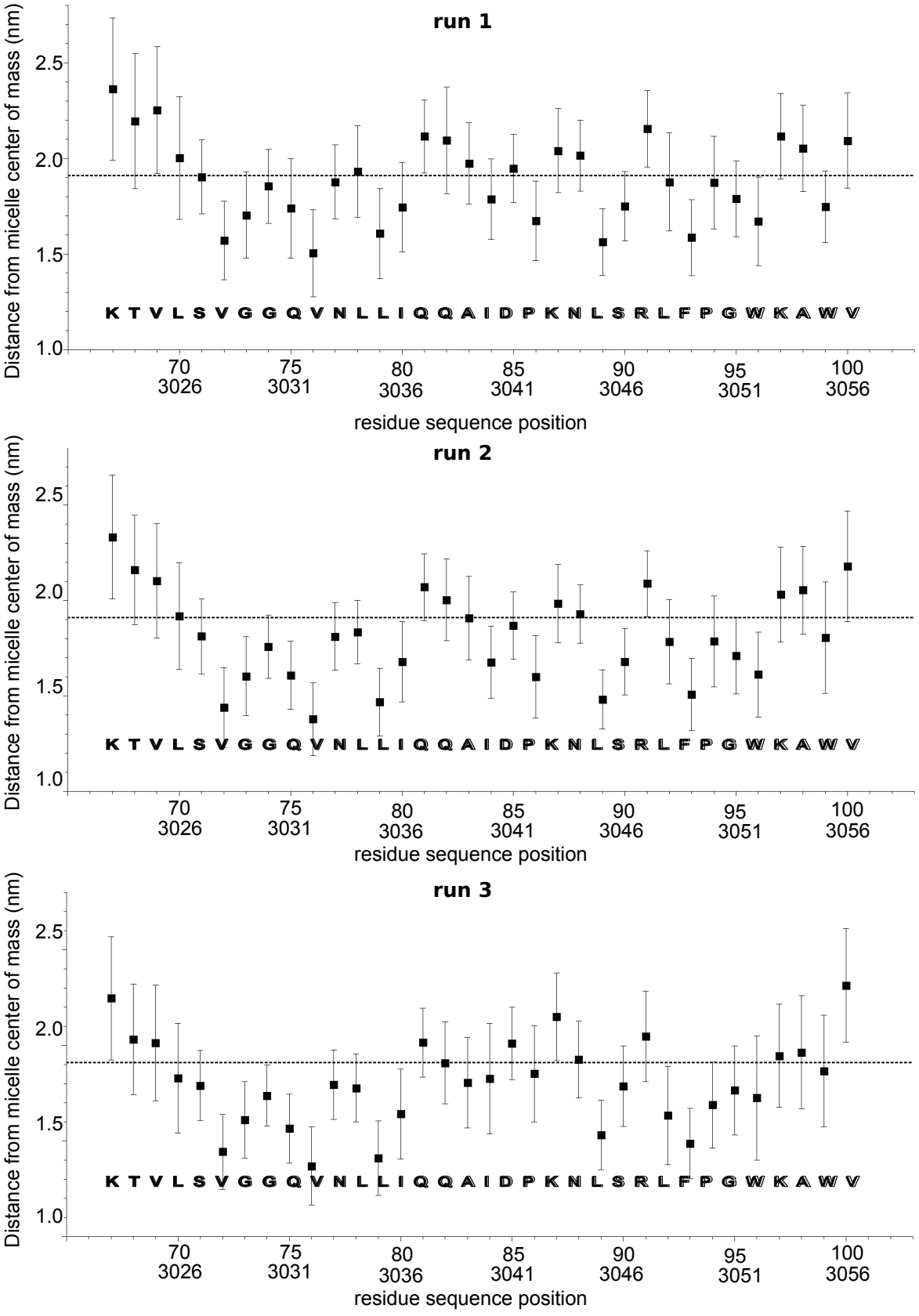
Abd Rahim et al.
SI Fig. S6



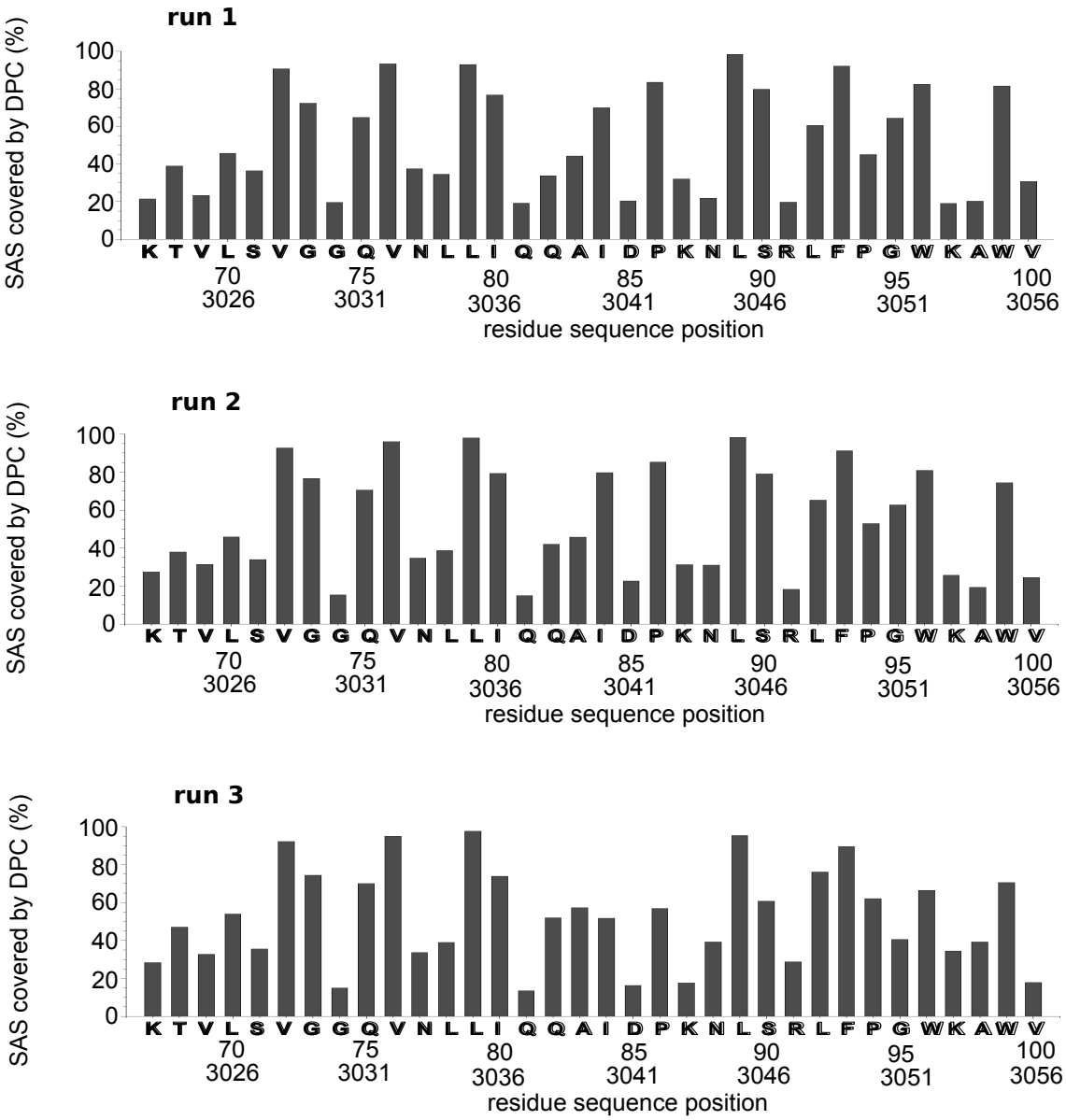
Abd Rahim et al.
SI Fig. S7



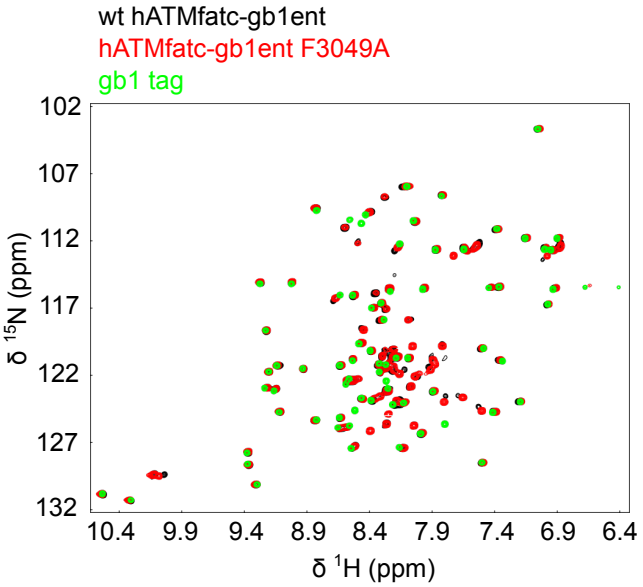
Abd Rahim et al.
SI Fig. S8

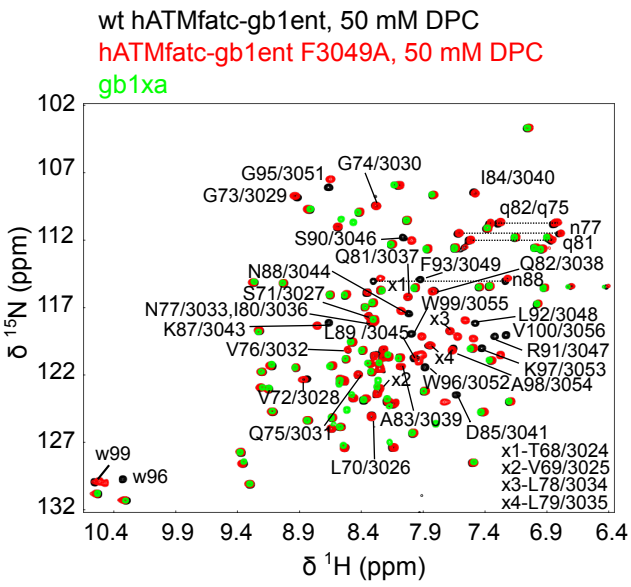


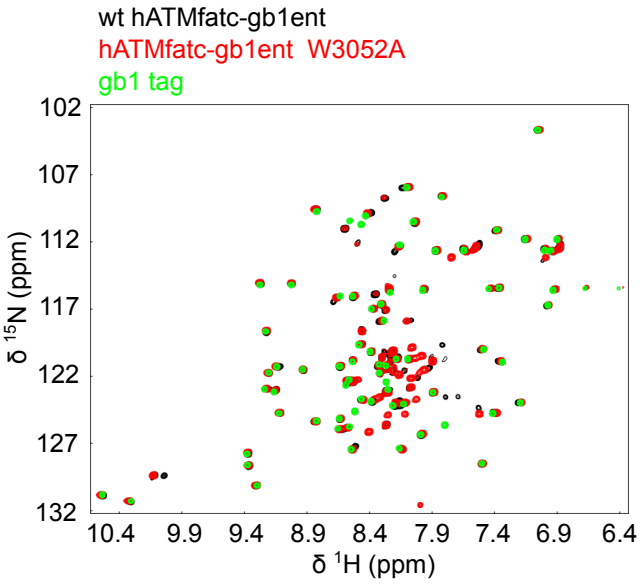
Abd Rahim et al.
SI Fig. S9

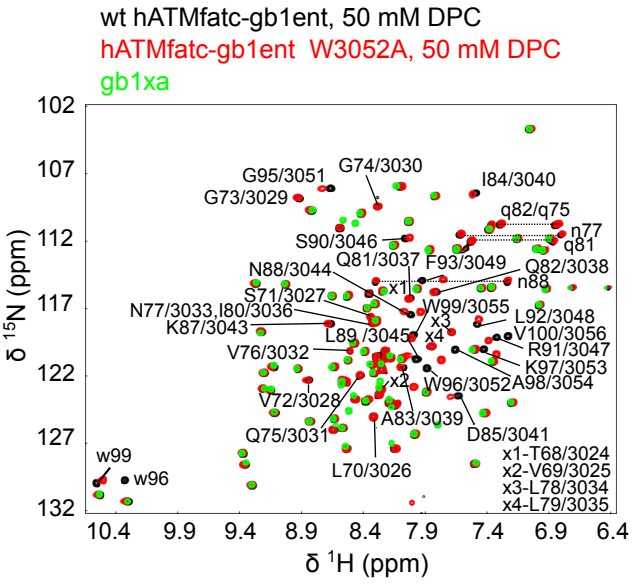


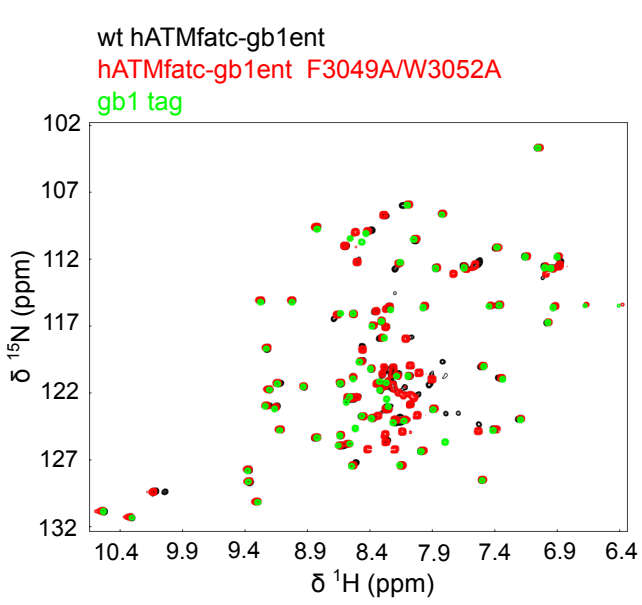
Abd Rahim et al.
SI Fig. S10

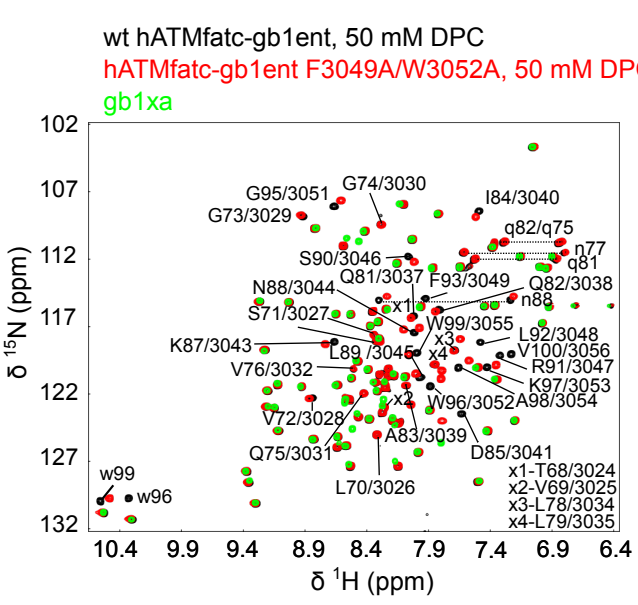




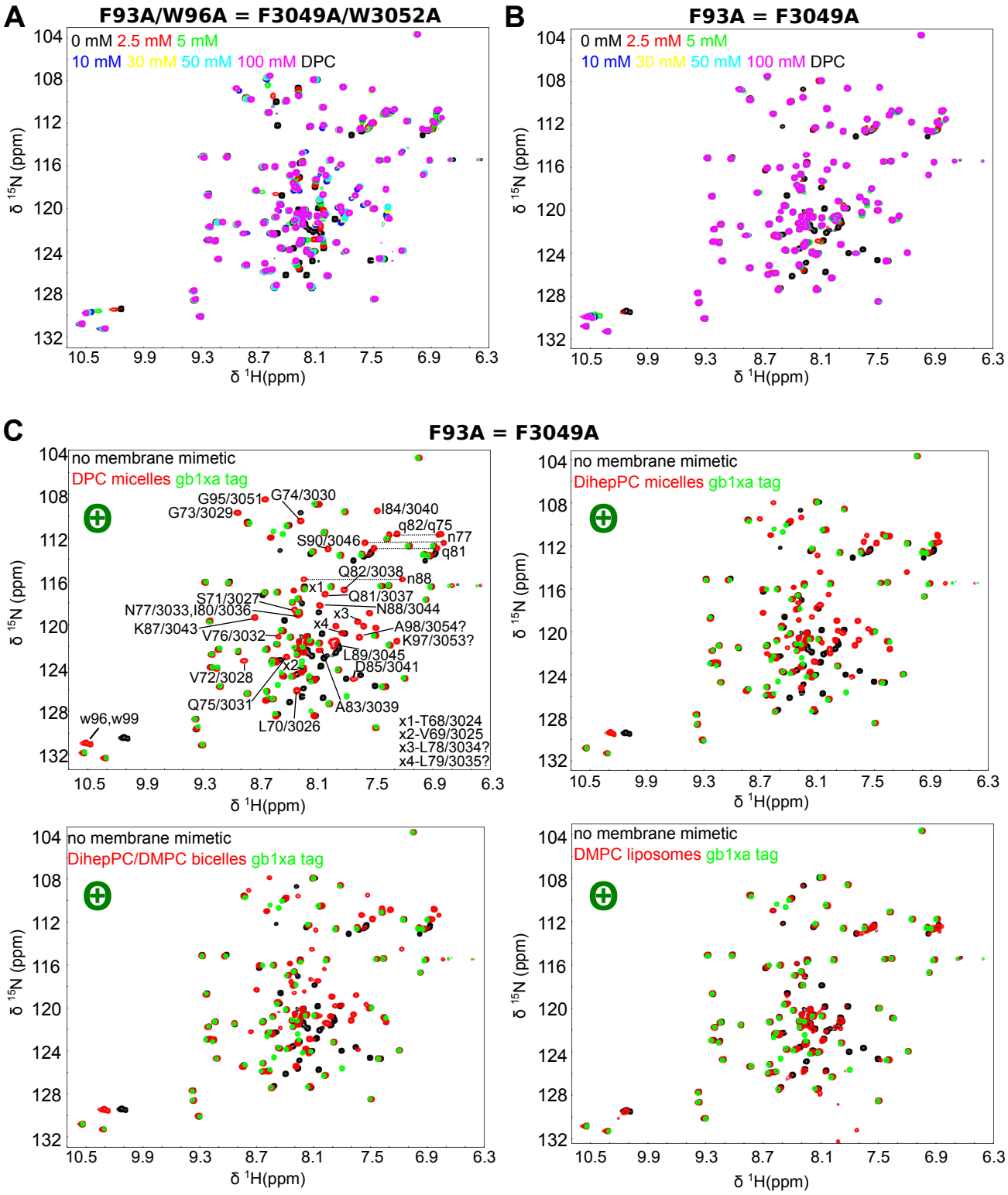








Abd Rahim et al.
SI Fig. S11



Abd Rahim et al.
SI Fig. S12