Biophysical Journal, Volume *120*

Supplemental information

IgG1 conformational behavior: elucidation of the N-glycosylation role via molecular dynamics

Simona Saporiti, Chiara Parravicini, Carlo Pergola, Uliano Guerrini, Mara Rossi, Fabio Centola, and Ivano Eberini

SUPPORTING MATERIAL

IgG1 conformational behavior: elucidation of the N-glycosylation role *via* molecular dynamics

Simona Saporiti[1], Chiara Parravicini[1], Carlo Pergola[2], Uliano Guerrini[1], Mara Rossi[3], Fabio Centola[3¶] and Ivano Eberini[4¶*]

[1]Dipartimento di Scienze Farmacologiche e Biomolecolari, Università degli Studi di Milano, Milano, Italy

[2]Analytical Development Biotech, Merck Serono S.p.A., Rome, Italy; an affiliate of Merck KGaA, Darmstadt, Germany

[3]Global Analytical Pharmaceutical Science and Innovation, Merck Serono S.p.A., Rome, Italy; an affiliate of Merck KGaA, Darmstadt, Germany

[4]Dipartimento di Scienze Farmacologiche e Biomolecolari & DSRC, Università degli Studi di Milano, Milano, Italy

* Corresponding author
E-mail: ivano.eberini@unimi.it

[¶] These authors contributed equally to this work

SUPPORTING METHODS

Molecular dynamics simulations protocol - replicas

Two replicas of molecular dynamics simulations 1 μs long were carried out for each system (aglycosylated, G0 and G0F antibodies) by AMBER 20 software (1) handled by the MOE graphical user interface (2). As well as for the first simulation, the AMBER10:EHT force field (3) was used to parametrize the systems and the TIP3P water model was chosen for solvent parametrization. The systems were configured by MOE: the Langevin thermostat was applied for temperature control, the Monte Carlo barostat was used to set constant pressure, and the simulations were carried out in the NPT ensemble (T=300 K, P=100 kPa). Exactly as in the first simulation, sample time was set to 10 ps and the integration time step to 2 fs. Systems were minimized for 5000 steps and a heating phase was performed for 100 ps. Then, an equilibration phase of 100 ps in the NVT and another one of 200 ps in the NPT ensemble were carried out before the production step.

SUPPORTING FIGURES & TABLES

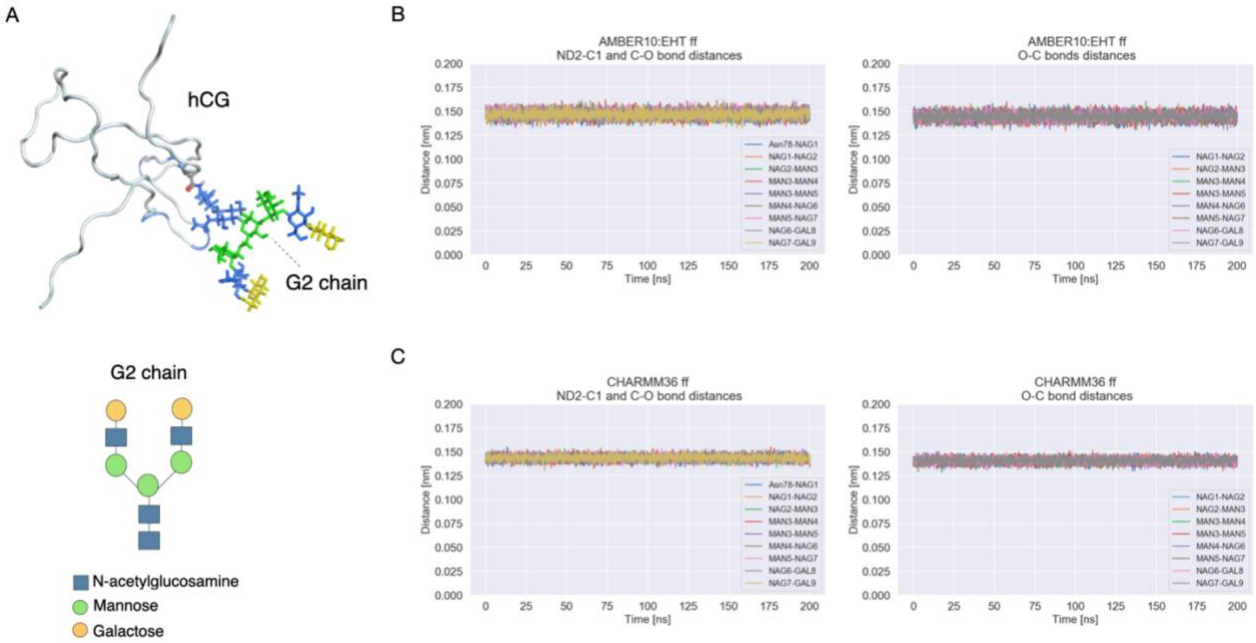


Figure S1. AMBER10:EHT parameters validation for glycans treatment in our study. (A) On the top, the NMR structure of glycosylated hCG used as reference and represented as a ribbon loop. Glycans are represented as sticks colored according to the SNFG system. On the bottom a schematic representation of G2 glycan chain according to the SNFG system. (B) Glycosydic bond distances computed during the MD simulation performed by NAMD 2.13 with AMBER10:EHT force field. (C) Glycosydic bond distances computed during the MD simulation performed by GROMACS 2020.1 with CHARMM36 force field. In both simulations the bond length range is conserved among all sugars couples and among the protein and glycans.

Table S1: Experimentally calculated glycosidic bonds distances of a G2 chain.

Residue Couple	C-O distance (nm)	O-C distance (nm)	ND2-C1 distance (nm)
Asn78-NAG1	-	-	0.144
NAG1-NAG2	0.144	0.139	-
NAG2-MAN3	0.144	0.138	-
MAN3-MAN4	0.144	0.140	-
MAN3-MAN5	0.144	0.140	-
MAN4-NAG6	0.143	0.138	-
MAN5-NAG7	0.144	0.139	-
NAG6-GAL8	0.145	0.139	-
NAG7-GAL9	0.144	0.139	-

The table reports NOE distances calculated for a G2 chain linked to human chorionic gonadotropin (hCG) for which the 3D structure was experimentally solved by NMR (1HD4.pdb). Sugars couples are numbered in a progressive way and the following abbreviations are used: N-acetylglucosamine (NAG), mannose (MAN), galactose (GAL).

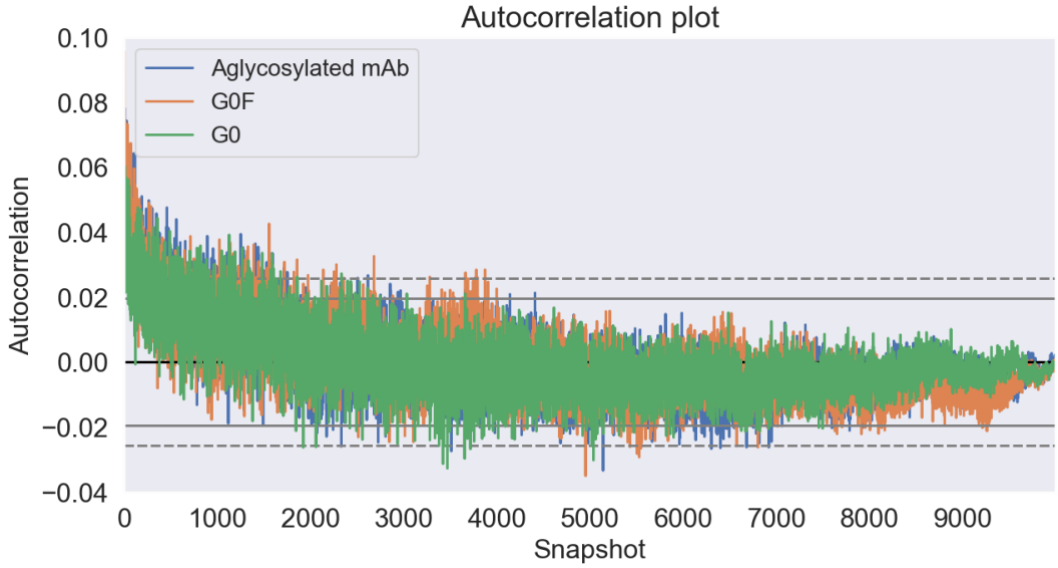


Figure S2. Autocorrelation plot of potential energy. Autocorrelation plot of potential energy computed for three systems. In all cases the autocorrelation value stabilizes under the confidence band within the first 5000 snapshots that corresponds to 50 ns of simulation.

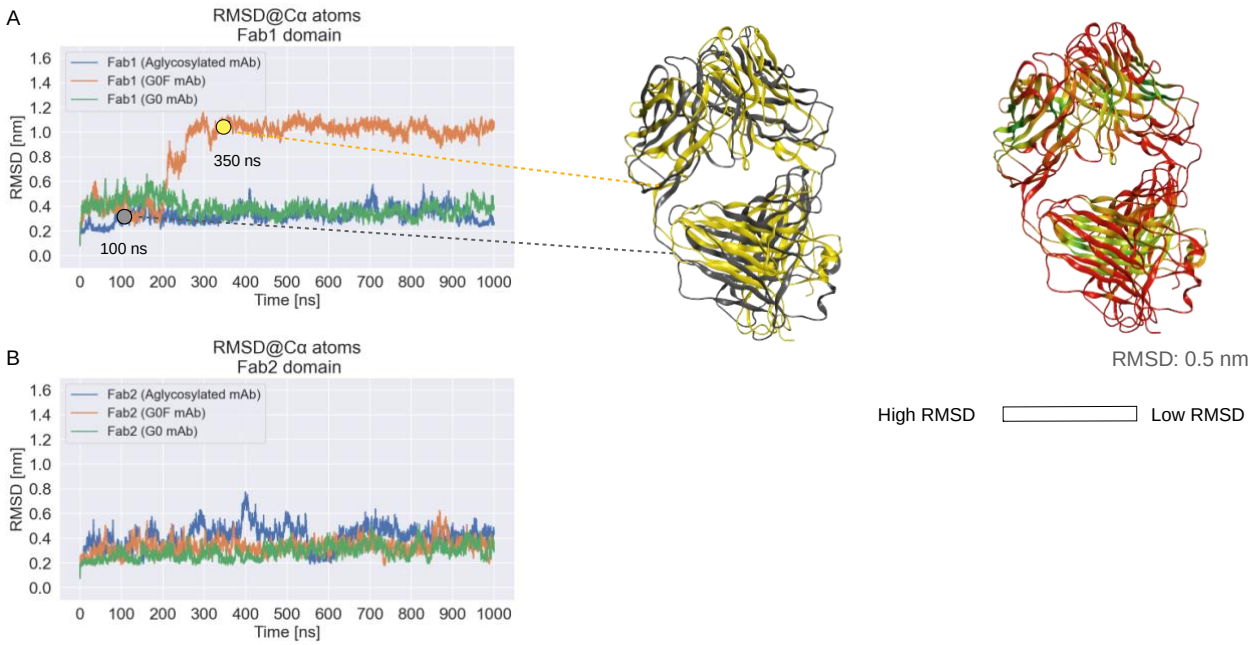


Figure S3. RMSD of single Fab domains. (A) On the left **t**he RMSD profile of Fab1 domains computed for aglycosylated (blue), G0 (green) and G0F (orange) antibodies. On the right, the structural superposition of two representative conformations of G0F Fab1 sampled during the dynamics, corresponding to 100 ns and 350 ns checkpoints. The Fab structures are rendered as ribbons colored both according to constant color and to an RMSD gradient, showing a variation of domain orientation that explains the shift in RMSD profile. (B) The RMSD profile of Fab2 domains computed for aglycosylated (blue), G0 (green) and G0F (orange) antibodies. All the domains show comparable and stable trends further confirming the convergence of simulations.

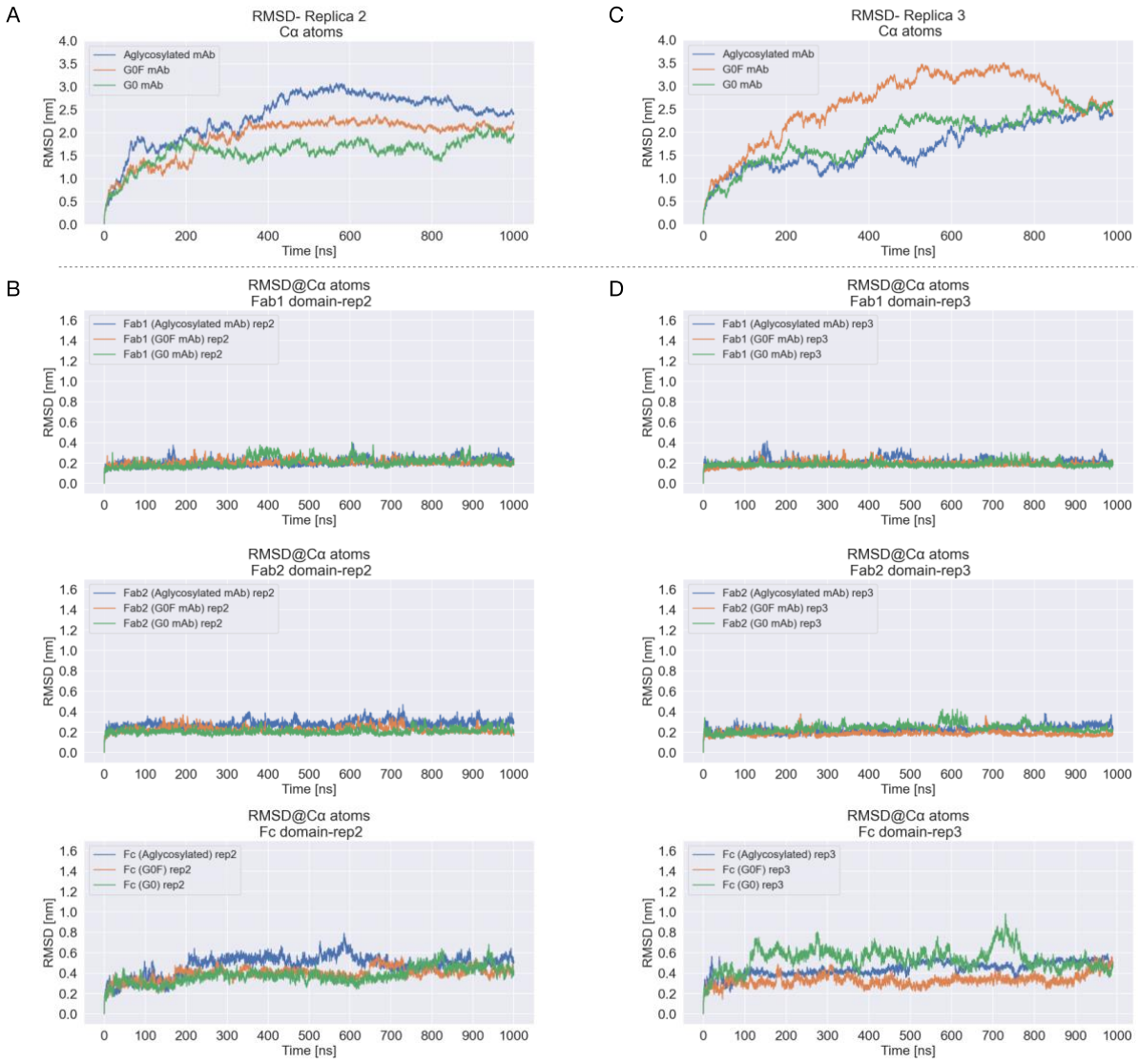


Figure S4: RMSD analysis of replicas 2 and 3. (A) RMSD of C-alpha atoms computed for aglycosylated (blue), G0 (green) and G0F (orange) mAbs in replica 2, showing three different trends as in replica 1. (B) RMSD of single antibody domains in replica 2 (Fab1, Fab2 and Fc), showing that all of them conserve their structure during the dynamics. (C) RMSD of C-alpha atoms computed for aglycosylated (blue), G0 (green) and G0F (orange) mAbs in replica 3, showing three different trends as in replicas 1 and 2. (D) RMSD of single antibody domains in replica 3 (Fab1, Fab2 and Fc), showing that all of them conserve their structure during the dynamics, with some variation observed for Fc in G0 antibody.

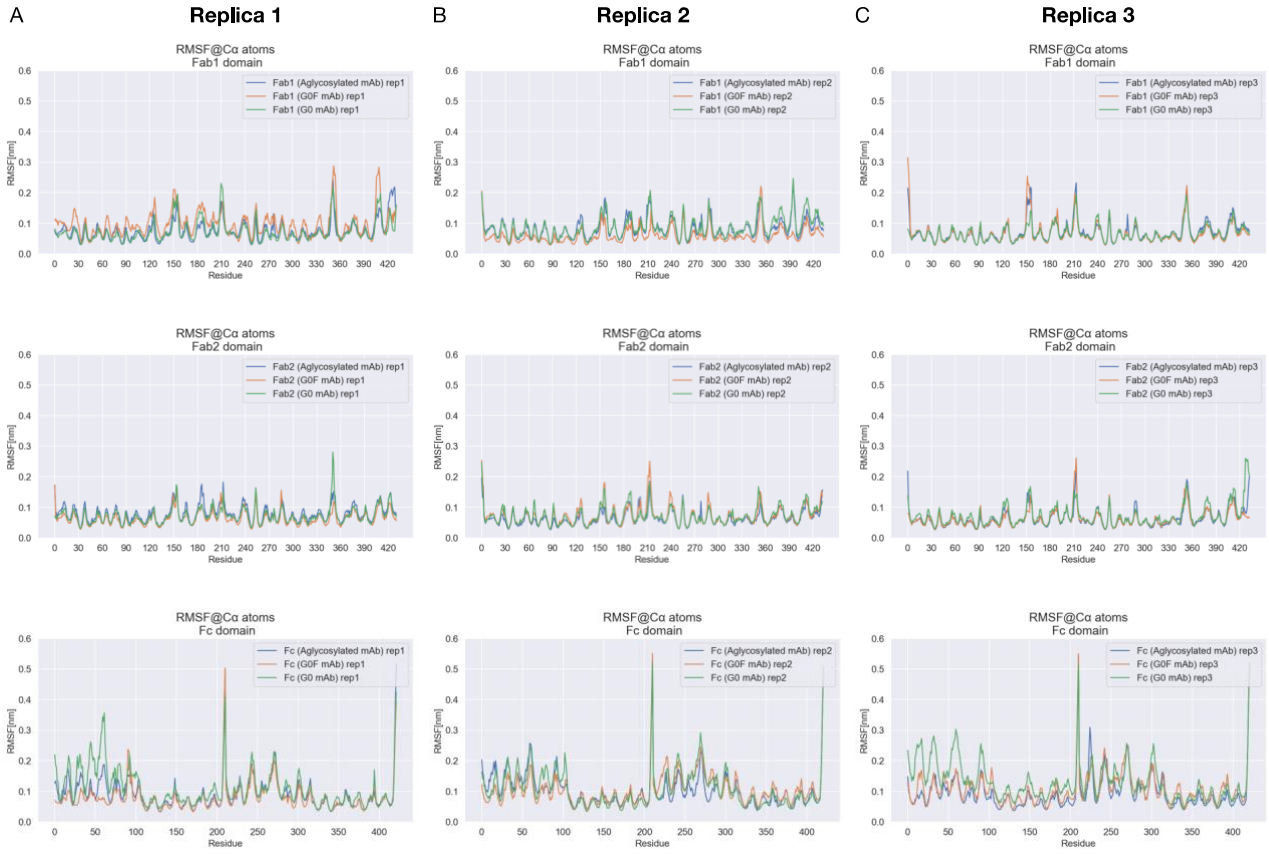


Figure S5: RMSF of C-alpha atoms in the three replicas. In panels (A-C) the RMSF of C-alpha atoms computed for each antibody domain in replicas 1, 2 and 3, respectively. The calculation was performed excluding the first 200 ns of trajectories, since considered as equilibration phase, with respect to the mean structure. All domains show comparable fluctuation trends, with small differences recognized for Fab1 G0F in replica 1, that fluctuates more than others, explaining also the different RMSD trend of this domain (see Fig. S3) and for the first CH2 domain of Fc G0 that fluctuates more in all the replicas, suggesting a conformational change in this region.

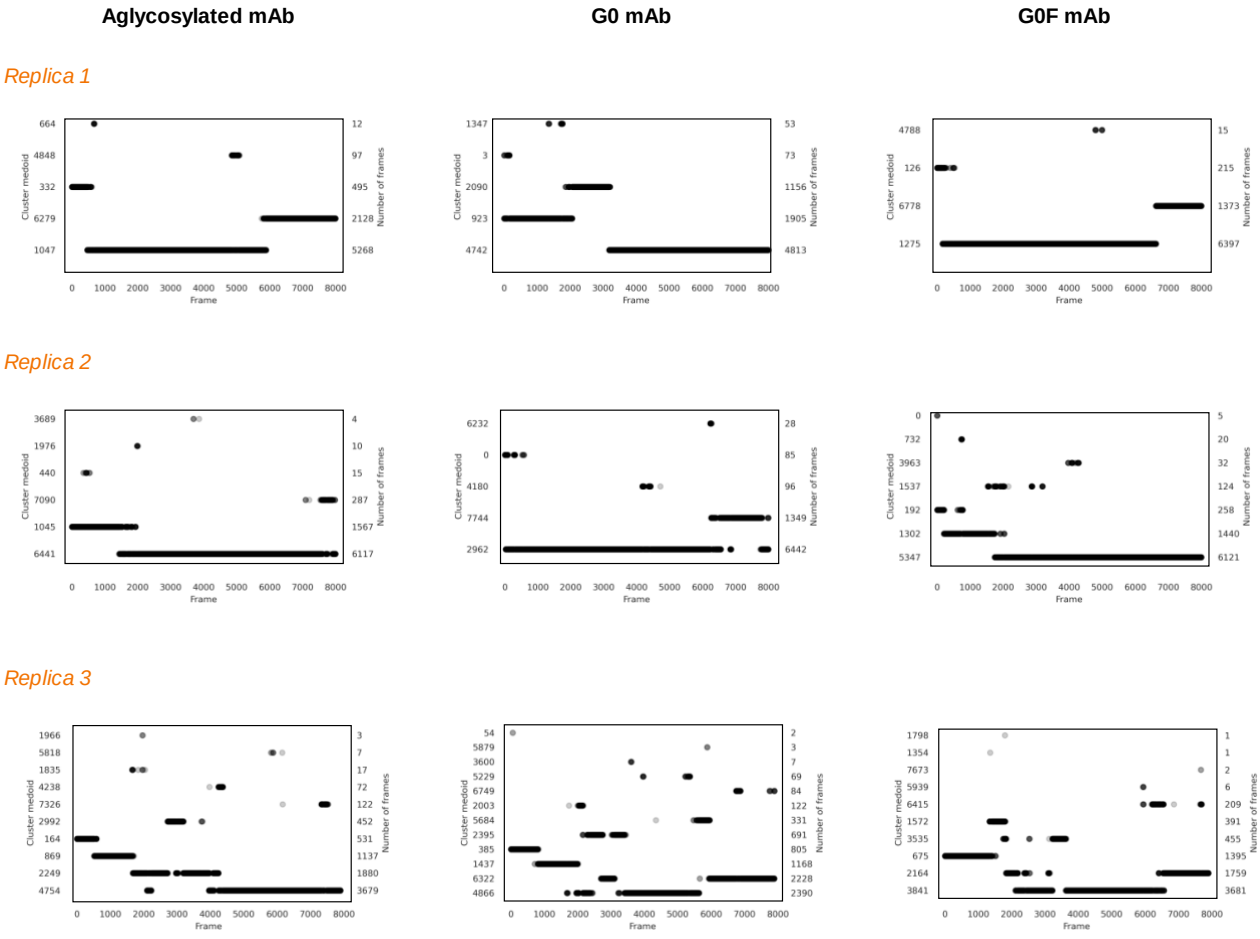


Figure S6: Cluster analysis. Cluster matrices showing the identified groups with an RMSD-based 7 Å threshold. On x-axis, the timeframes (the first 200 ns were excluded from calculation), on y-axes, the number of frames included in each cluster and the frame corresponding to the medoid structure.

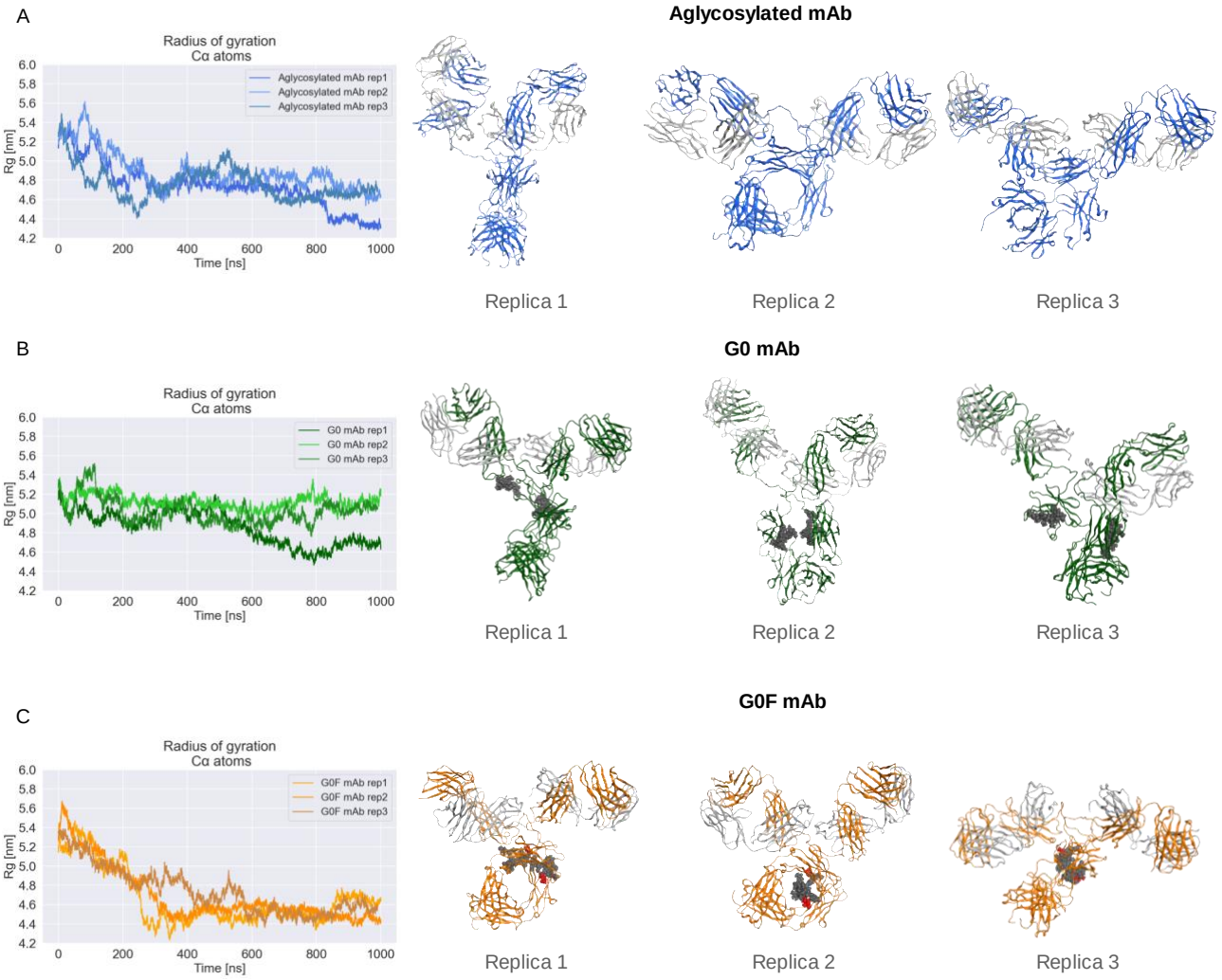


Figure S7: Radius of gyration of C-alpha atoms by replica. (A) The Rg computed for the aglycosylated mAb in the three simulations and the corresponding medoid structures isolated by clustering show the propensity of this antibody to reach different states. (B) The Rg computed for the G0 mAb in the three simulations and the corresponding medoid structures isolated by clustering show the propensity of this antibody to maintain a Y-shaped form. (C) The Rg computed for the G0F mAb in the three simulations and the corresponding medoid structures isolated by clustering show the propensity of this antibody to assume a T-shaped form.

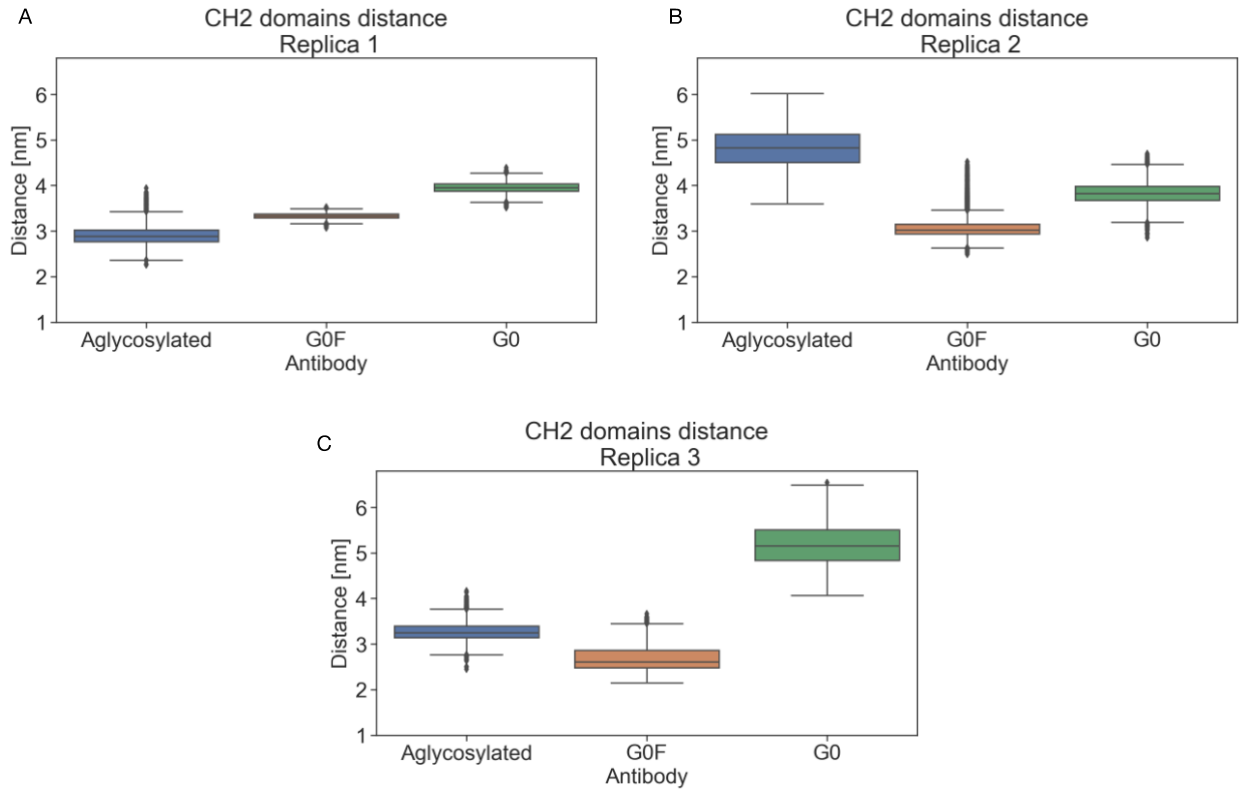


Figure S8: Distribution of the distances between CH2 domains by replica. (A-C) Box plots of CH2 domains distances computed between the two glycosylated Asn residues (Asn301, Asn297 according to the standard EU numbering) in the last 6,000 timeframes of each simulation. Globally, this analysis shows how the CH2 domains of Fc portion in G0 antibody are more distant (approx. 4-5 nm) than in G0F mAb which values span between approx. 2.5 and 3.5 nm. Considering the aglycosylated antibody, distance values can vary a lot, exploring conformations similar to both G0 and G0F, because of the higher instability of this antibody. On y-axis the distance in nm is reported; on x-axis the antibody species. Outliers are shown as diamonds.

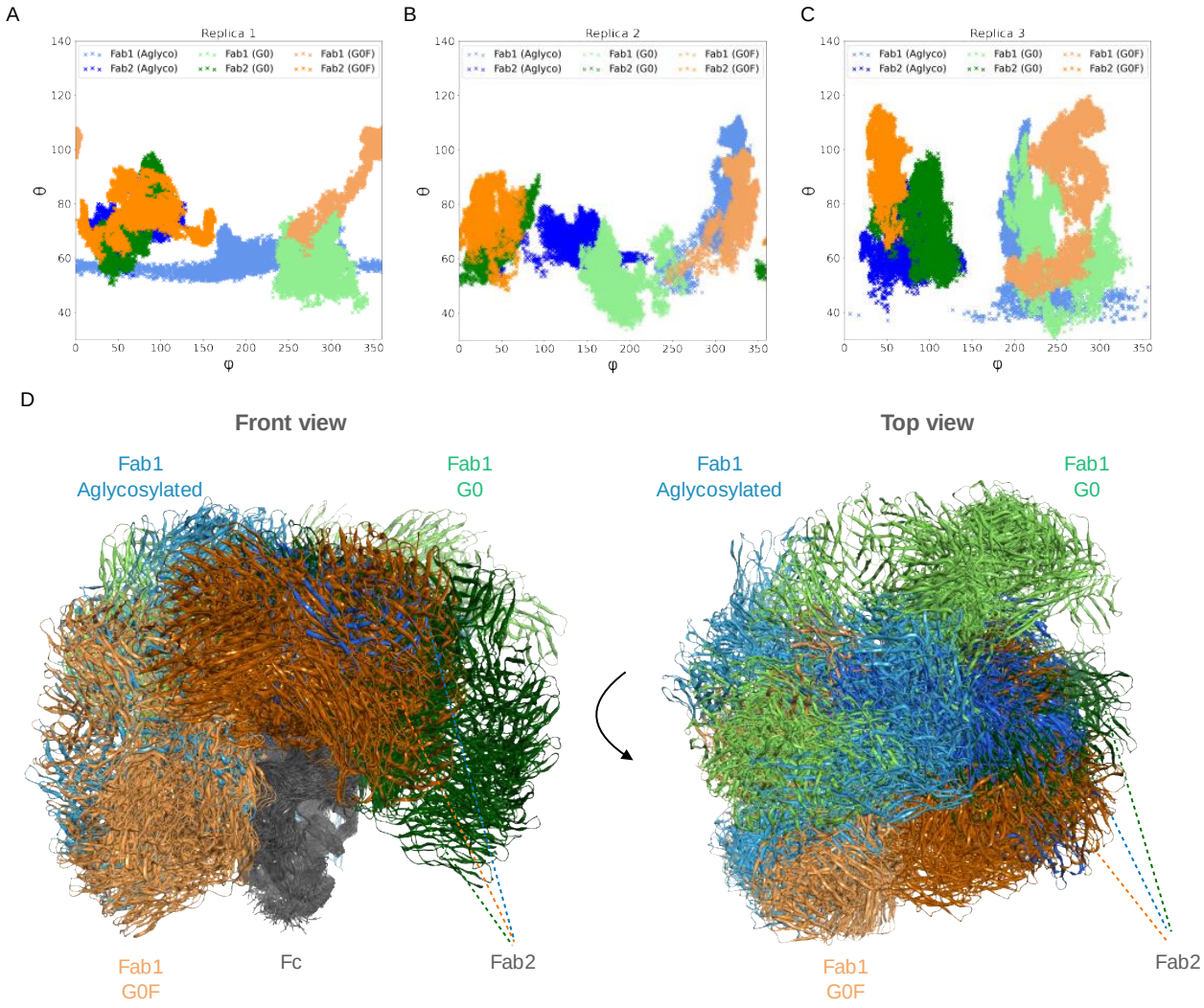


Figure S9: Angles φ and θ by replica and structural superposition of 30 conformations per antibody. (A-C) Scatter plots of φ and θ angles computed for Fab1 and Fab2 of each antibody in each replica. As a result, in all the simulations Fab2 positions are quite conserved exploring very similar values of φ and θ angles, while Fab1 domains show different trends. (D) The structural alignment of 30 conformations per each mAb (10 per replica) better show that the Fab2 position is conserved in all antibodies, while Fab1 reaches different conformations according to the glycosylation pattern. The aglycosylated antibody results to explore more than others, probably due to the absence of glycans allosteric modulation.

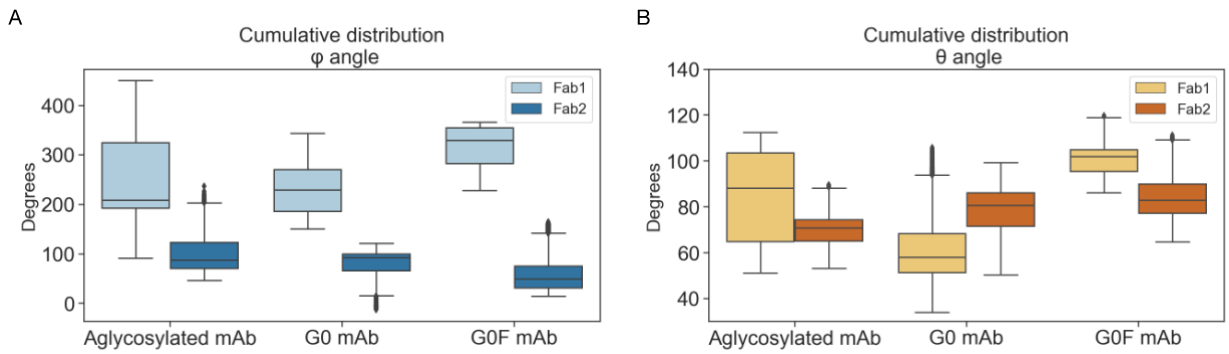


Figure S10: Cumulative distribution of φ and θ angles. Box plots representing the distribution of φ (A) and θ (B) angles for Fab1 and Fab2 considering 18,000 timeframes as the sum of the last 6,000 timeframes of each simulation. To better represent φ angle values and to avoid graphical issues due to the periodicity, a value of 360° was summed to the original angle when φ > 90°, while a value of 360° was subtracted when φ > 270°. As a result of the distribution, in all the species Fab2 explores less than Fab1 with more comparable values on both φ and θ dimensions. Moreover, focusing on Fab1, the aglycosylated antibody shows the highest variability on both angles, the G0 mAb results to explore especially on phi, and the G0F one tends to change both φ and θ, suggesting in the first case a rotation of the Fab and in the second one the collapse onto the Fc.

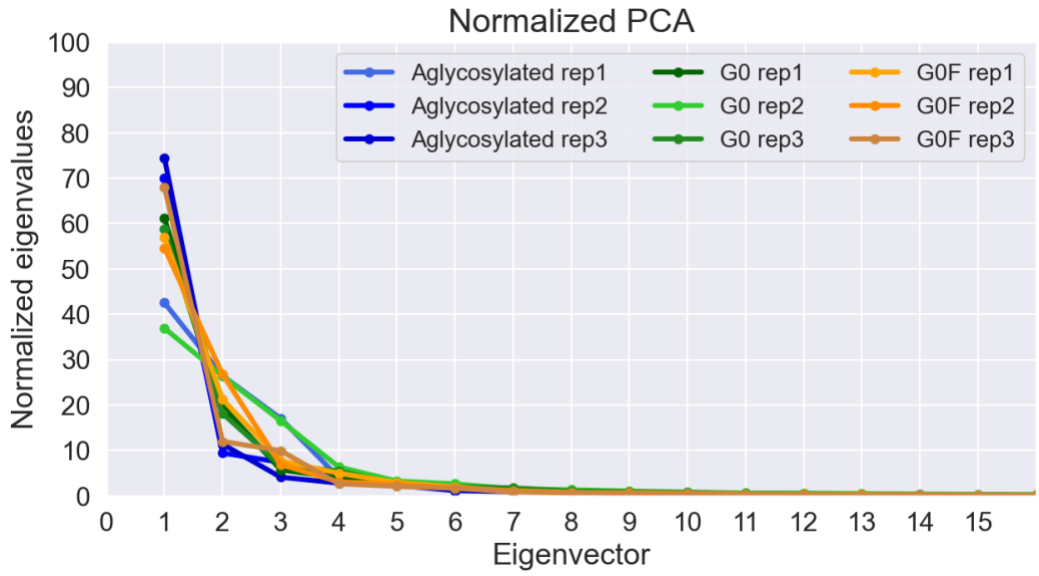


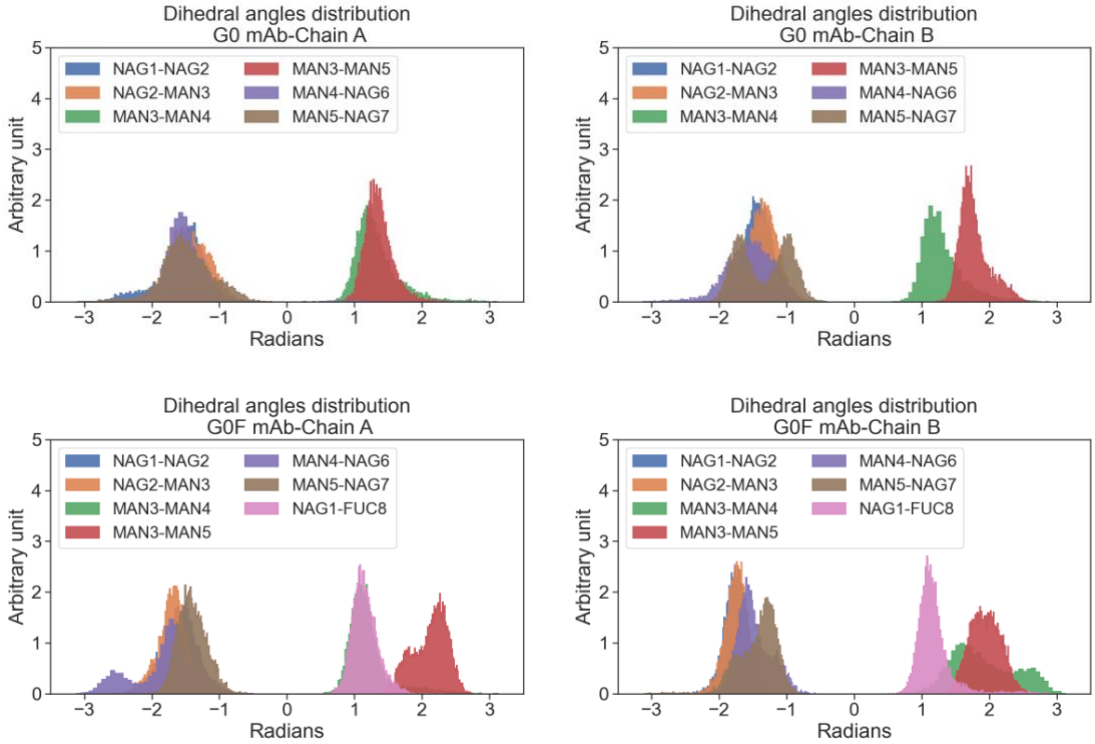
Figure S11: Normalized principal component analysis by replica. In this plot, the normalized eigenvalues *vs* eigenvector numbers. The sum of the first three component describes for each antibody more than 80% of its dynamics, that can be considered sufficient to cover the principal motions of the proteins. However, the 95% of motion is globally described by the first 5 PC in the aglycosylated system, by the first 7 in the G0 one and by the first 6 in G0F antibody.

Table S2: Eigenvalues corresponding to the first three eigenvectors of replica 1.

	Aglycosylated mAb	G0 mAb	G0F mAb
Eigenvector 1	338.99	518.02	573.60
Eigenvector 2	210.73	167.13	214.44
Eigenvector 3	135.82	48.28	77.40

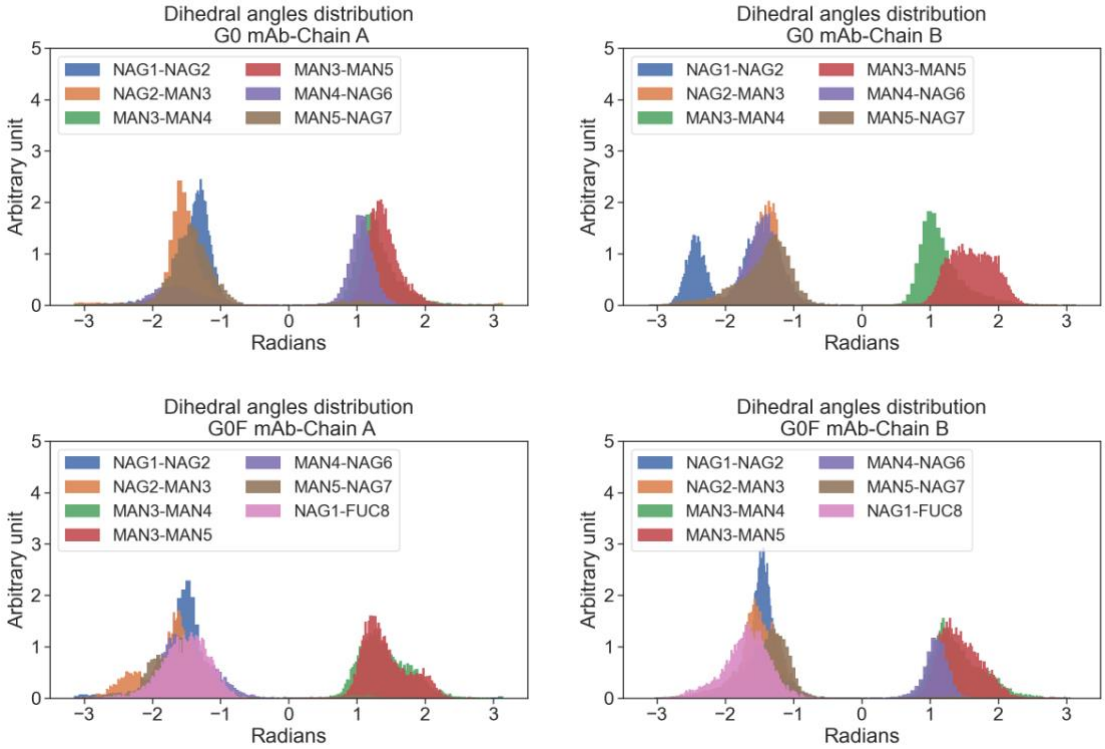
A

Replica 1



B

Replica 2



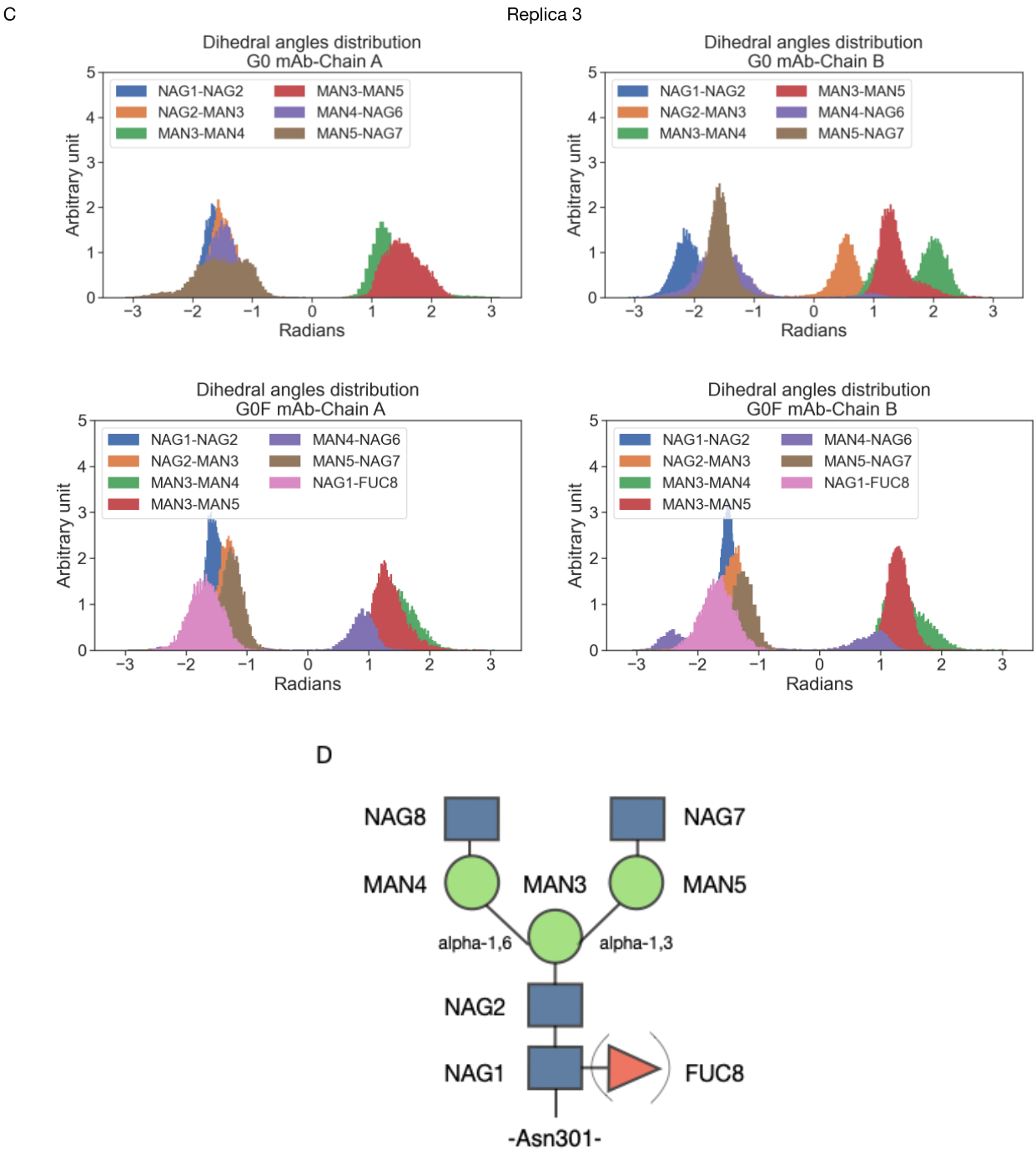


Figure S12: Normalized dihedral-angle distributions of glycan chains by replica. The dihedral-angle distributions of G0 and G0F chains in replica 1 (A), replica 2 (B) and replica 3 (C). All the angles show a unimodal distribution that suggests the stability of glycans conformation. (D) The schematic representation of G0/G0F chains with the numbering used in the plots.

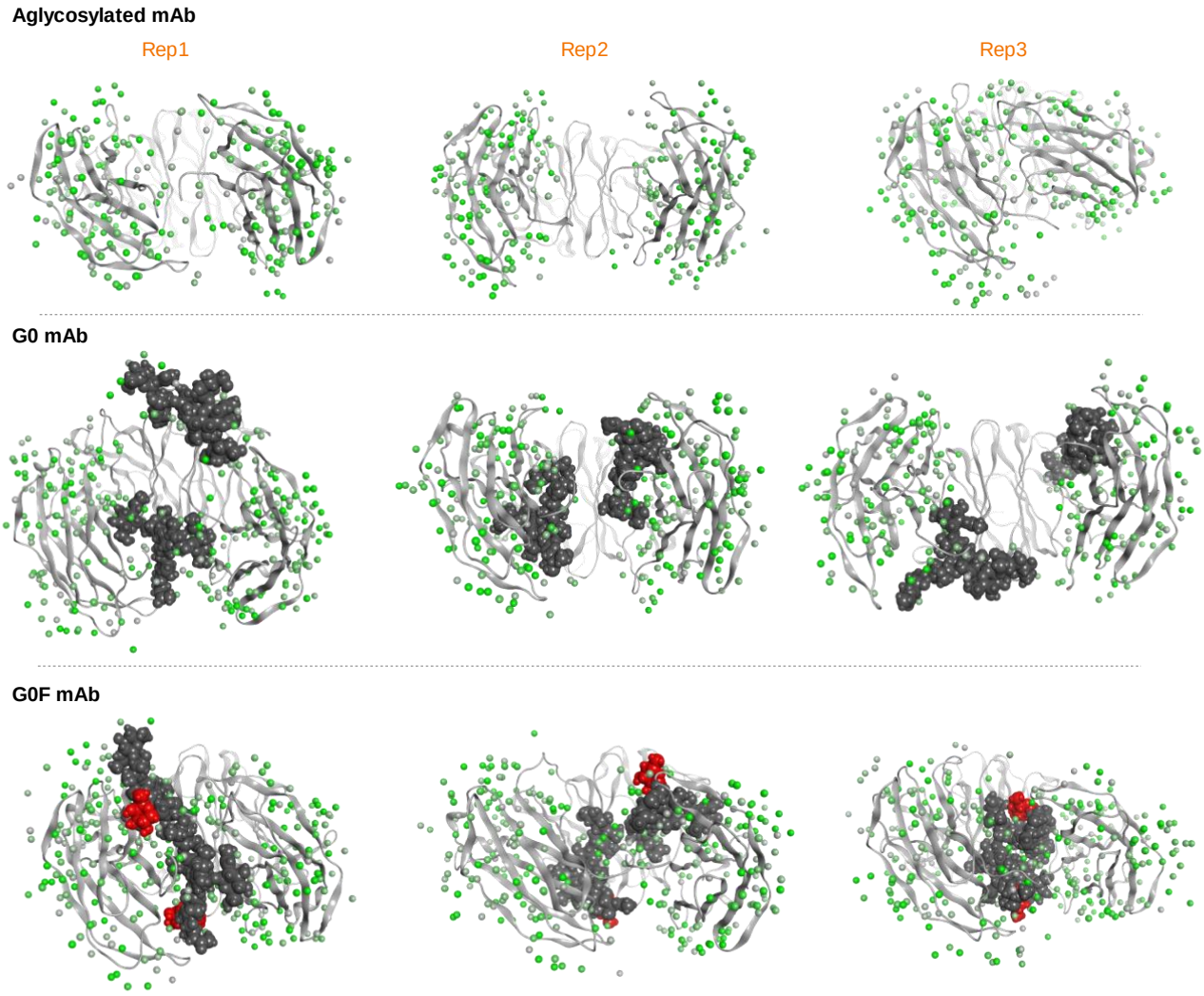


Figure S13: Solvent analysis by replica. Solvent analysis performed by 3D-RISM on aglycosylated, G0 and G0F mAbs in the three replicas. In the picture, the top view of CH2 domains is shown as grey ribbons and glycans as dark grey spheres, with fucose in red. Water molecules with a negative hydration free energy value (dG) are shown as green spheres, pointing out that in G0F Fc the water placement within the cavity is more favored than in others.

SUPPORTING REFERENCES

1. D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, C. Jin, K. Kasavajhala, M.C., and P.A.K. 2021. Amber 2021. .
2. MOE. 2021. Chemical Computing Group Inc. Molecular Operating Environment (MOE); Chemical Computing Group Inc. 1010 Sherbooke St. West, Suite# 910: Montreal, QC, Canada,. .
3. Case, D.A., Darden, T.A., Cheatham, T.E., Simmerling, C.L., Wang, J., Duke, R.E., Luo, R., Crowley, M.R.C.W., Walker, R.C., Zhang, W. and Merz, K.M.Case, D.A., Darden, T.A., Cheatham, T.E., Simmerling, C.L., Wang, J., Duke, R.E., Luo, R., Crowley, M.R.C.W, K.M. 2008. AMBER 10. San Francisco: University of California.

APPENDIX

Evaluation of the convergence of simulations

The evaluation of systems convergence was performed by the block analysis technique and specifically computing structural observables and principal component analysis (PCA) on trajectory segments 200 ns long. In order to avoid background noise due to the high flexibility of the hinge, the analysis was performed on structured domains, namely the two Fabs and the Fc. The distribution of RMSD and Rg together with the RMSF profiles are provided below, showing very similar median values between the last four blocks, and suggesting that the convergence is reached after approx. 200 ns. Then, PCA was performed for single domains, isolating the first two principal motions identified for each segment. According to the results, in several cases the first two segments (0-400 ns) present different projections with respect to the other three, that instead shows overlapped plots. Overall, this result suggests that principal modes can be considered stabilized after 400 ns of simulation. Considering both the structural observables and the PCA, we decided to exclude the first 200 ns of MD for the cluster and H-bonds analysis.

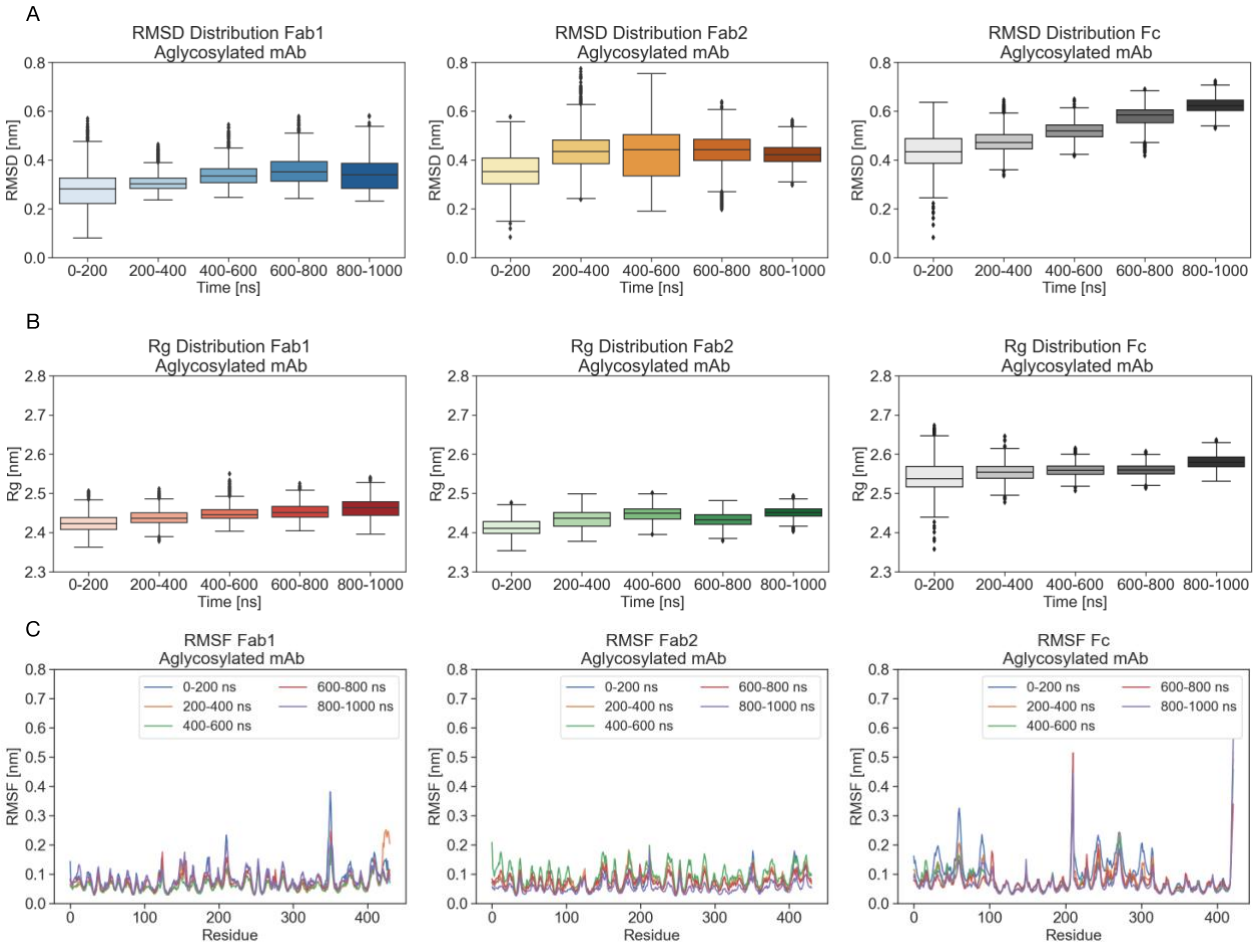


Figure S14: RMSD and Rg distribution and RMSF profiles of aglycosylated mAb domains. (A-B) Box plots showing the distribution of RMSD and Rg computed on C-alpha atoms for each antibody domain in each trajectory segment. Outliers are represented as diamonds. (C) RMSF profiles of C-alpha of each antibody domain in each segment.

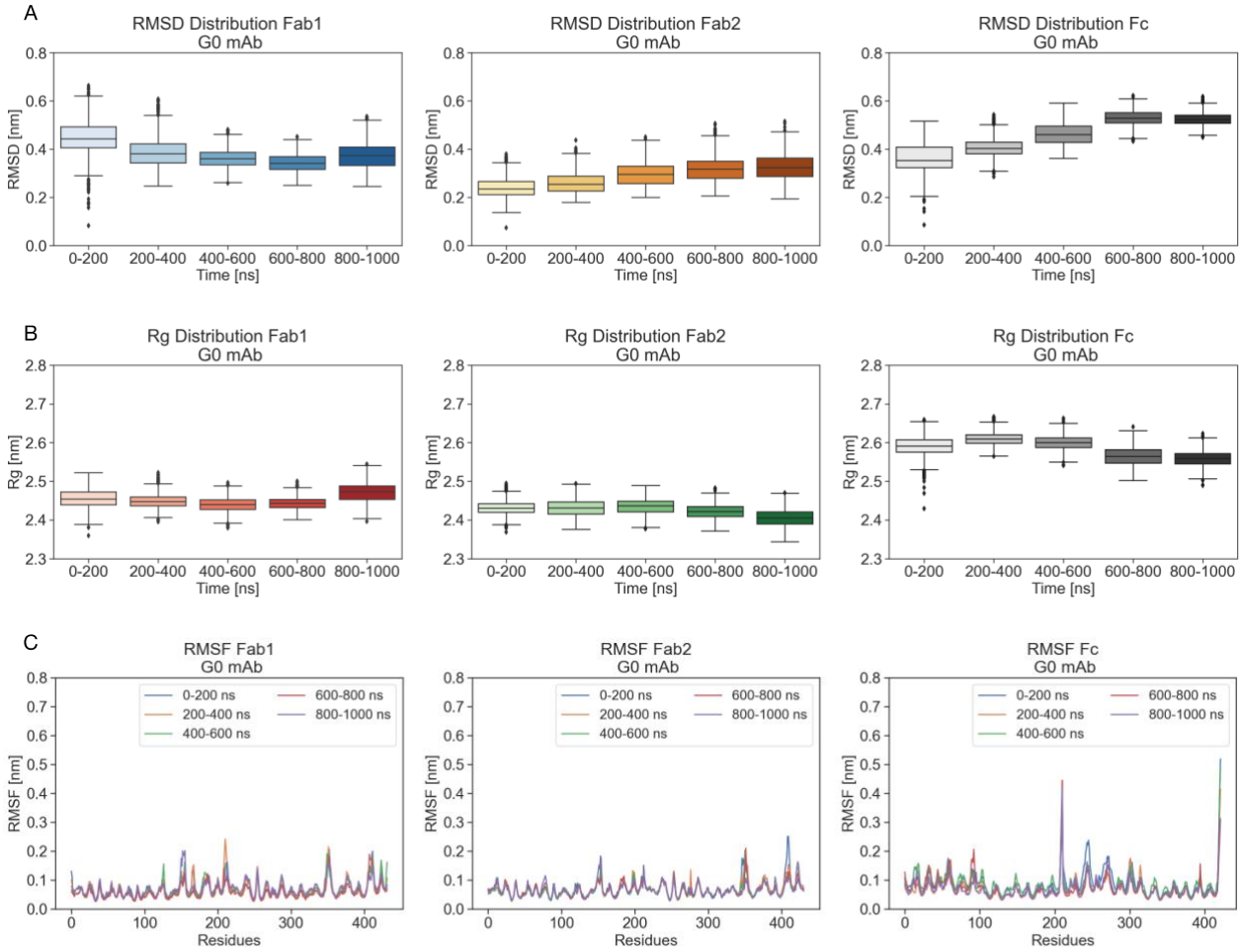


Figure S15: RMSD and Rg distribution and RMSF profiles of G0 mAb domains. (A-B) Box plots showing the distribution of RMSD and Rg computed on C-alpha atoms for each antibody domain in each trajectory segment. Outliers are represented as diamonds. (C) RMSF profiles of C-alpha of each antibody domain in each segment.

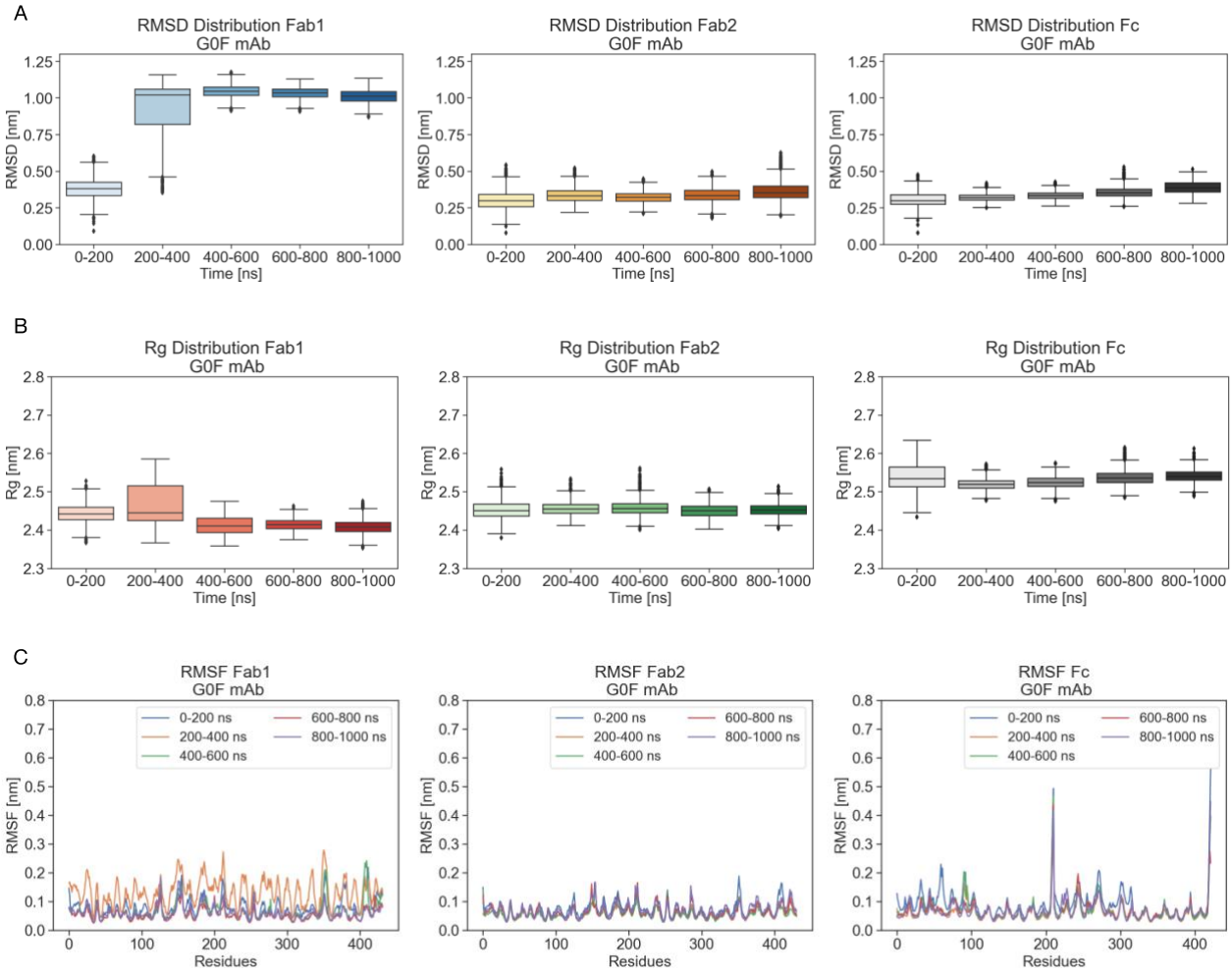


Figure S16: RMSD and Rg distribution and RMSF profiles of G0F mAb domains. (A-B) Box plots showing the distribution of RMSD and Rg computed on C-alpha atoms for each antibody domain in each trajectory segment. Outliers are represented as diamonds. (C) RMSF profiles of C-alpha of each antibody domain in each segment.

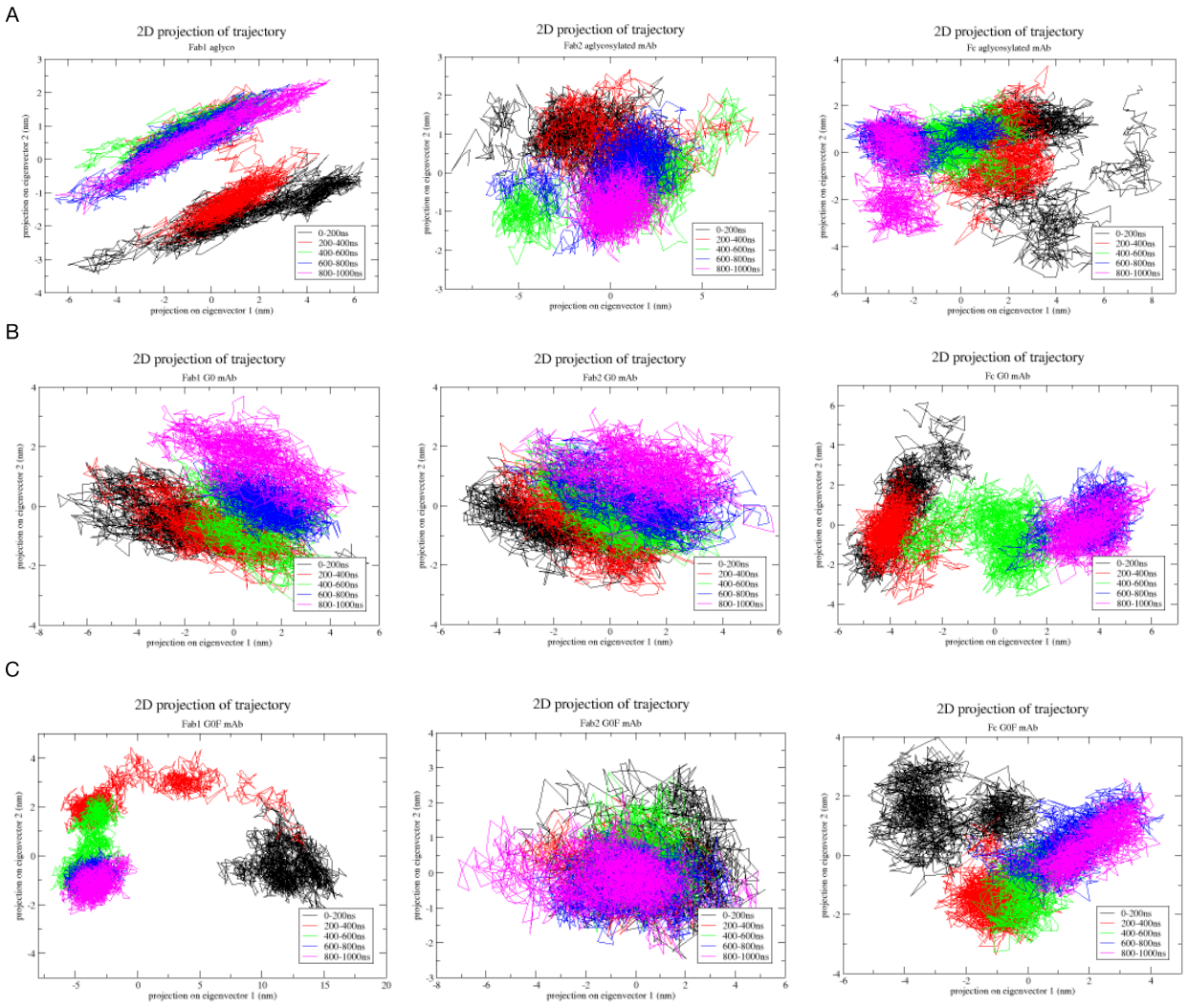


Figure S17: Principal component analysis. Bi-dimensional projections of the first two principal components computed per aglycosylated (A), G0 (B) and G0F (C) mAb domains in each trajectory segment.