Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

S.I. - Influence of metal binding on the conformational landscape of neurofilament peptides

David Silva-Brea,[1] David de Sancho,[1,2] Xabier Lopez,[2]

E-mail:

First and Second Coordination Shell calculations.

Table 1: Δ_fH and Δ_fG values for First Coordination Shell and Second Coordination Shell systems.

Index	Systems	ΔH (kcal/mol)	ΔG (kcal/mol)
1	$[Al(H_2O)_5(SerP)]^{+1}$	-78.8	-73.6
2	$[Al(H_2O)_4(SerP)]^{-1}$	-128.7	-118.5
3	$[Al(H_2O)_4(SerP)(Glu)]$	-100.9	-91.7
4	$[Al(H_2O)_{46}(Glu)_2]^{+1}$	-61.0	-53.1
5	$[Al(H_2O)_3(SerP)_2(Glu)]^{-2}$	-139.4	-124.5
6	$[Al(H_2O)_3(SerP)(Glu)_2]^{-1}$	-114.0	-101.3
1’	$[Al(H_2O)_5(Lys@SerP)]^{+2}$	-61.3	-51.0
2’	$[Al(H_2O)_4(Lys@SerP)_2]^{+1}$	-138.6	-119.0
3’	$[Al(H_2O)_4(Lys@SerP)(Glu)]^{+1}$	-85.5	-75.5
3*’	$[Al(H_2O)_4(Lys@SerP)(Lys@Glu)]^{+2}$	-134.6	-116.7
4’	$[Al(H_2O)_4(Lys@Glu)(Glu)]^{+2}$	-51.5	-43.5
5’	$[Al(H_2O)_3(Lys@SerP)_2(Lys@Glu)]^{+1}$	-177.5	-150.6
6’	$[Al(H_2O)_3(Lys@SerP)(Lys@Glu)(Glu)]^{+1}$	-133.3	-112.3

DFT calculations

The FCS and SCS systems only consider the first two coordination shells. This allows quick calculations with small systems but is limited by the lack of interactions with the rest of the peptide. When FCS/SCS systems are converted to the entire peptide systems, some of the resultant structures share the same FCS/SCS. This is represented in the following Figure 1 where, for example, the 6’ indexed FCS/SCS system is converted into two different species from different groups: $NF13Al^{S2}_{S2E10E11}$ and $NF13Al^{S7}_{S7E10E11}$.

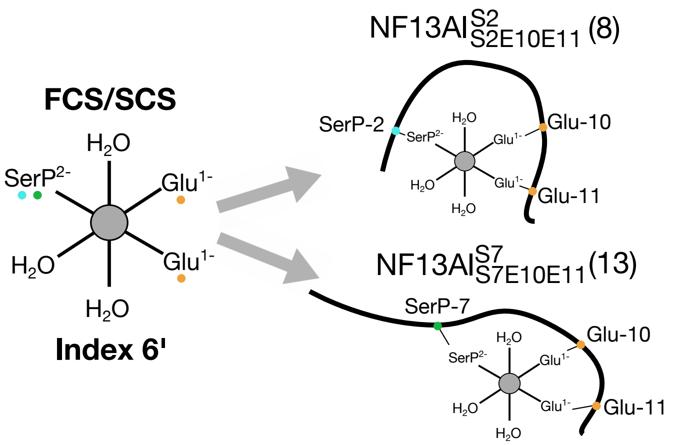

Figure 1: Representation of the conversion of FCS/SCS systems to entire peptide systems. The grey ball in the center corresponds to the Al^{3+} cation, binding 6 ligands. The same FCS/SCS can be observed in the two right cases, but these systems consider the whole 13-residue peptide instead of only the side-chains of the residues.

Images of systems.

Images of the 13-residue + aluminum systems. Reference systems are shown in Figure 2, systems from group S2 shown in Figure 3, systems from group S7 shown in Figure 4 and systems from group S2S7 shown in Figure 5.

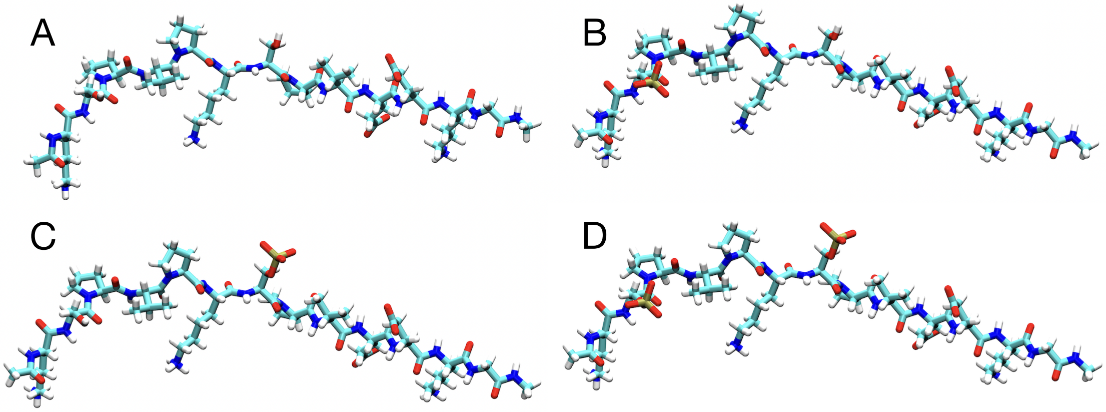

Figure 2: Reference systems. The peptide is considered both with and without any phosphorylation. This allows to characterize the impact of the phosphorylation by itself. This structures are in a configuration of β-sheet. A: $NF13$; B: $NF13S2$; C: $NF13S7$; D: $NF13S2S7$.

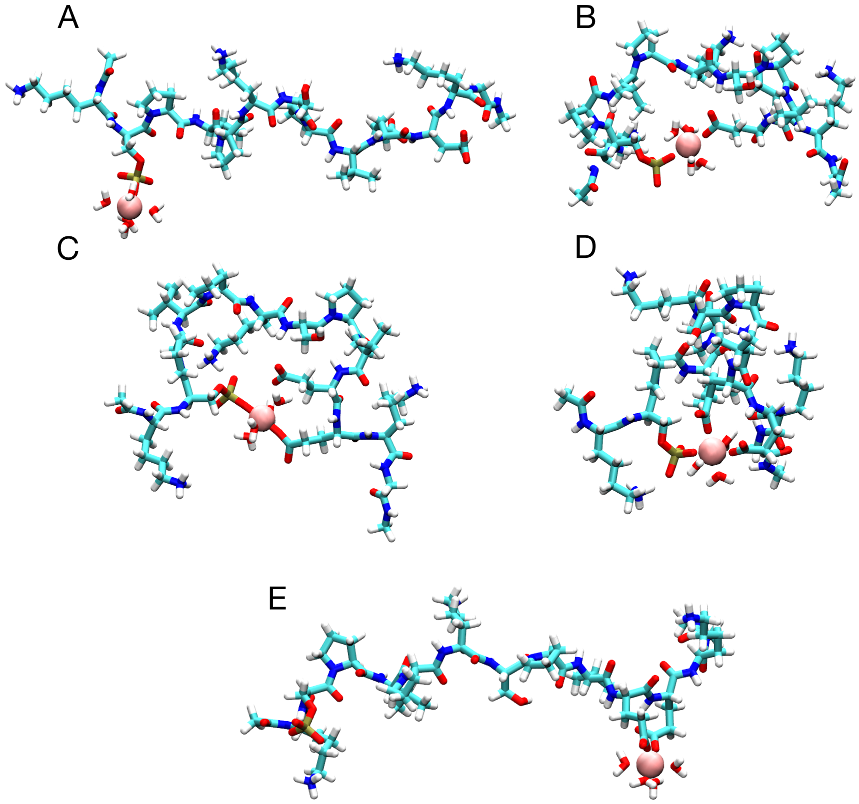

Figure 3: Systems from the S2 group. In this group the serine in the position 2 is phosphorylated, referred in the work by SerP-2. Aluminum can be bound to the mentioned SerP-2 and the glutamic acids in the position 10 and 11, Glu-10 and Glu-11 respectively. A: $NF13Al^{S2}_{S2}$; B: $NF13Al^{S2}_{S2E10}$; C: $NF13Al^{S2}_{S2E11}$; D: $NF13Al^{S2}_{S2E10E11}$; E: $NF13Al^{S2}_{E10E11}$.

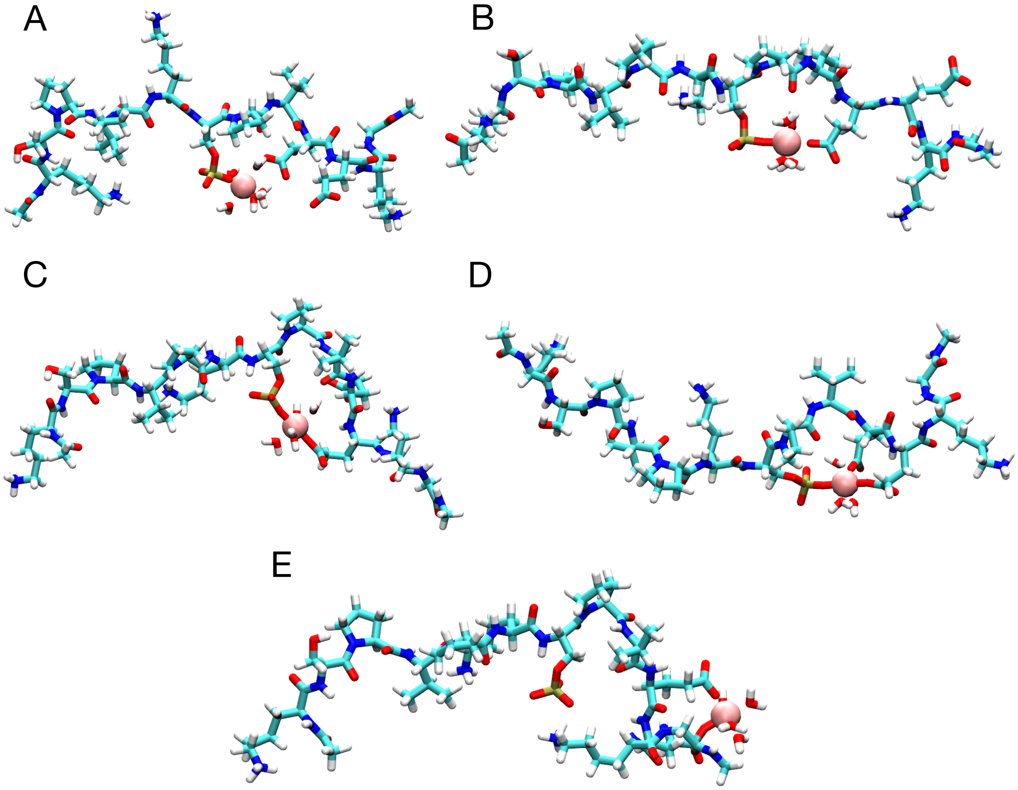

Figure 4: Systems from the S7 group. In this group the serine in the position 7 is phosphorylated, referred in the work by SerP-7. Aluminum can be bound to the mentioned SerP-7 and the glutamic acids in the position 10 and 11, Glu-10 and Glu-11 respectively. A: $NF13Al^{S7}_{S7}$; B: $NF13Al^{S7}_{S7E10}$; C: $NF13Al^{S7}_{S7E11}$; D: $NF13Al^{S7}_{S7E10E11}$; E: $NF13Al^{S7}_{E10E11}$.

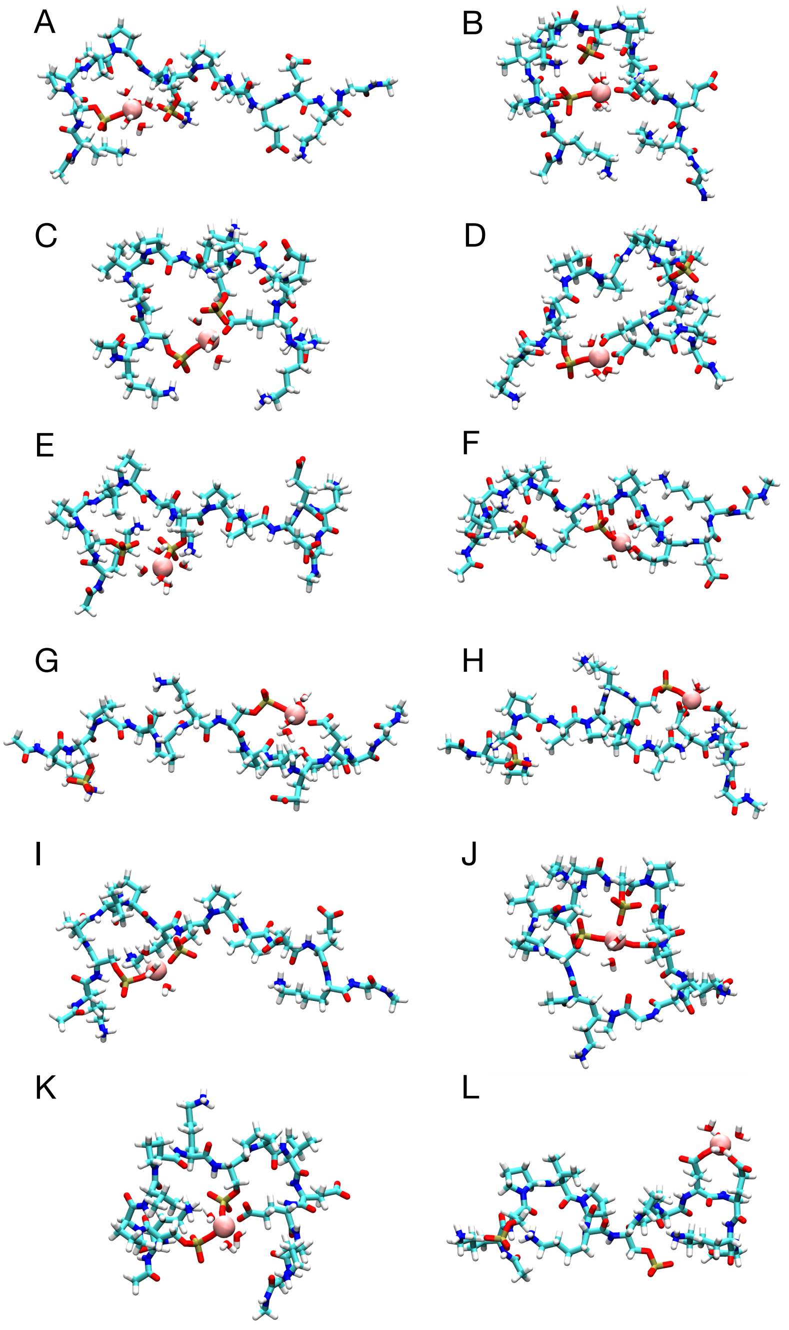

Figure 5: Systems from the S2S7 group. In this group both serines in the position 2 and 7 are phosphorylated. Aluminum can be bound to both SerP-2 and SerP-7 and the glutamic acids in the position 10 and 11, Glu-10 and Glu-11 respectively. A: $NF13Al^{S2S7}_{S2}$; B: $NF13Al^{S2S7}_{S2E10}$; C: $NF13Al^{S2S7}_{S2E11}$; D: $NF13Al^{S2S7}_{S2E10E11}$; E: $NF13Al^{S2S7}_{S7}$; F: $NF13Al^{S2S7}_{S7E10}$; G: $NF13Al^{S2S7}_{S7E11}$; H: $NF13Al^{S2S7}_{S7E10E11}$; I: $NF13Al^{S2S7}_{S2S7}$; J: $NF13Al^{S2S7}_{S2S7E10}$; K: $NF13Al^{S2S7}_{S2S7E11}$; L:$NF13Al^{S2S7}_{E10E11}$.

PARENT Entropy Calculations.

The weirdness of the MIST entropy values calculated with PARENT is their negative sign. A brief explanation is given here but more detailed information can be acquired in these conceptual works V. Hnizdo et al.,[1] D. Suarez et al.[2] work. It is also recommended the works of PARENT authors,.[3,4] PARENT entropy calculation is based on the semi-classical entropy integral,

$$S = -k_B \int dp \int dq \rho(p,q) ln[h^s \rho(p,q)] \tag{1}$$

where the h^s factor ($h = 2\pi\hbar$) is introduced to ensure the dimensionless of the logarithm argument.

This entropy expression can be factorized, upon some mathematical work, into two different contributions: the momentum one (S_m) and the spatial one (S_s). The first term is constant at constant temperature and we are not interested in it because we seek for entropy differences between states (i.e bound vs unbound). We want the second term. However, there is no way to distribute the h^s to the resulting two integrals in a way that these integrals yield physically meaningful units. Therefore, and having in mind that momentum part is constant, the factor is shifted entirely to this part. The other part, the spatial part (S_s), is what is calculated in the PARENT MIST entropy calculations, do not bear physical units and they are often negative. Fortunately, upon substraction (when comparing two systems) and multiplication by the gas constant, they bear correct physical units and are equal to the conventional total entropy change ΔS. This spatial entropy S_s, when referred to the entropy associated with the DOF of a single molecule, is often known as configurational entropy S_{conf}

Simulations were typically long enough to reach convergence for most systems, although occasionally, we had to extend the simulations to 2 μs. An example is shown in Figure 6

In the following Figure 7 the values of S_{MIST} and ΔS_{MIST} without excluding the "hard" bonds and angles DOF are shown. Notice that qualitative trends displayed in the case of "soft" entropy (used in the main text) are kept also in this figure, i.e S2 group entropy loss in greather than S7 group;

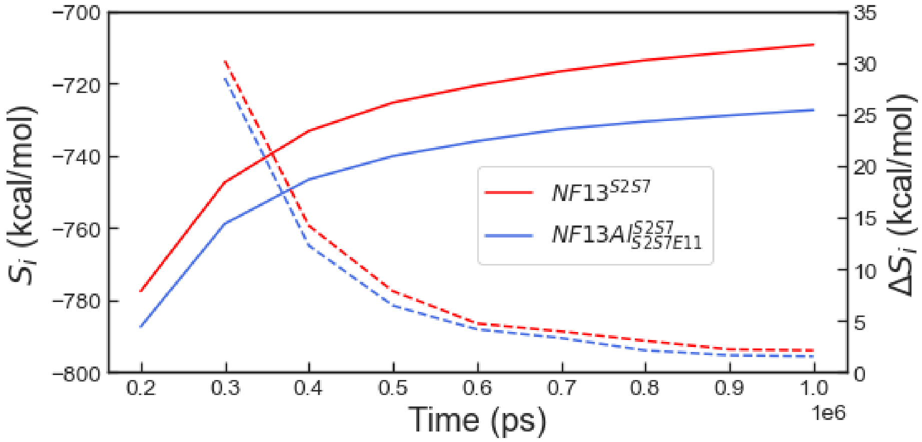

Figure 6: Convergence of the entropy shown for system $NF13Al^{S2S7}_{S2S7E11}$ and its group reference $NF13^{S2S7}$. Dashed line represents the change of entropy for each system.

greatest entropy losses are present in the S2S7 group; tricoordinated systems are the most affected by entropy loss, but the error bars associated (black vertical lines) are much more bigger.

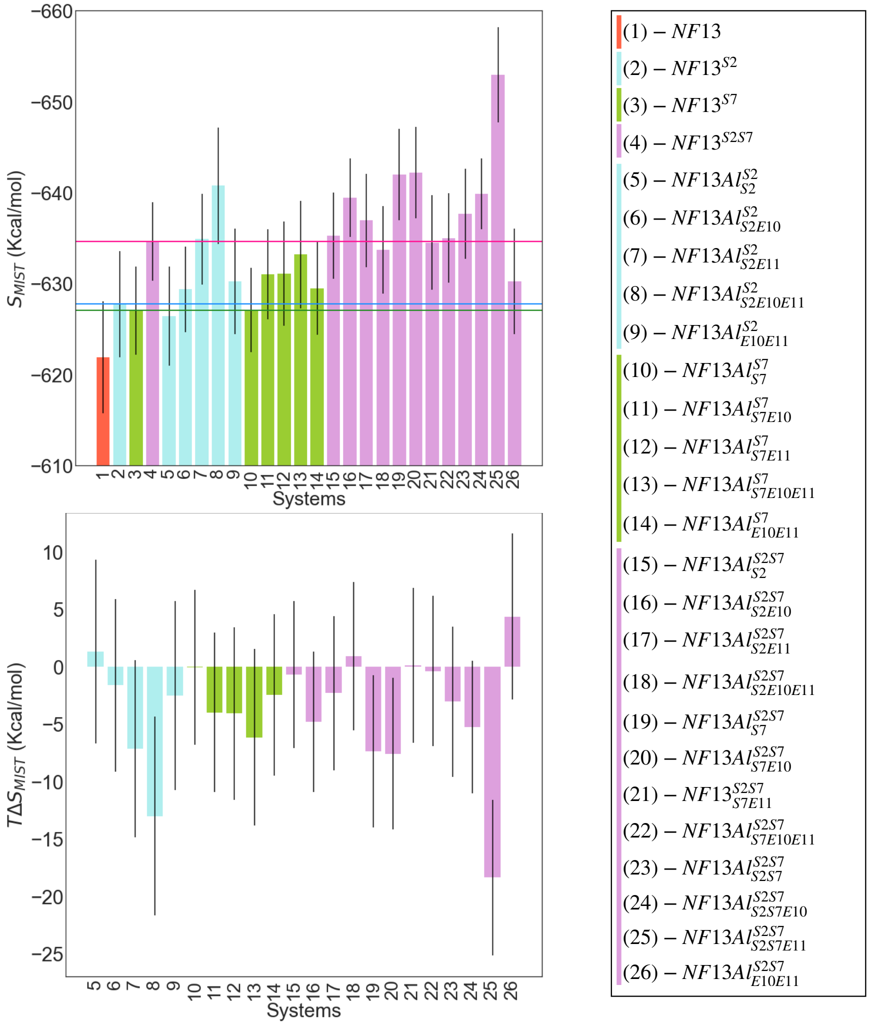

Figure 7: S_{MIST} and ΔS_{MIST} values obtained for all systems with the MIST scheme when considering all the DOF (bonds, angles and torsions). Notice that when the ΔS_{MIST} is calculated, the displayed error is often greater than the actual values, which makes any conclusions meaningless.

Entropy and energy tables

In this section the result of entropy and energy determined for each system using MIST and TURBOMOLE (section 3.4 of main text) are displayed in tables, separated by groups. For each case, the ΔX shows the mean value and the ΔXSD the standard deviation. All values are in kcal/mol. Remember that by the nature of the calculated entropy ΔS, this can not be directly compared to

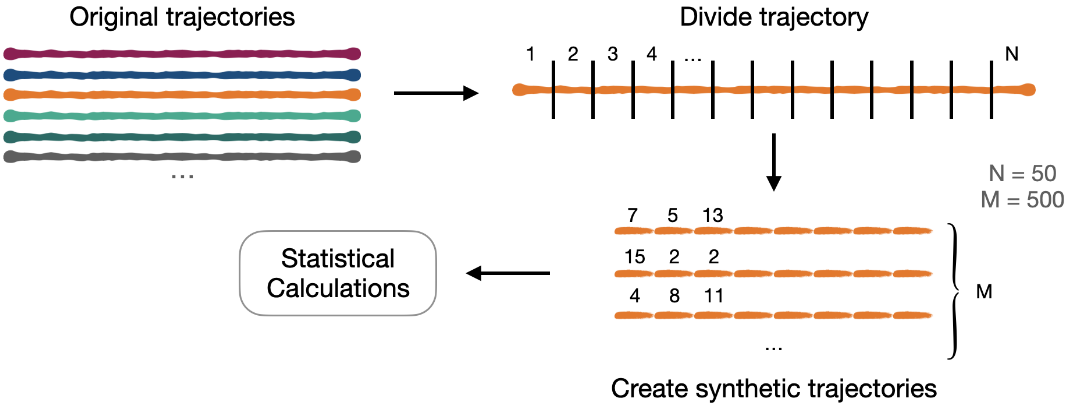

Figure 8: Scheme followed for the bootstrapping carried out for the S_{MIST} determination. Original trajectories are divided into N pieces (N=50), then pieces are arbitrarily combined to form M number of synthetic trajectories (M=500). This allows determination of the mean value and the standard deviation shown in the Figure 9.

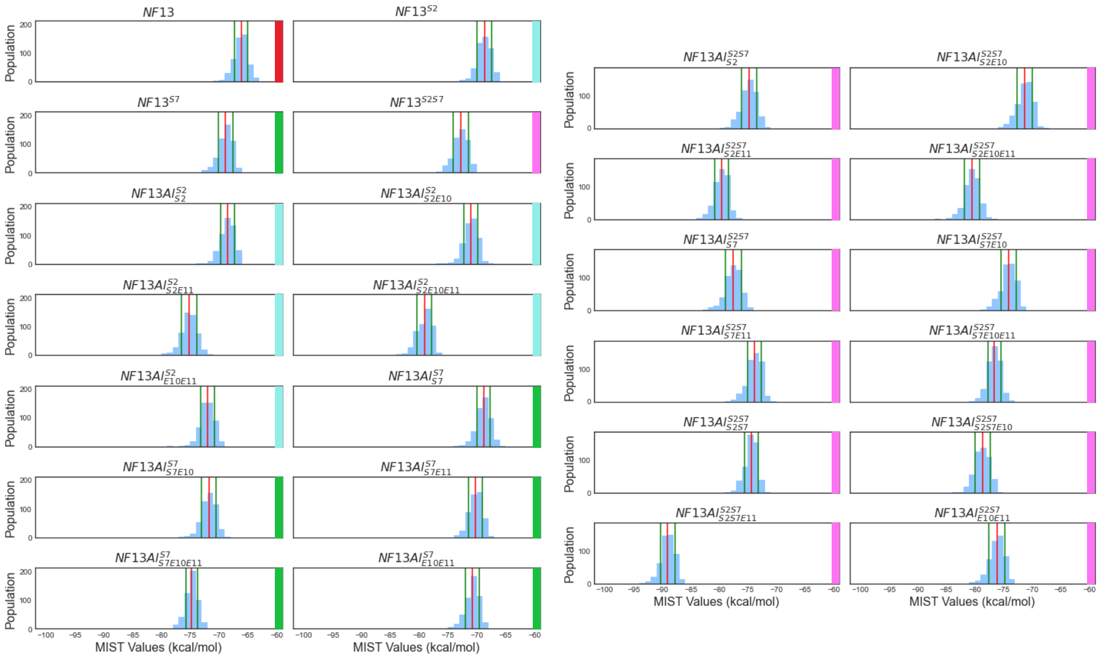

Figure 9: Distributions captured for each system after bootstrapping. Red line in the center corresponds to the mean value; green lines correspond to standard deviations. Color code of groups is represented by a bar in the right-side of the plot of each system.

the stabilization energy. For this reason, the determination of the stabilization energies comes exclusively from single-point energy calculations. Therefore, values are not additive and they should be used only for qualitative conclusions.

Table 2: Entropy and energy values for group S2.

Index	Systems	ΔE (kcal/mol)	ΔE SD (kcal/mol)	ΔS (kcal/mol)	ΔS SD (kcal/mol)
5	$NF13Al^{S2}_{S2}$	-19.5	2.7	**0.2**	1.2
6	$NF13Al^{S2}_{S2E10}$	-39.4	6.2	-2.3	1.2
7	$NF13Al^{S2}_{S2E11}$	-42.0	6.8	-6.5	1.3
8	$NF13Al^{S2}_{S2E10E11}$	**-71.6**	15.8	-10.4	1.3
9	$NF13Al^{S2}_{E10E11}$	-70.5	10.5	-3.3	1.2

Table 3: Entropy and energy values for group S7.

Index	Systems	ΔE (kcal/mol)	ΔE SD (kcal/mol)	ΔS (kcal/mol)	ΔS SD (kcal/mol)
10	$NF13Al^{S7}_{S7}$	-15.9	2.3	**0.1**	1.2
11	$NF13Al^{S7}_{S7E10}$	-51.2	6.4	-2.9	1.3
12	$NF13Al^{S7}_{S7E11}$	-52.5	6.3	-1.4	1.2
13	$NF13Al^{S7}_{S7E10E11}$	**-97.6**	12.0	-5.9	1.0
14	$NF13Al^{S7}_{E10E11}$	-65.7	12.3	-1.9	1.2

Table 4: Entropy and energy values for group S2S7.

Index	Systems	ΔE (kcal/mol)	ΔE SD (kcal/mol)	ΔS (kcal/mol)	ΔS SD (kcal/mol)
15	$NF13Al^{S2S7}_{S2}$	-49.3	6.4	-2.0	1.3
16	$NF13Al^{S2S7}_{S2E10}$	-82.1	14.6	**1.5**	1.4
17	$NF13Al^{S2S7}_{S2E11}$	-62.8	13.2	-6.8	1.2
18	$NF13Al^{S2S7}_{S2E10E11}$	-108.4	34.4	-7.7	1.3
19	$NF13Al^{S2S7}_{S7}$	-50.3	6.9	-4.7	1.4
20	$NF13Al^{S2S7}_{S7E10}$	-62.6	5.4	-1.3	1.3
21	$NF13Al^{S2S7}_{S7E11}$	-57.9	7.2	-1.1	1.3
22	$NF13Al^{S2S7}_{S7E10E11}$	-109.1	11.3	-3.8	1.1
23	$NF13Al^{S2S7}_{S2S7}$	-53.6	8.9	-1.60	1.2
24	$NF13Al^{S2S7}_{S2S7E10}$	**-128.3**	27.3	-5.9	1.3
25	$NF13Al^{S2S7}_{S2S7E11}$	-99.9	31.8	-16.2	1.3
26	$NF13Al^{S2S7}_{E10E11}$	-50.9	5.4	-3.4	1.4

MD Analysis

Ramachandran density plots

Ramachandran density plots were computed using the MDtraj Python package. Figure 10 shows the plot for the $NF13$ reference peptide (without phosphorylations and metalic ion) with a red bar on top. The other 3 reference systems (without metal but phosphorylated) are shown inside each group, in order to facilitate the analysis. All references appear in black. Colors show the different number of coordination and types of residues.

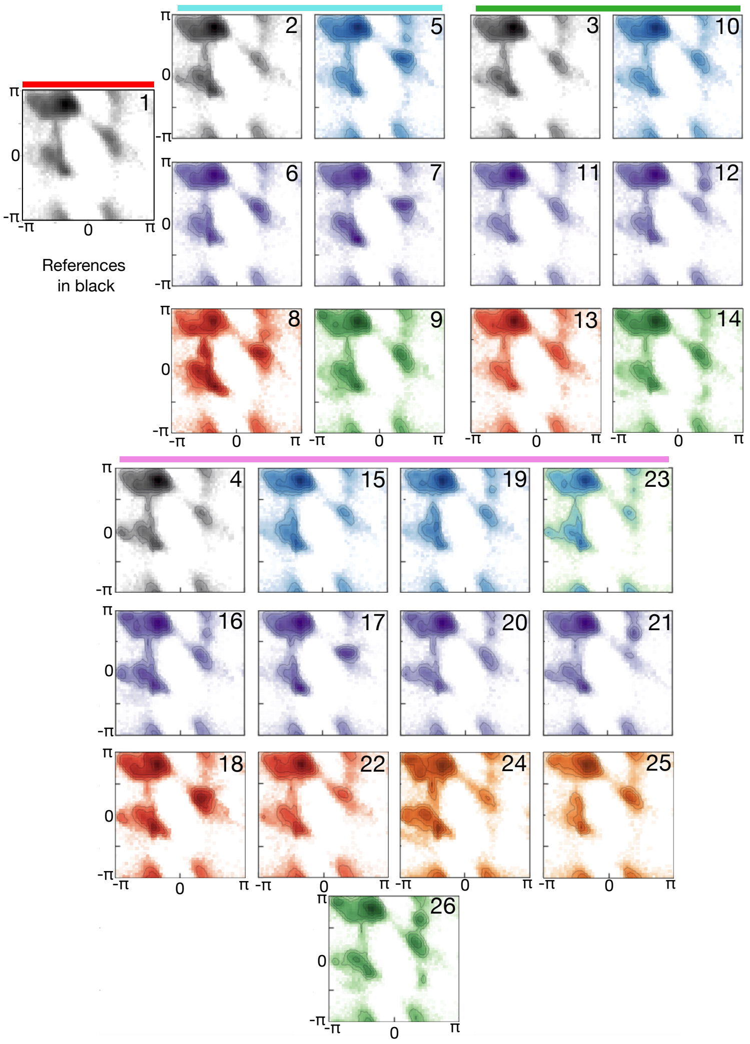

Figure 10: Ramachandran plots for all systems organized in groups. S2 group is displayed under the cyan bar, S7 group under the green bar and S2S7 group under the magenta bar. References are shown in black.

Ramachandran based SS Fraction

Figure 11: A) Representation of the regions defined to assign secondary structure fraction from the Ramachandran plots. Black dots represent the most populated point of the regions, expected to concentrate the greatest density. The rest of the plots show the fraction of SS assigned for each residue on each system. References are shown in the top part, with same color code as Figure 10 for each group (red for $NF13$).

Contact Maps

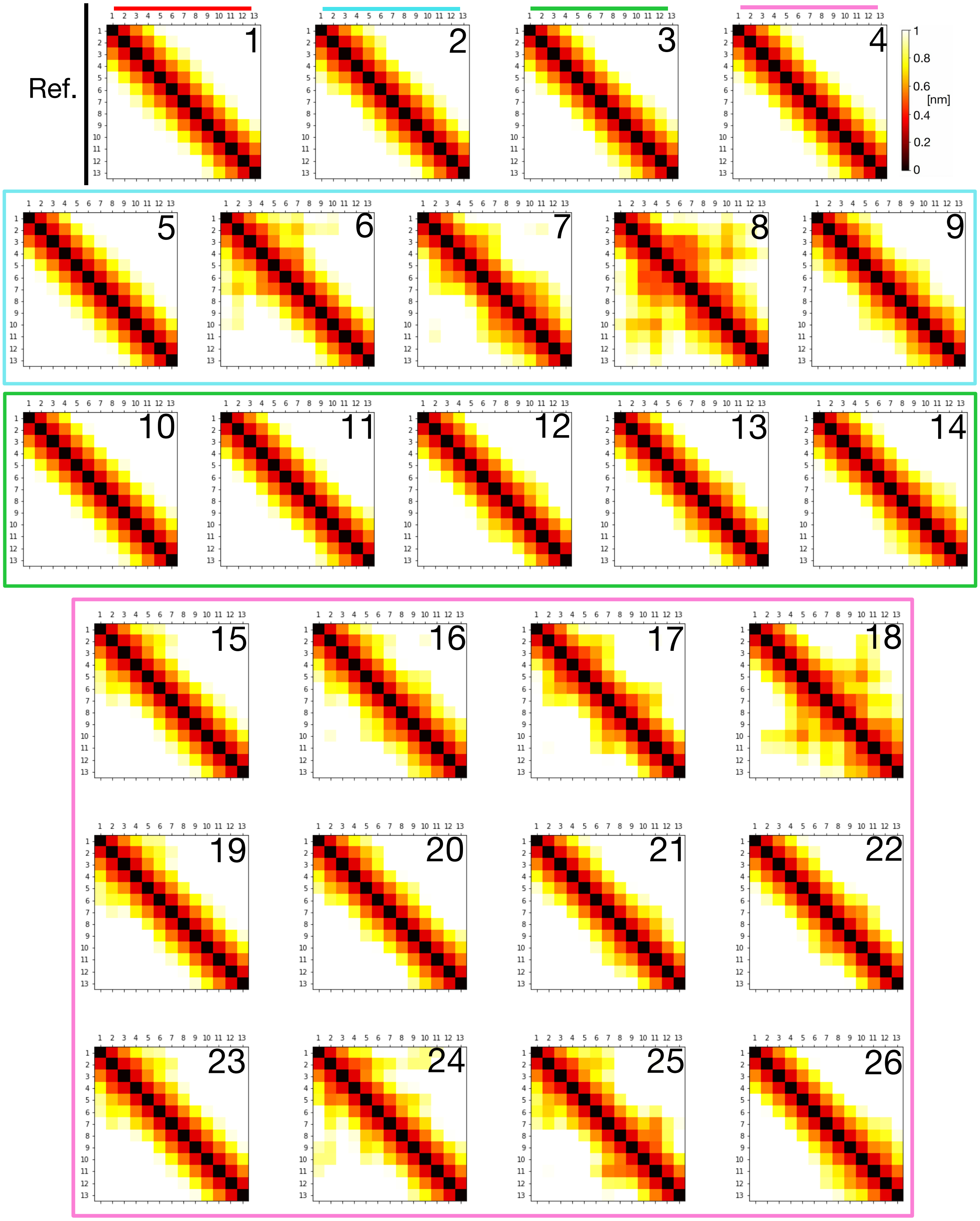

Figure 12: Contact maps of all systems by group. Top plots are for reference systems. The contacts between residues are shown in a range of 0 to 1 nm. Color code as of Figure 10.

DSSP Analysis

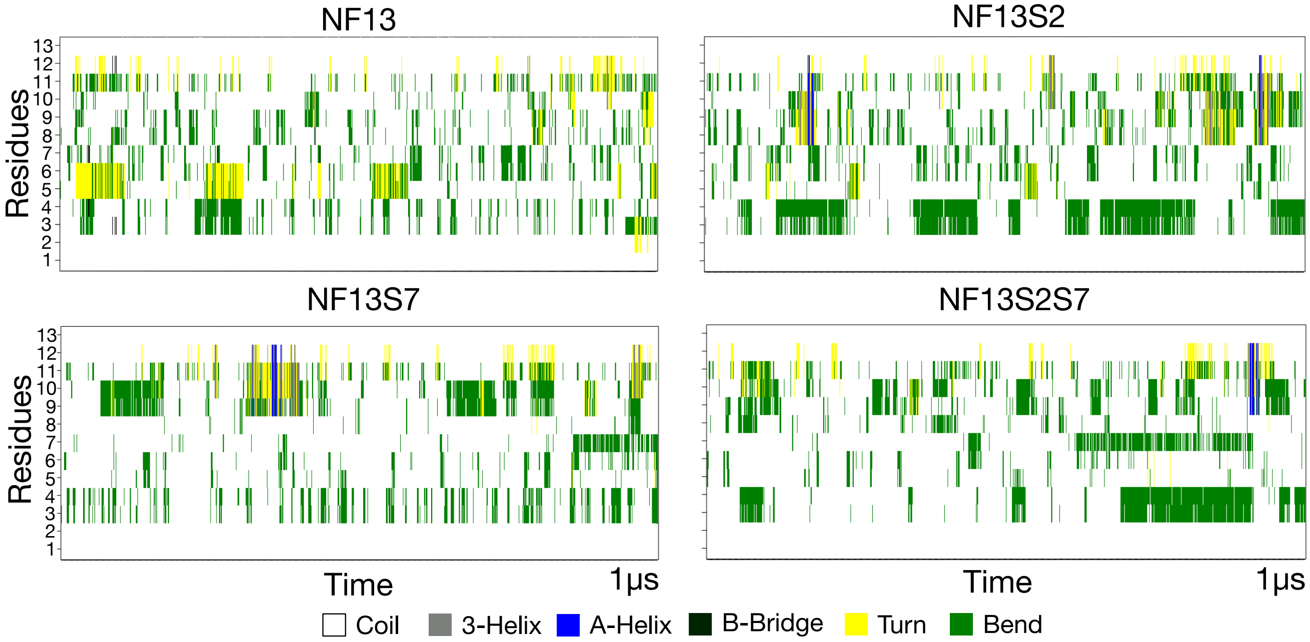

Figure 13: DSSP analysis for the reference systems.

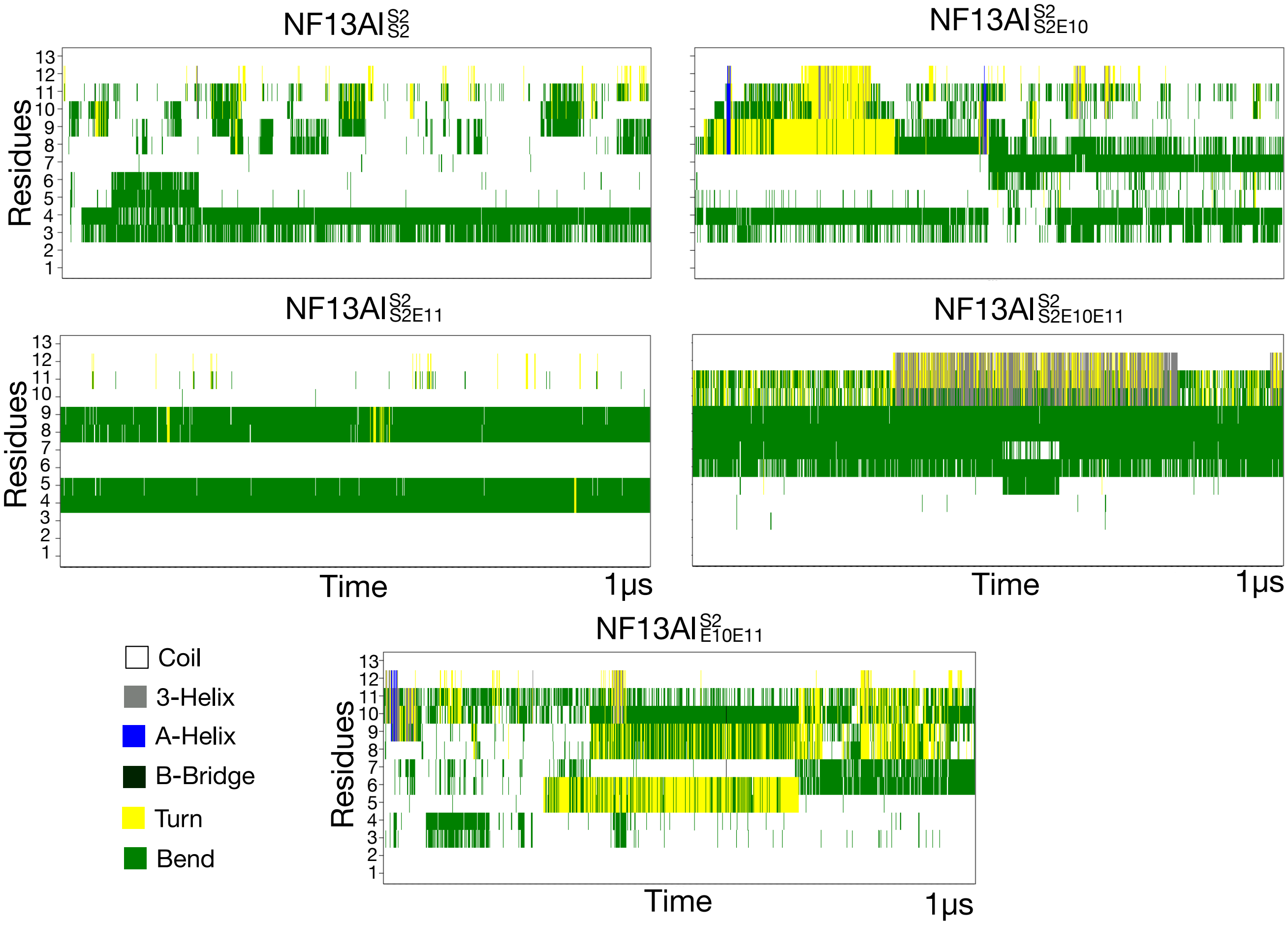

Figure 14: DSSP analysis for the S2 group of systems.

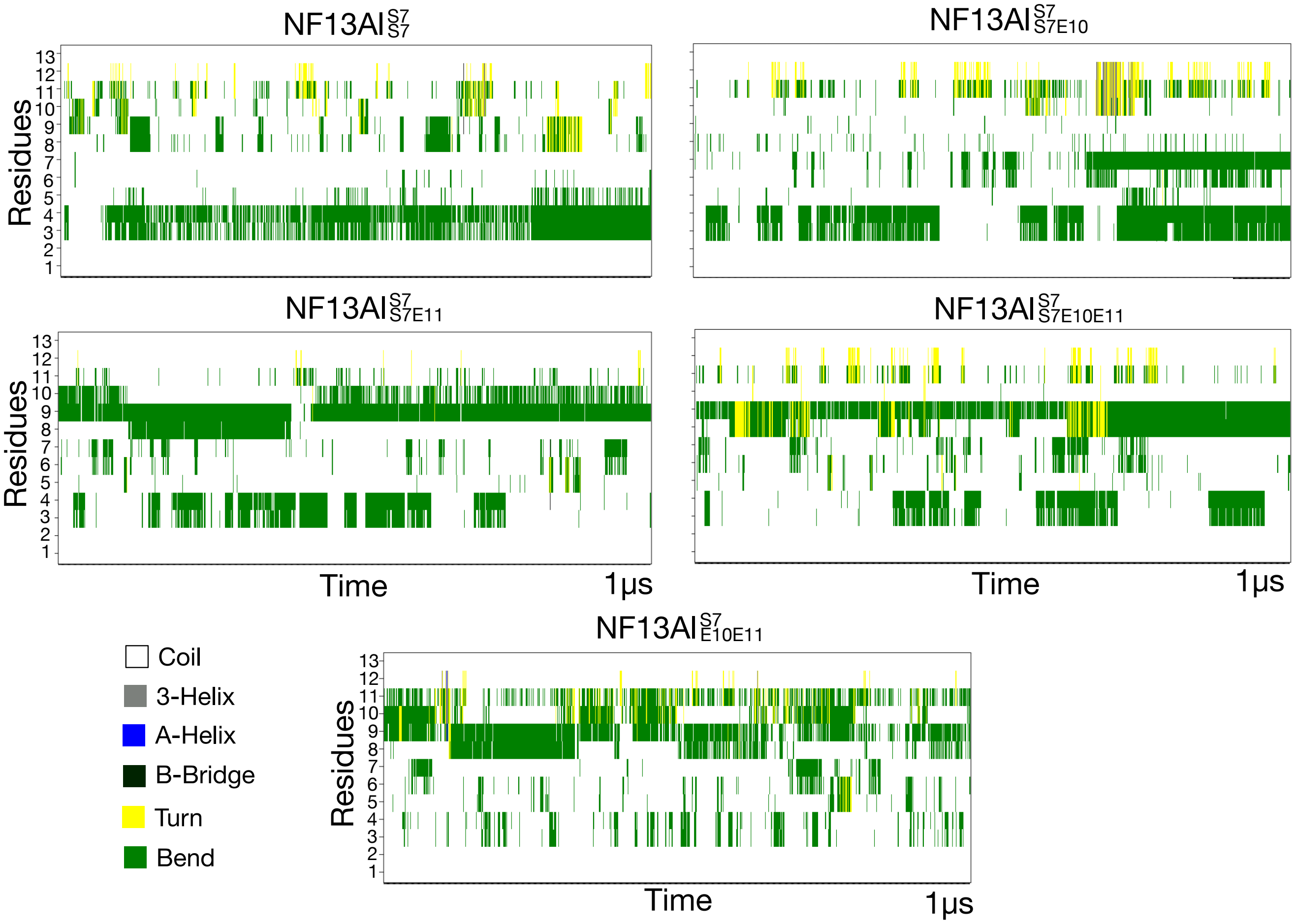

Figure 15: DSSP analysis for the S7 group of systems.

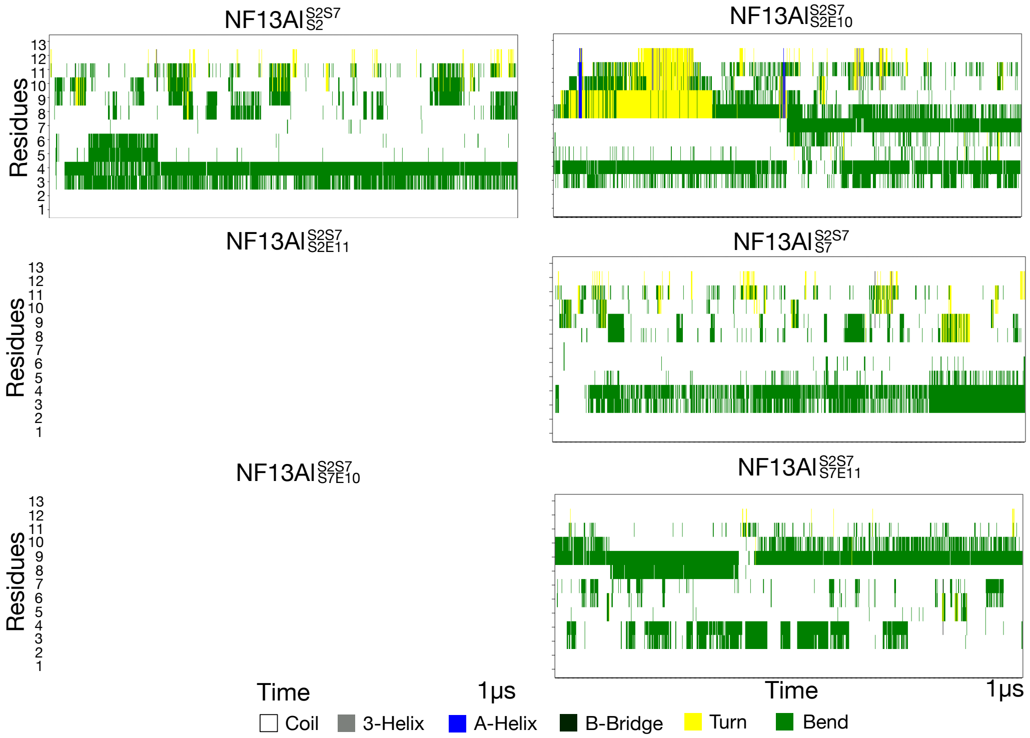

Figure 16: DSSP analysis for the S2S7 group of systems

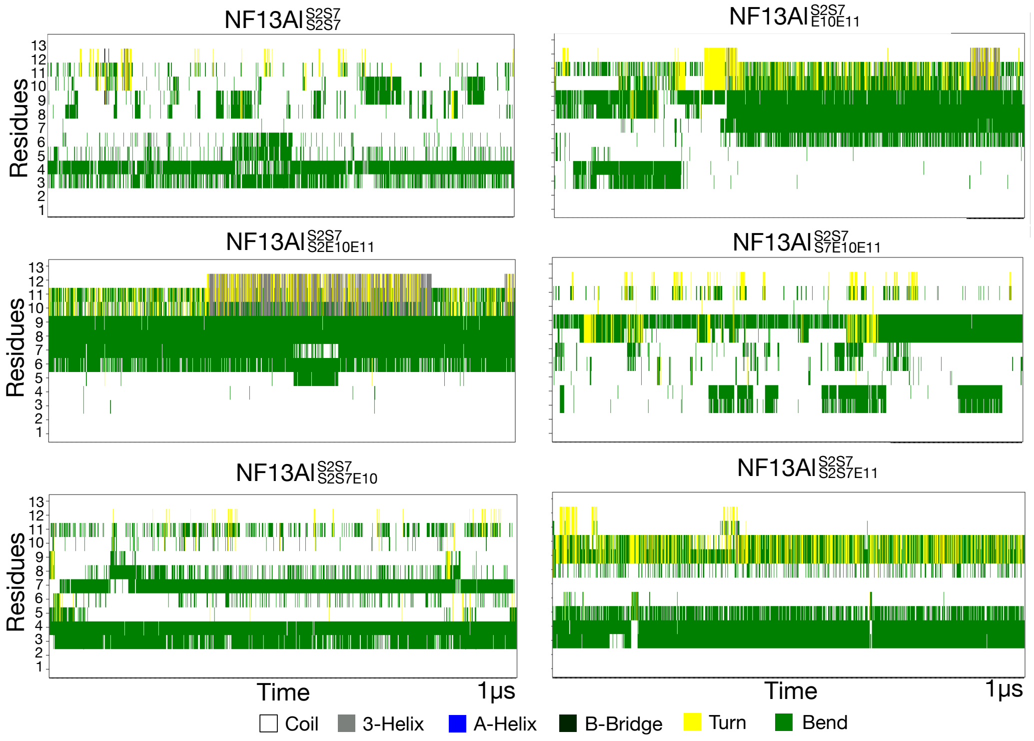

Figure 17: DSSP analysis for the S2S7 group of systems.

Radius of Gyration

Table 5: Mean values of the radius of gyration (Rg) for systems. Values are calculated from the figures below.

Index	Systems	Rg(nm)
1	$NF13$	1.05
2	$NF13^{S2}$	1.04
3	$NF13^{S7}$	1.08
4	$NF13^{S2S7}$	1.06
5	$NF13Al^{S2}_{S2}$	1.06
6	$NF13Al^{S2}_{S2E10}$	0.88
7	$NF13Al^{S2}_{S2E11}$	0.84
8	$NF13Al^{S2}_{S2E10E11}$	0.74
9	$NF13Al^{S2}_{E10E11}$	0.88
10	$NF13Al^{S7}_{S7}$	1.11
11	$NF13Al^{S7}_{S7E10}$	1.13
12	$NF13Al^{S7}_{S7E11}$	1.05
13	$NF13Al^{S7}_{S7E10E11}$	1.10
14	$NF13Al^{S7}_{E10E11}$	1.03
15	$NF13Al^{S2S7}_{S2}$	0.98
16	$NF13Al^{S2S7}_{S2E10}$	0.84
17	$NF13Al^{S2S7}_{S2E11}$	0.87
18	$NF13Al^{S2S7}_{S2E10E11}$	0.82
19	$NF13Al^{S2S7}_{S7}$	0.99
20	$NF13Al^{S2S7}_{S7E10}$	1.01
21	$NF13Al^{S2S7}_{S7E11}$	1.03
22	$NF13Al^{S2S7}_{S7E10E11}$	1.04
23	$NF13Al^{S2S7}_{S2S7}$	1.00
24	$NF13Al^{S2S7}_{S2S7E10}$	0.82
25	$NF13Al^{S2S7}_{S2S7E11}$	0.80
26	$NF13Al^{S2S7}_{E10E11}$	1.03

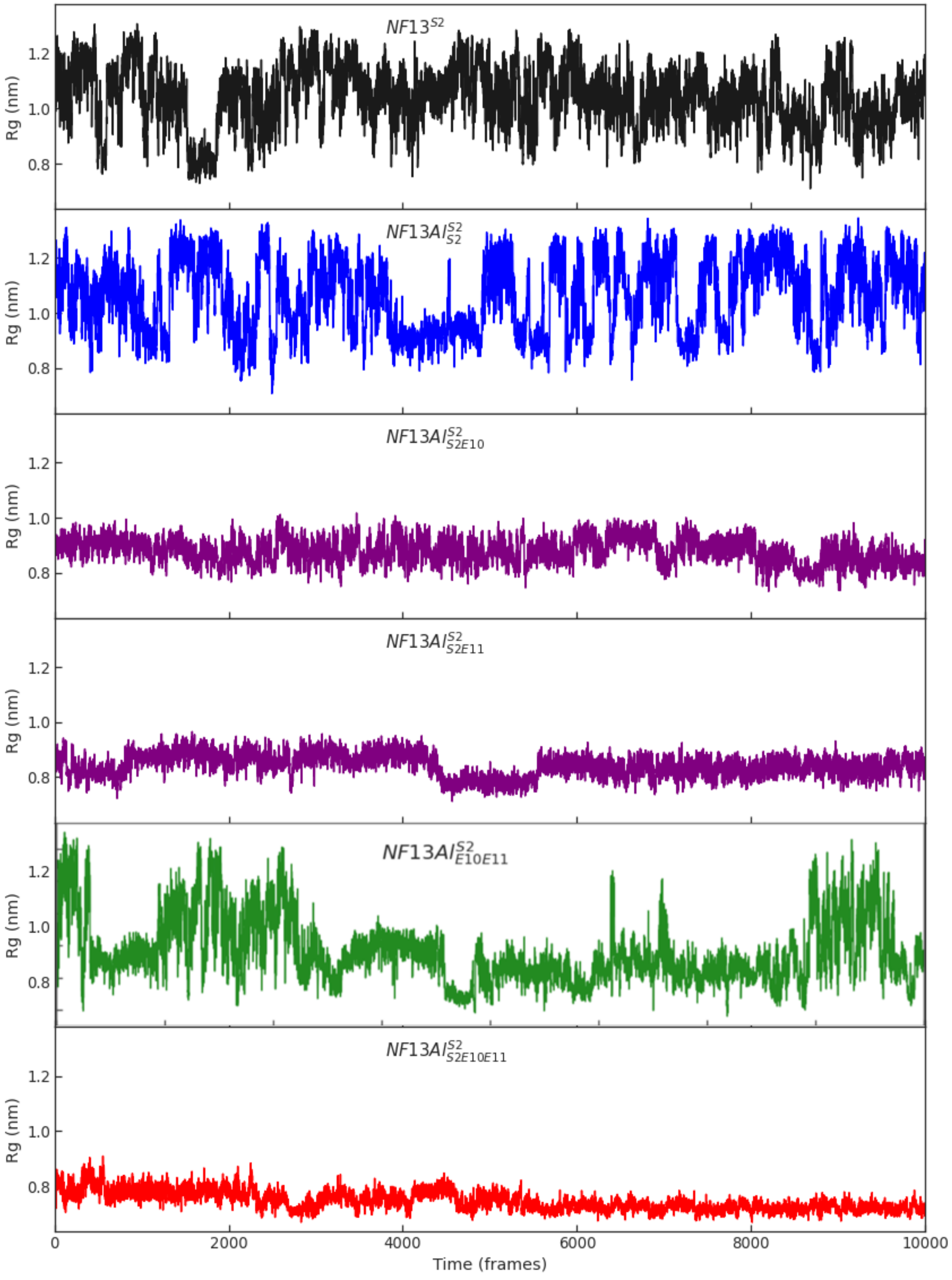

Figure 18: Radius of Gyration for S2 group. Reference system of the group displayed in black.

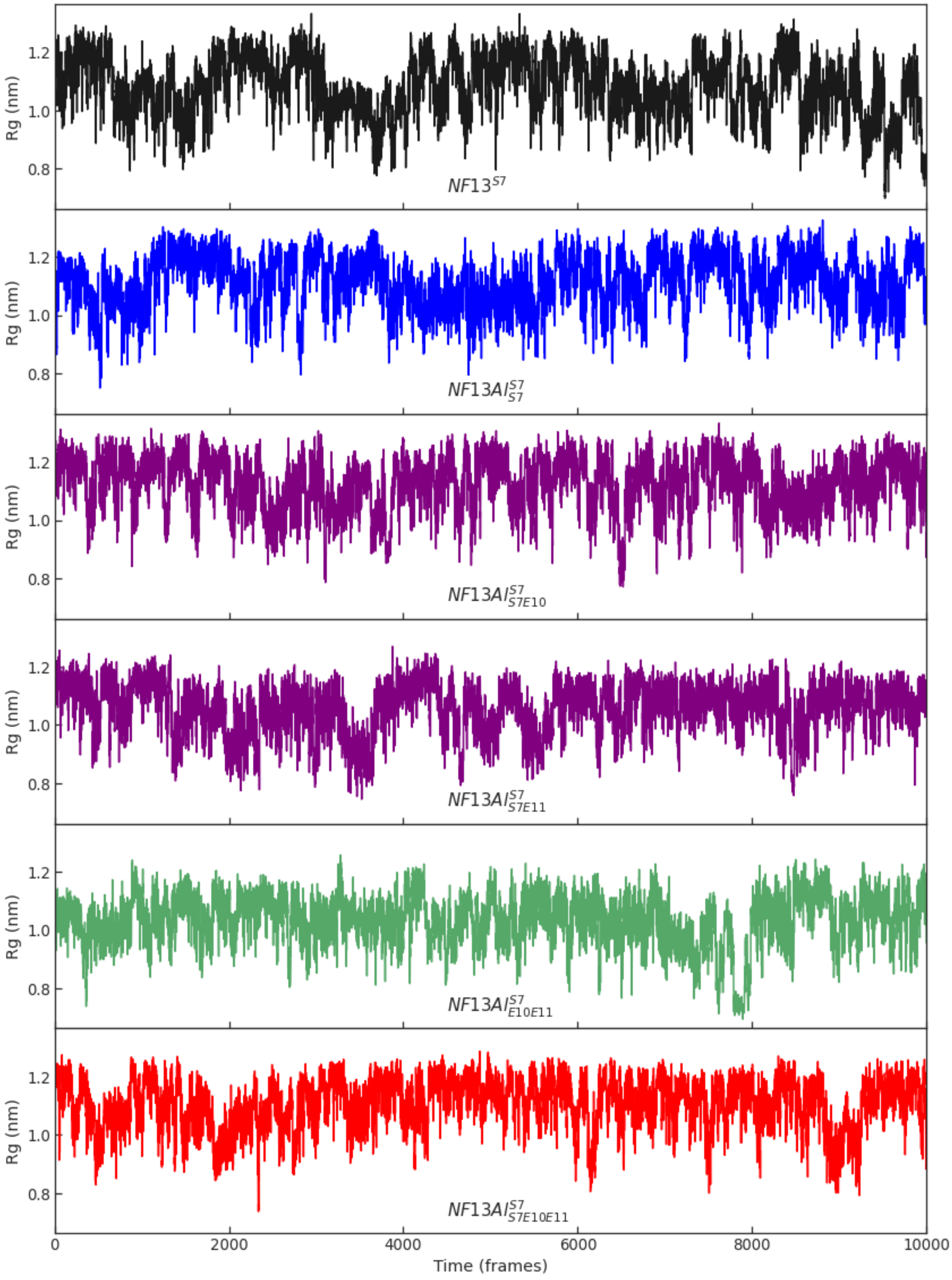

Figure 19: Radius of Gyration for S7 group. Reference system of the group displayed in black.

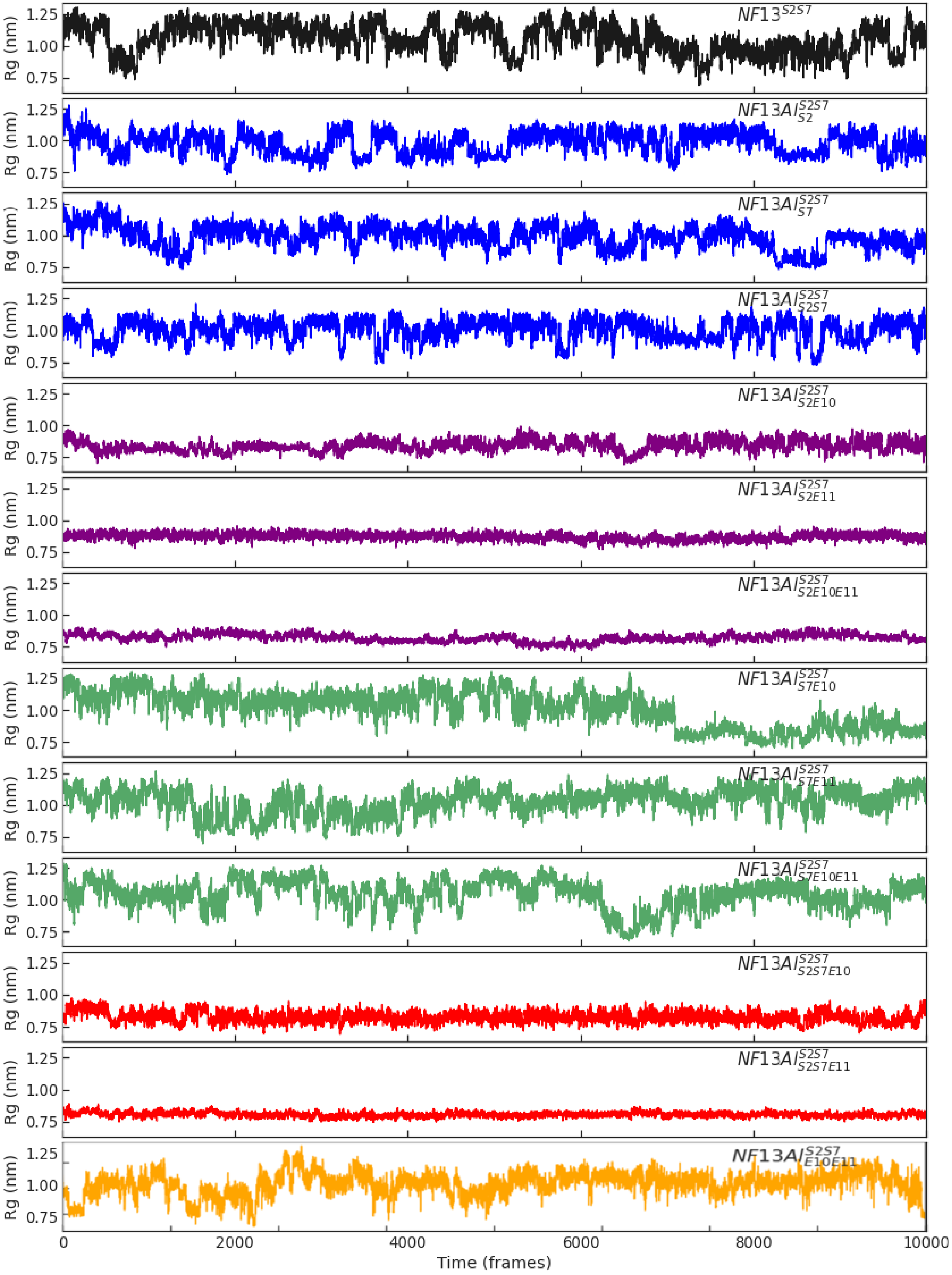

Figure 20: Radius of Gyration for S2S7 group. Reference system of the group displayed in black.

Clustering Pie Charts

Results of the PCA are shown in Figure 21. The resultant energy of each system corresponds to the mean value of the energy distribution formed by the clusters explored by each system during its trajectory.

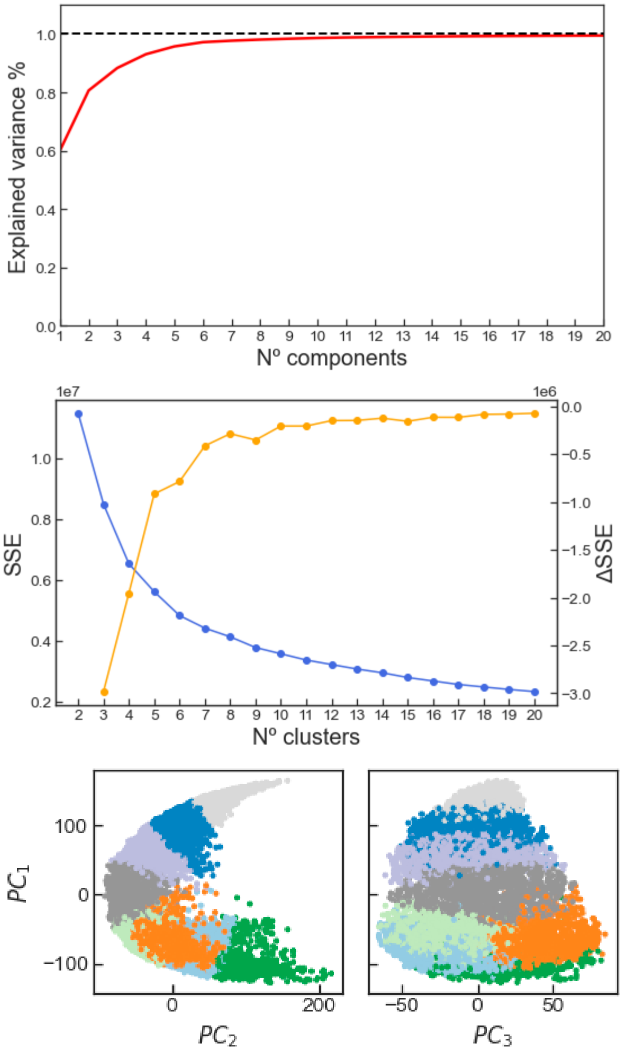

Figure 21: (a) Explanation of the variance of the data set by the number of components, (b) fit of data with the number of components using the SSE measure, (c) PC1 vs PC2 and PC3 for 8 clusters.

References

(1) Hnizdo, V.; Gilson, M. K. Thermodynamic and Differential Entropy under a Change of Variables. *Entropy* **2010**, *12*, 578–590.

(2) Suárez, D.; Díaz, N. Direct methods for computing single-molecule entropies from molecular simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2015**, *5*, 1–26.

(3) Fleck, M.; Polyansky, A. A.; Zagrovic, B. PARENT: A Parallel Software Suite for the Calculation of Configurational Entropy in Biomolecular Systems. *J. Chem. Theory Comput.* **2016**, *12*, 2055–2065.

(4) Fleck, M.; Zagrovic, B. Configurational Entropy Components and Their Contribution to Biomolecular Complex Formation. *J. Chem. Theory Comput.* **2019**, *15*, 3844–3853.