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

Electronic Supporting Information

Analysing the microenvironment of 2-*p*-toluidinylnaphthalene-6-sulfonate (TNS) in solvents and in different conformational states of proteins with relation to its fluorescence properties: a computational study

Neshatul Haque, Krishnakanth Baratam, N. Prakash Prabhu*

**E-mail: nppsl@uohyd.ernet.in*

APPENDIX

Electronic Supporting Information

Appendix A

List of all the simulations

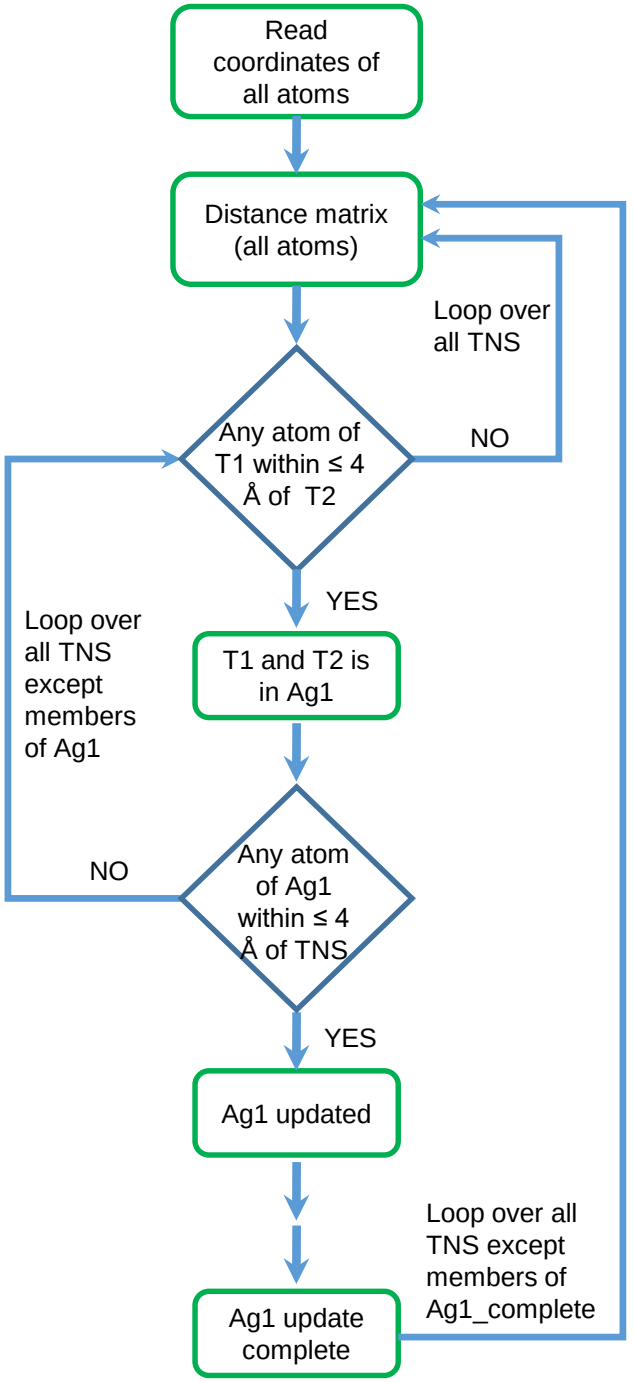
Name/ Label	Protein	Solvent	Number of TNS	Simulation Time (ns)
BSA	+	Water	-	50
A0	+	Water	-	50
A2	+	Water	-	50
A3	+	Water	-	50
R0	+	Water	-	50
R2	+	Water	-	50
R3	+	Water	-	50
BSA	+	Water	50	100
A0	+	Water	50	100
A2	+	Water	50	100
A3	+	Water	50	100
R0	+	Water	50	100
R2	+	Water	50	100
R3	+	Water	50	100
Water	-	Water	64	100
Ethanol	-	Ethanol	64	100
DMF	-	DMF	64	100
DMSO	-	DMSO	64	100
			Total Simulation time	1450 ns (1.45 μs)

Appendix B

Characterization of aggregates

Aggregates were calculated using the distance coordinates between every atom in a particular time frame during simulation.

1. A distance matrix consisting of all the atoms was created.
2. All the atoms of each TNS (T1) was compared against all the atoms of another TNS (T2) molecule. If the distance was ≤ 4 Å, the molecules were considered to be forming an aggregate (Ag1).
3. Then, Ag1 was checked against all the atoms of another TNS, excluding T1 and T2. If the distance condition (≤ 4 Å) was met, then Ag1 was updated.
4. If Ag1 cannot be updated anymore which means that the aggregate is complete.
5. The steps 2 and 3 were repeatedly performed on all TNS molecules, excluding the members of Ag1. The new aggregates were numbered as Ag2, Ag3, and so on.
6. In brief, TNS molecules within 4 Å to each other were assumed to be forming aggregates. The number of aggregates were calculated for each frame of the simulation and averaged out.



Appendix C

Number of Hydrogen bonds:

- Last 10 ns data was used for the hydrogen bond analysis.
- The hydrogen bonds were calculated with the help of the g_hbond tool in GROMACS.
    - The OH and NH groups were considered as donors. Atoms N and O were considered as acceptors.
    - Bond formation was considered, if the distance between the donor and acceptor was ≤ 3.5 Å and the angle between hydrogen -donor - acceptor was ≤30º
- All the hydrogen bonds formed by each TNS with protein, other TNS molecules and solvents were obtained.
- In house R-script was written to sort different H-bonds and to count them.
- Care was taken not to over-count the hydrogen bond number, e.g all the three oxygen of TNS showing hydrogen bond with an amino acid was counted as one.

Appendix D

Order of contacts and amino acid specificity

- Order of contacts was calculated for last 10 ns at an interval of every100 ps (i.e., for 101 frames).
- For each TNS, list of atoms in contact with other components of the system (protein, solvent and other TNS molecules) was obtained using Gromacs tool for every chosen time frame.

  If the distance between the atoms was ≤4 Å, they were assumed to be in contact.
- The file consisted of all the possible interactions of all the TNS molecules. In-house R-code was written to dissect each type of interactions, as given below.
- For 101 frames, with 50 TNS molecules (or 64 TNS in the cases of solvent-only simulations) in each frame, the interactions were classified into TNS-protein, TNS-TNS and TNS-solvent, and the number of interactions were counted.
- The number of each interactions were divided by 101 to calculate the average of interactions/frame.
- These values were further divided by 50 (or 64 in the cases of solvent-only simulations) to evaluate the fraction of TNS molecules interacting with protein or with solvent or with other TNS molecules.
- The number of TNS-protein contacts was in the range of 1 to 50 and TNS-TNS & TNS-solvent contacts were in the range of 1 to 90
- For TNS-Protein contacts, three categories were assigned such that zero, lower and higher order for 0, 1-20 and above 21 contacts, respectively.
- For TNS-TNS and TNS-solvents contacts, three categories were assigned such that zero, lower and higher order for 0, 1-50 and above 51 contacts, respectively.
- From TNS-Protein contact information, the individual amino acids interacting with TNS were counted. The amino acid-TNS contacts were classified into TNS-side chain and TNS-main-chain contacts based on atom labels.
- The counted contacts were normalized with respect to the number of each amino acid residues present in the protein.

FIGURES

Electronic Supporting Information

Figure S1

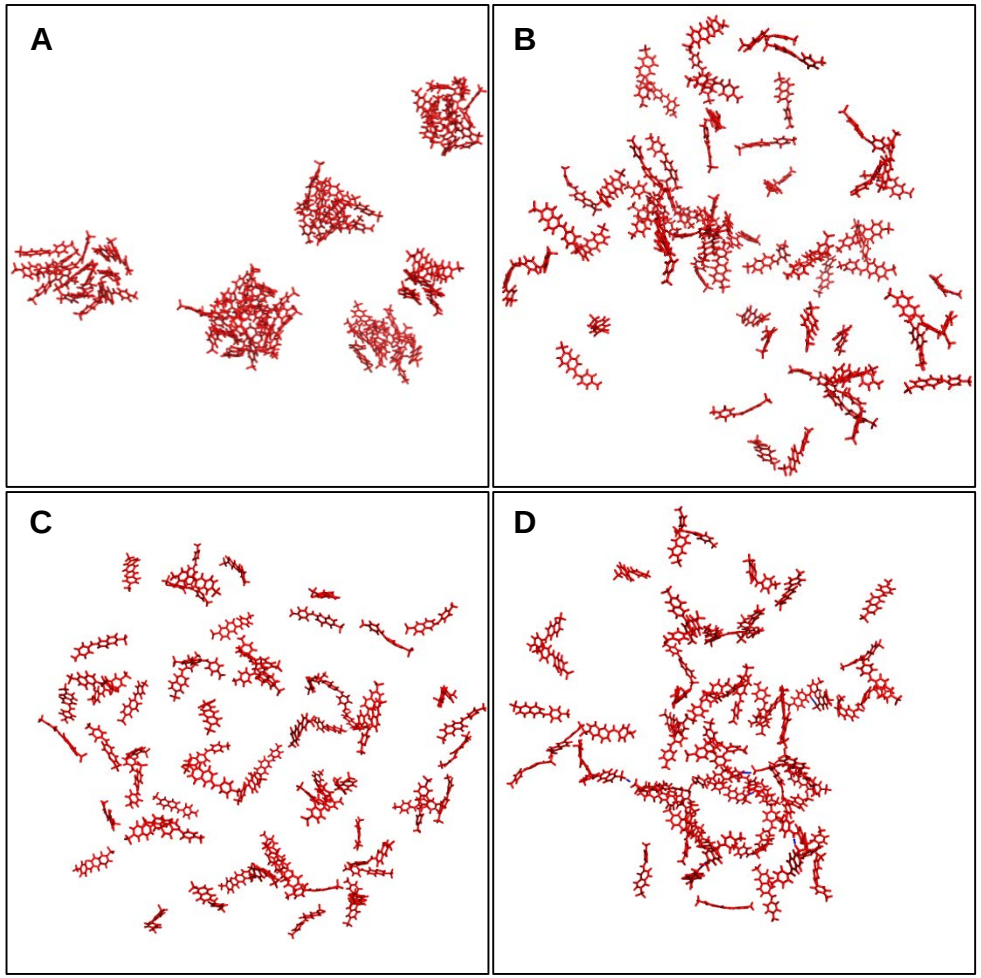


Fig. S1 Snapshots of positions of 64 TNS molecules in (A) water, (B) ethanol, (C) DMF and (D) DMSO obtained at the end of 100 ns simulation

Figure S2

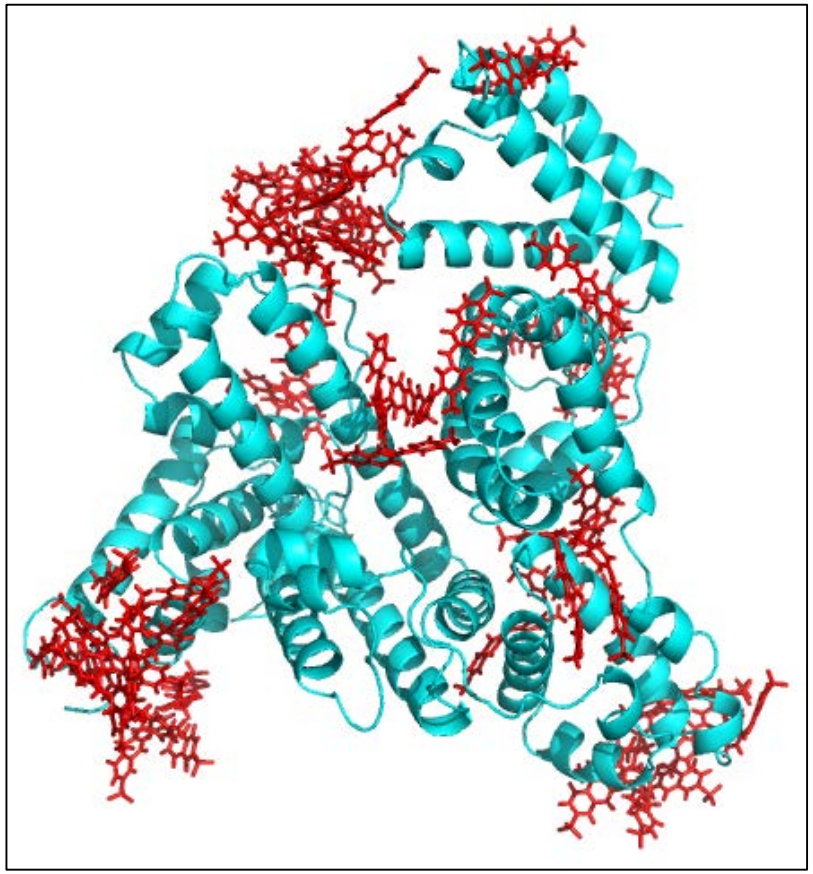


Fig. S2 Snapshot of position of 50 TNS molecules in the presence of BSA obtained at the end of 100 ns simulation. Cartoon diagram represents the protein and stick models represent TNS.

Figure S3

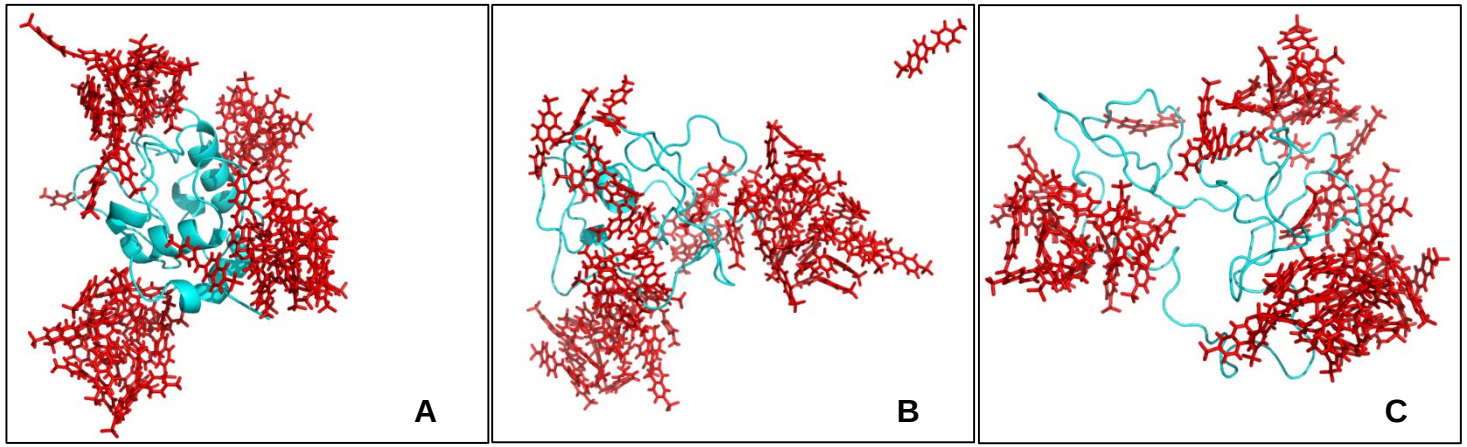


Fig. S3 Snapshots of positions of 50 TNS molecules in the presence of (A) native and (B and C) two-denatured conformations of α-LA obtained at the end of 100 ns simulation. Cartoon diagrams represent the protein and stick models represent TNS.

Figure S4

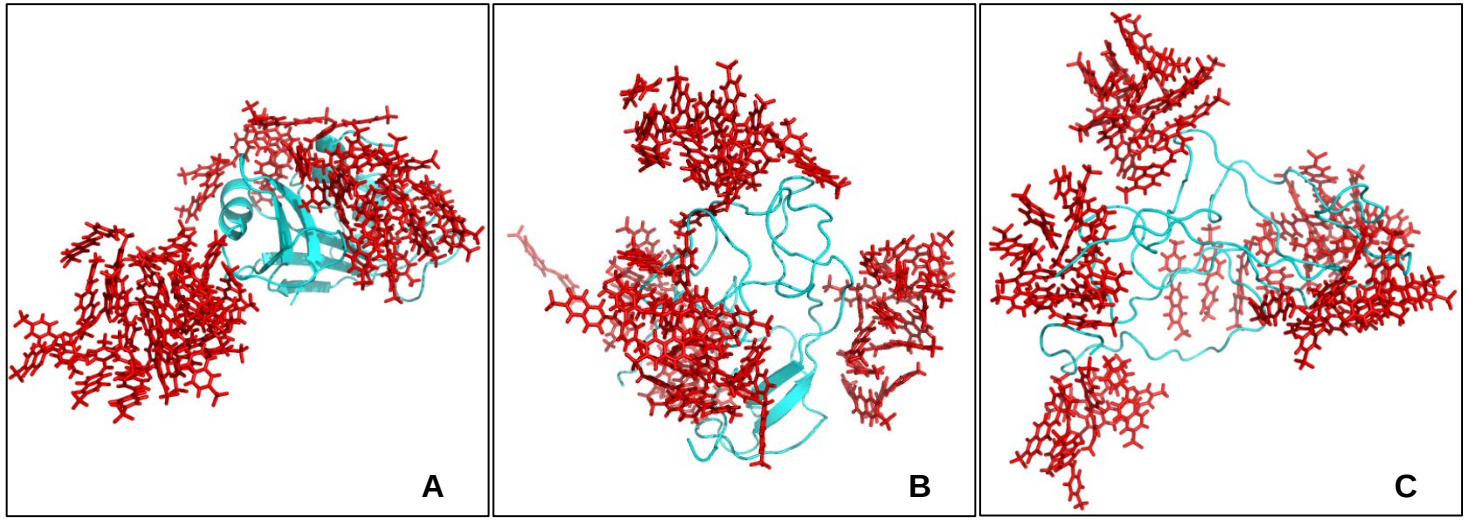


Fig. S4 Snapshots of positions of 50 TNS molecules in the presence of (A) native and (B and C) two-denatured conformations of RNase A obtained at the end of 100 ns simulation. Cartoon diagrams represent the protein and stick models represent TNS.

Figure S5

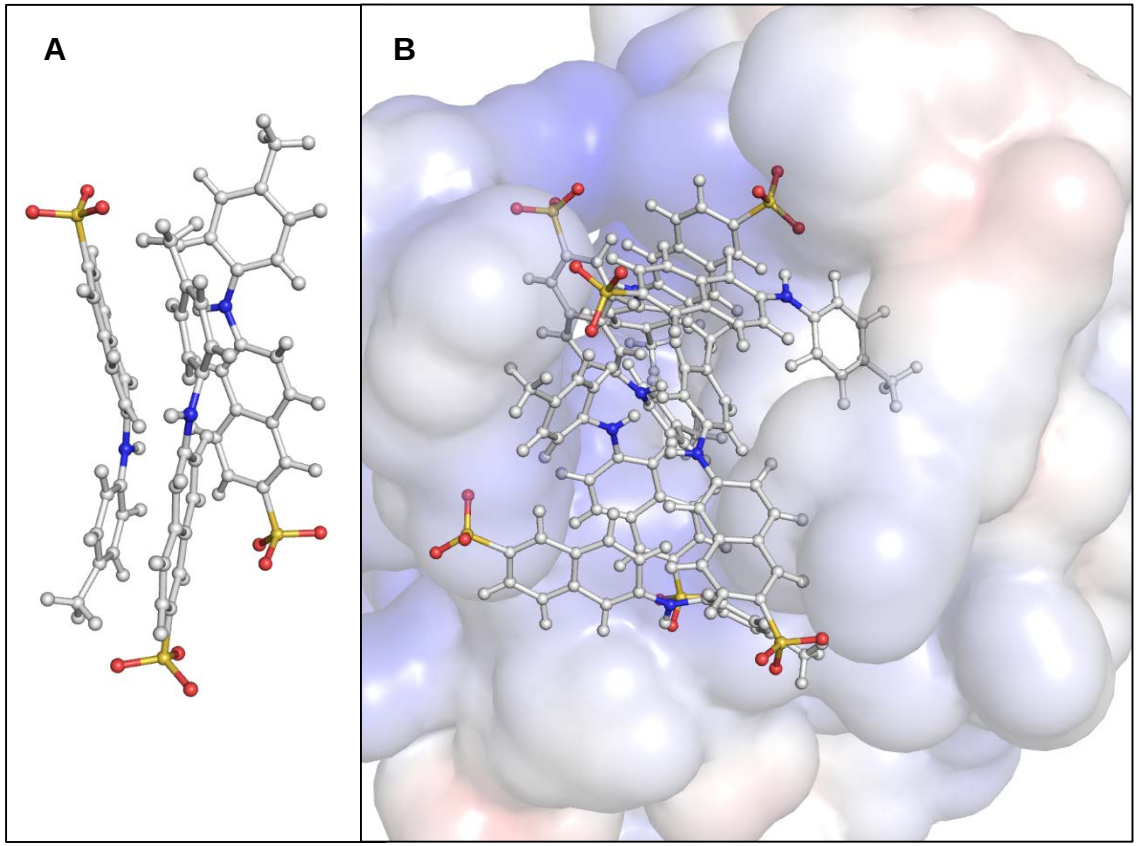


Fig. S5 Representative snapshots of TNS aggregates formed in (A) water and (B) on the surface of protein R2.

Figure S6

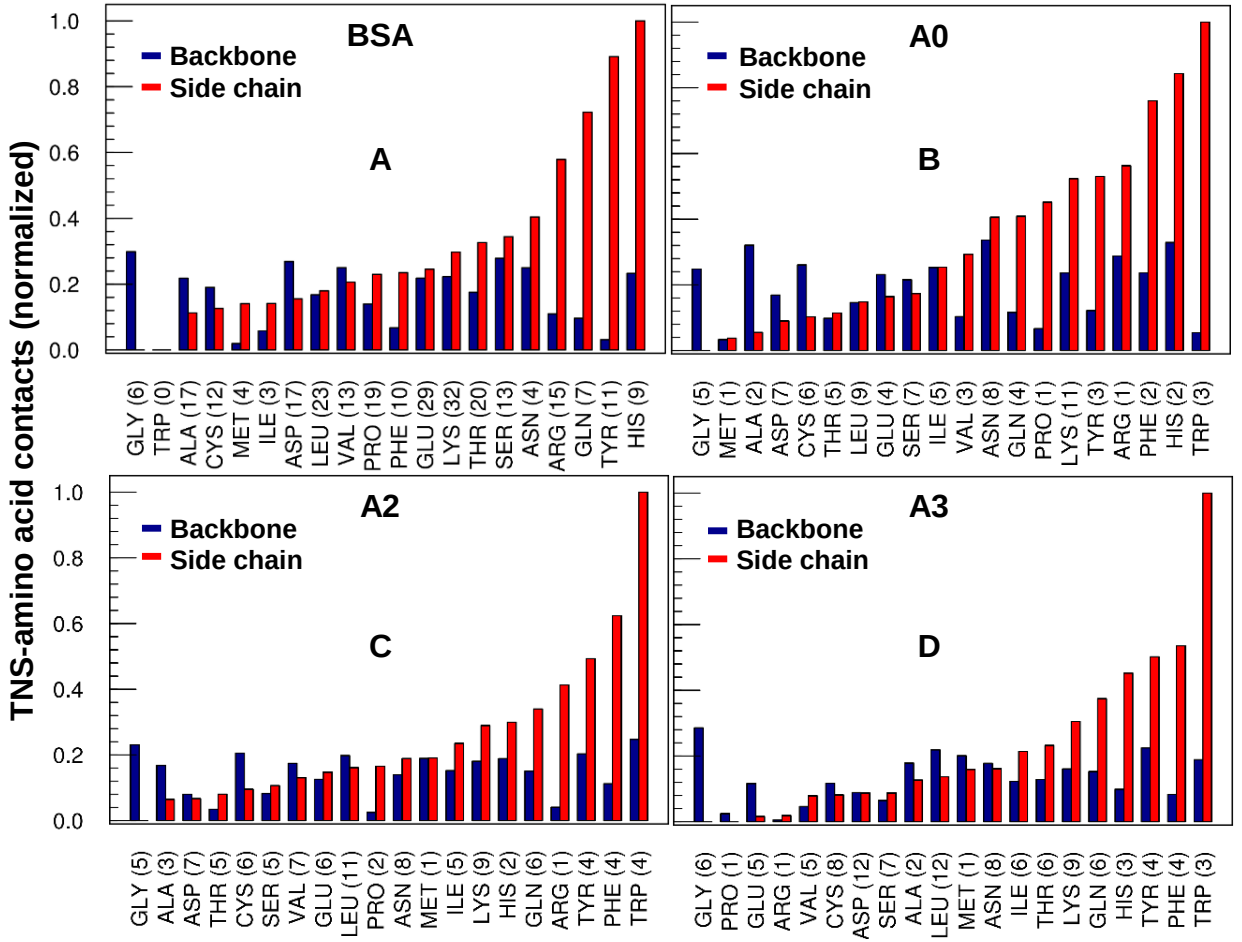


Fig. S6 The number of contacts formed between TNS and backbone (blue) or side chain (red) atoms of each amino acid residue during the last 10 ns of MD simulation is counted and normalized with respect to the number of respective amino acid in the protein. The plots present the normalized TNS-amino acid contacts in (A) BSA, (B) native state of α-LA and (C & D) in denatured states of α-LA.

Figure S7

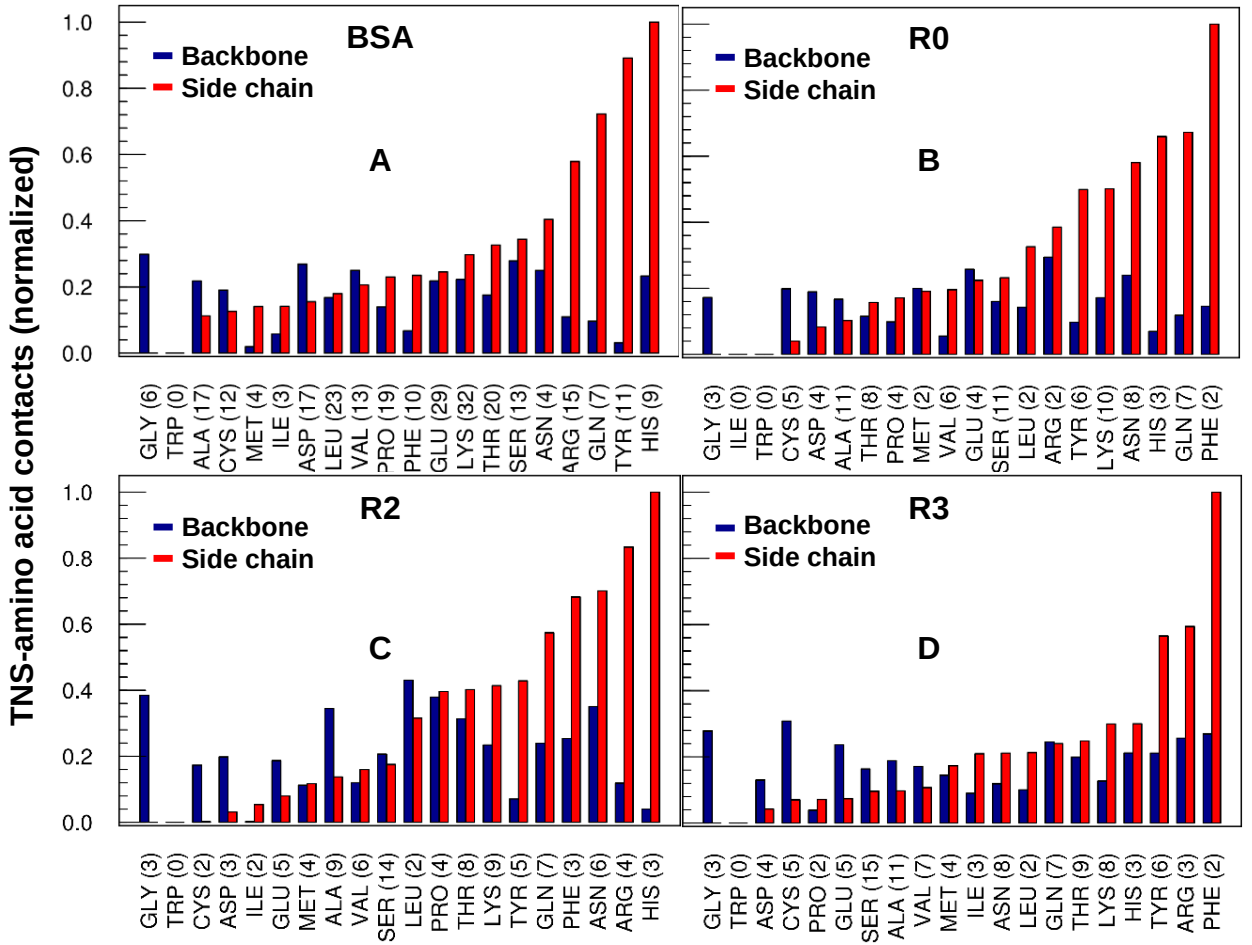


Fig. S7 The number of contacts formed between TNS and backbone (blue) or side chain (red) atoms of each amino acid residue during the last 10 ns of MD simulation is counted and normalized with respect to the number of respective amino acid in the protein. The plots present the normalized TNS-amino acid contacts in (A) BSA, (B) native state of RNase A and (C & D) in denatured states of RNase A.

Figure S8

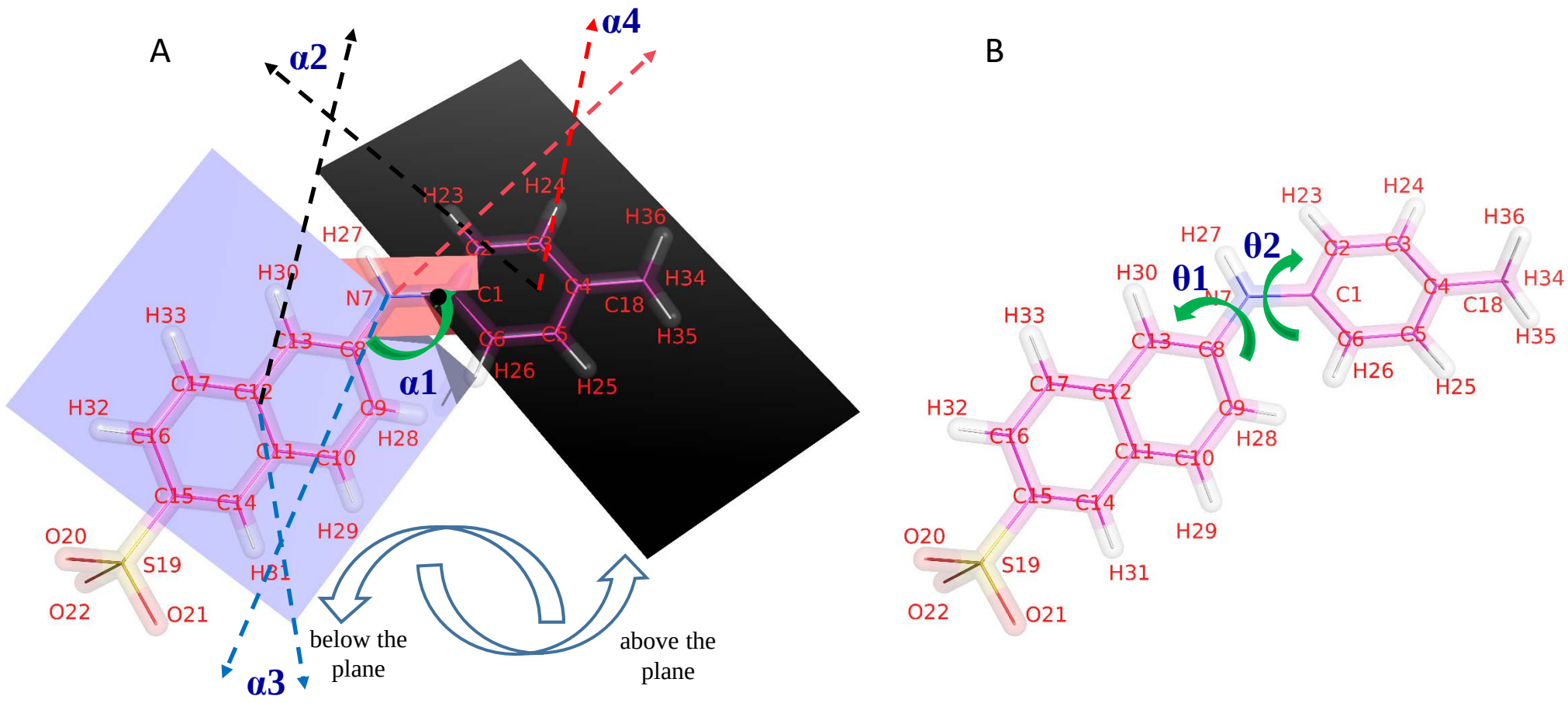


Fig. S8 Four angles and two dihedrals defined for conformational analysis of TNS. (A) With reference to the center nitrogen atom N7 as the visual plane (pink), the naphthyl (blue) and tolyl (black) aromatic rings are represented below and above the plane, respectively. The angles defined are: α1- angle between C8-N7-C1, α2- angle between the planes of naphthyl (blue) and tolyl (black) rings, α3- angle between the plane of naphthalene ring and the center nitrogen atom N7, and α4 - angle between the plane of tolyl ring and center nitrogen N7. (B) The dihedral angles defined are: θ1 - rotational angle of naphthyl ring with reference to the plane of central nitrogen N7 and θ2 - rotational angle of tolyl ring with reference to the plane of central nitrogen N7.

Figure S9

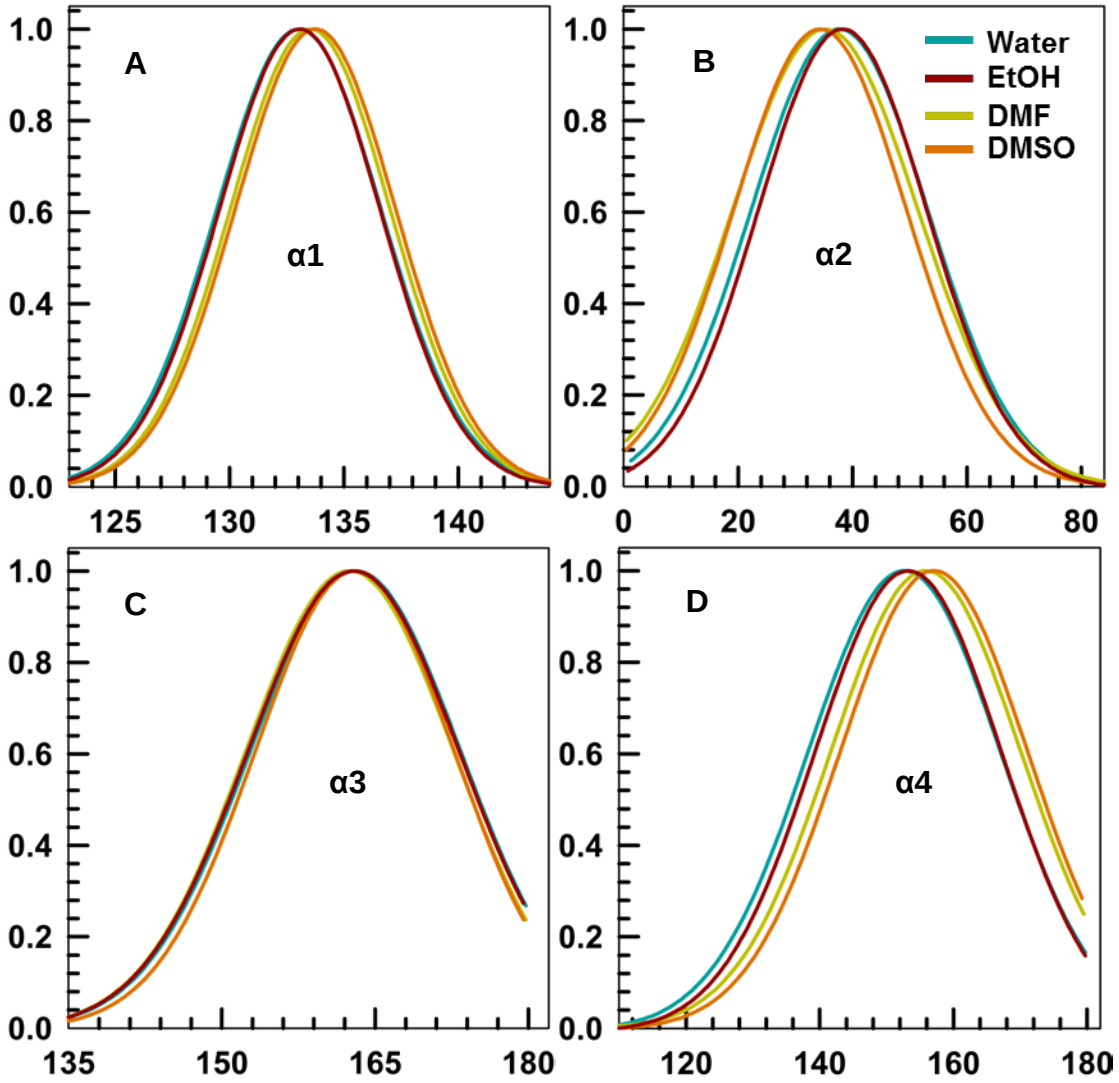


Fig. S9 Conformational changes of TNS in different solvents were analyzed using four angles (refer Fig. S8). The distributions of angles (A) α1, (B) α2, (C) α3, and (D) α4 during MD simulation in water (blue), ethanol (red), DMF (dark yellow), and DMSO (orange) are presented. The mean values and standard deviations of the distributions are presented in Table S1 (in ESI).

Figure S10

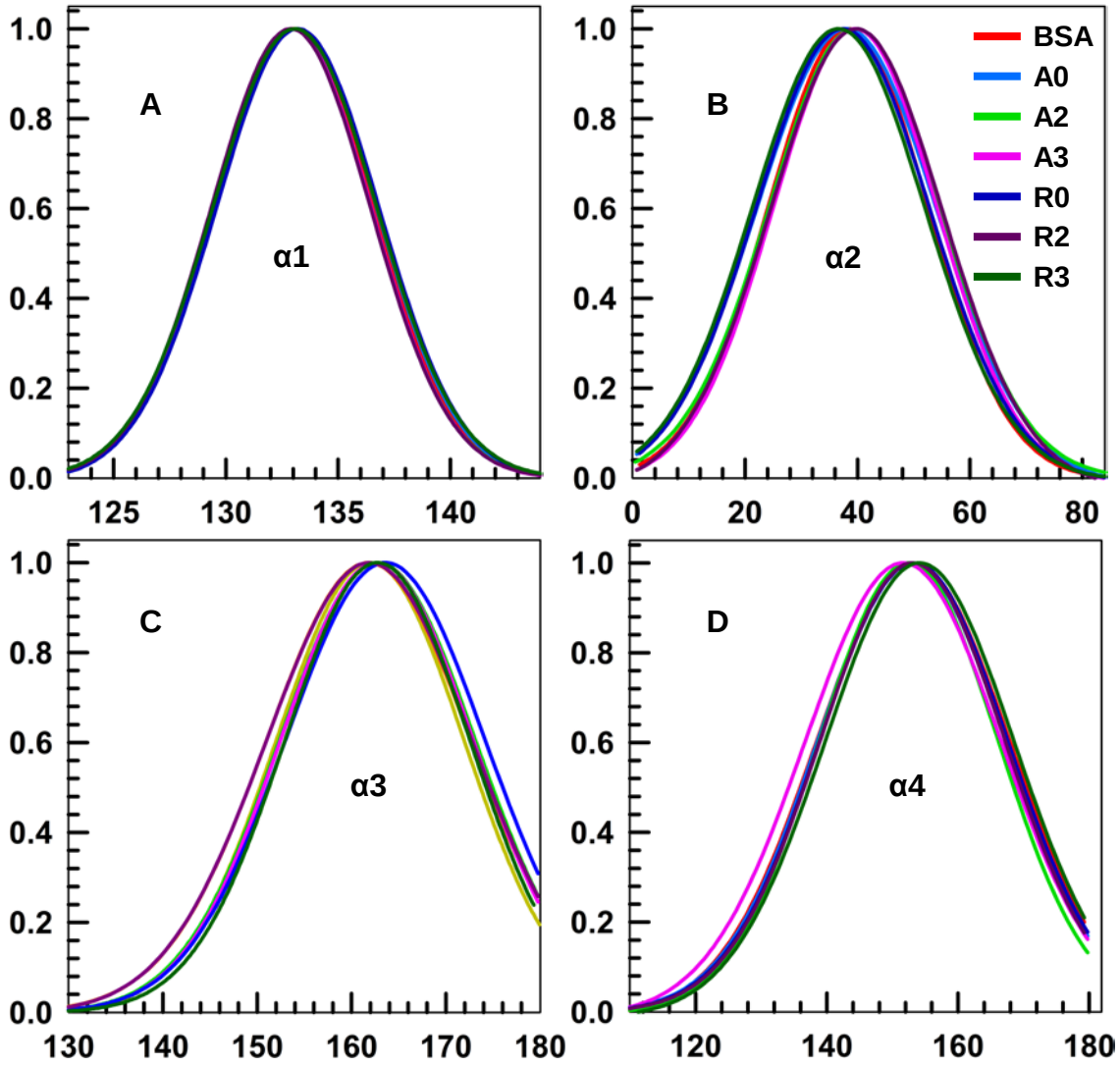


Fig. S10 Conformational changes of TNS in different proteins were analyzed using four angles (refer Fig. S8). The distributions of angles (A) α1, (B) α2, (C) α3, and (D) α4 during MD simulation in the presence of native (BSA, A0, and R0) and denatured (A2, A3, R2 and R3) proteins are presented. The mean values and standard deviations of the distributions are presented in Table S1 (in ESI).

Figure S11

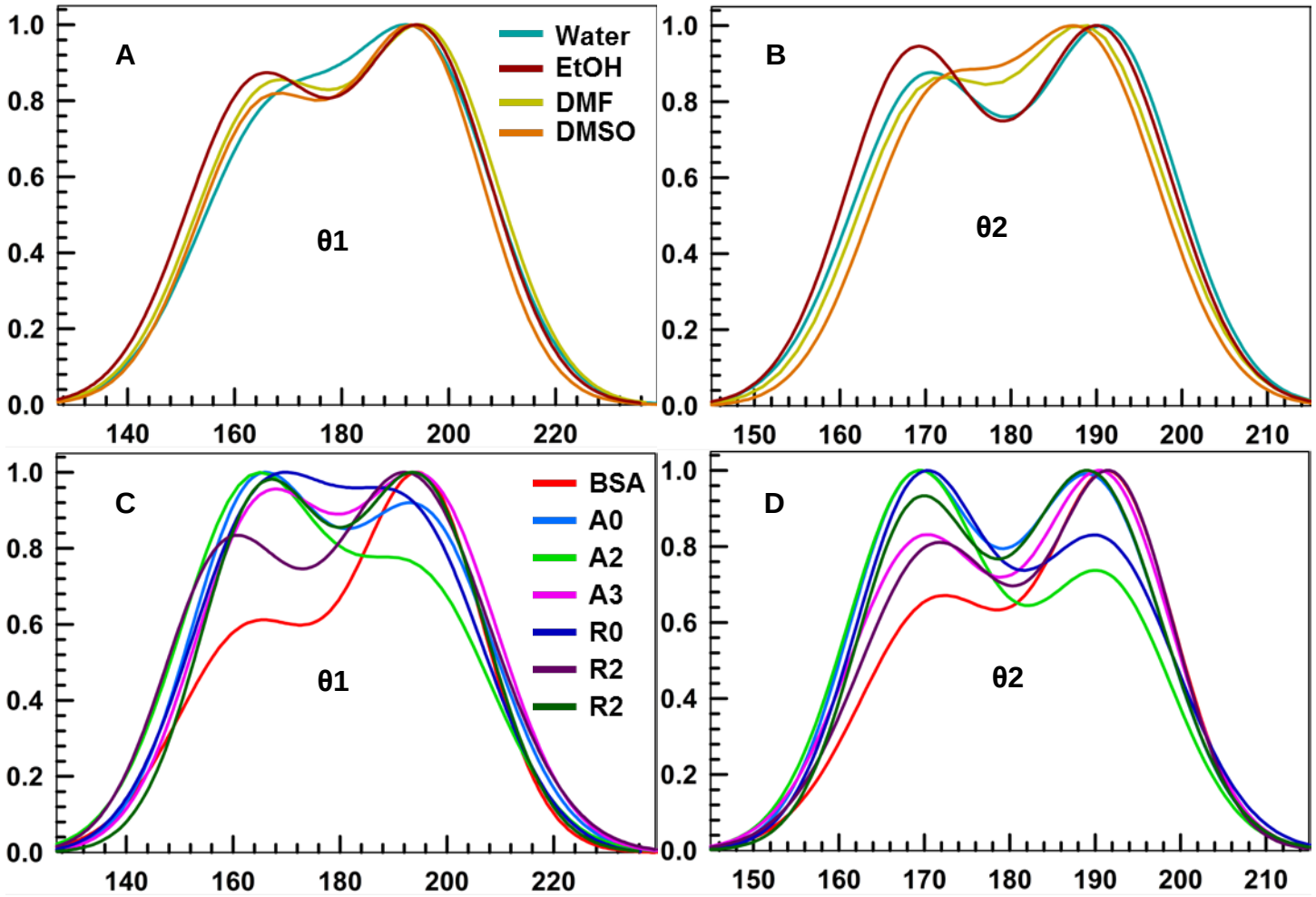


Fig. S11 TNS conformational changes analyzed using two dihedrals, θ1 and θ2 (refer Fig. S8). The distribution of dihedral angles during MD simulation in different solvents (A and B) and in different conformations of proteins (C and D) are shown. The mean values and standard deviations of the distributions are presented in Table S2 (in ESI).

Figure S12

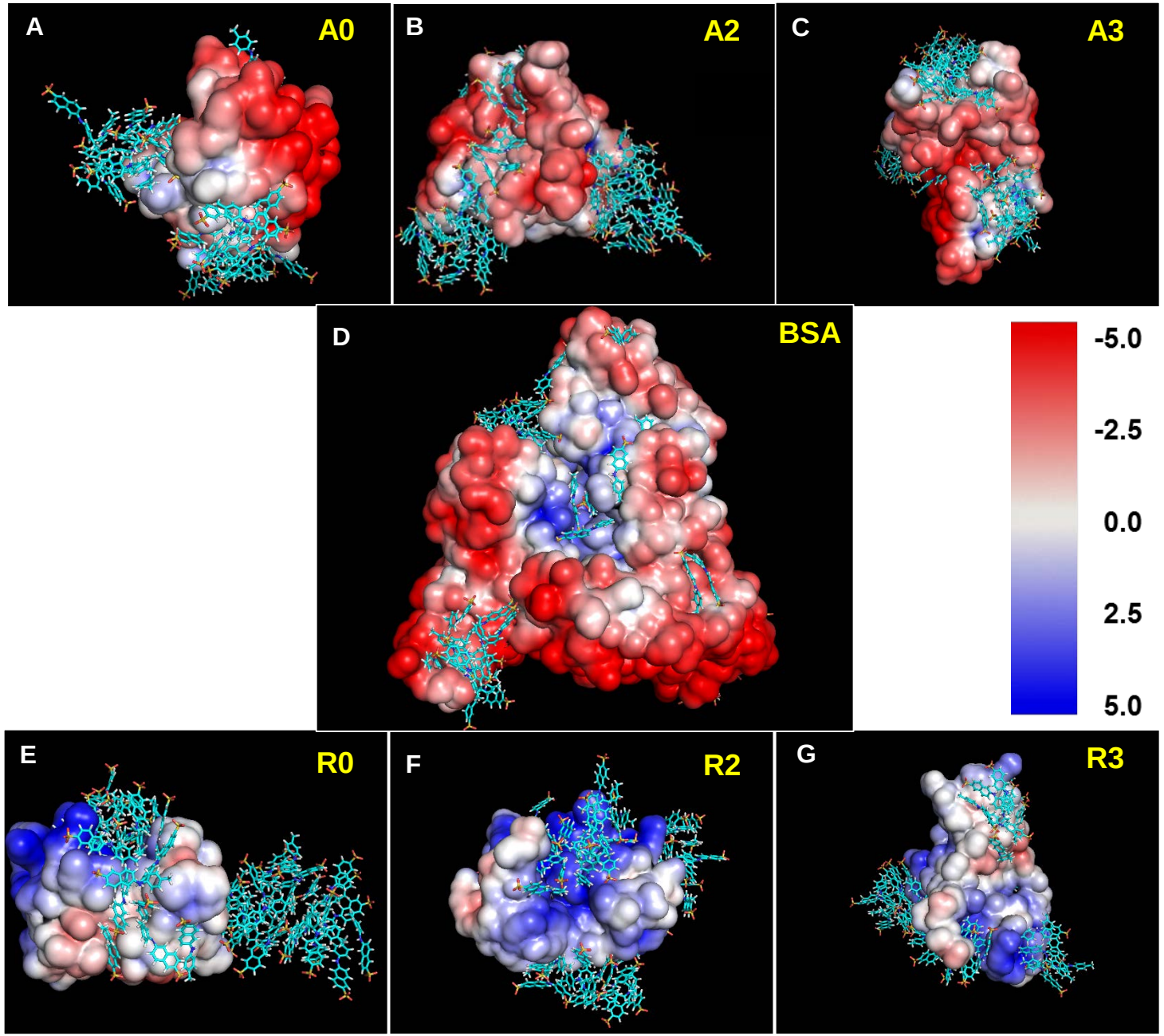


Fig. S12 Electrostatic surface maps of (A) native state of α-LA, (B & C) denatured states of α-LA, (D) native state of BSA, (E) native state of RNase A, and (F & G) denatured states of RNase A constructed using APBS are presented as surface models with bound TNS molecules as stick models. The bar scale represents the color codes of charge distribution.

Figure S13

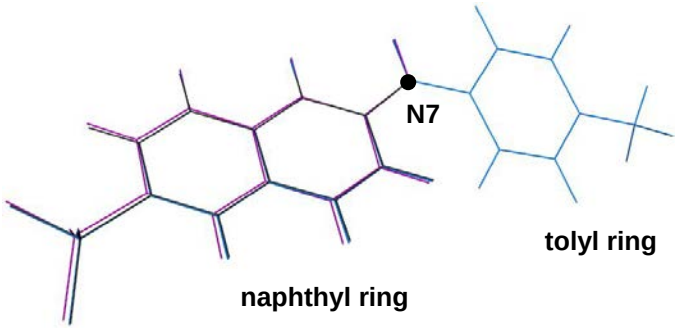


Fig. S13 Representation of most accessible conformation of TNS in water (blue), DMF (red) and BSA (black). The central nitrogen atom N7 is labeled for reference.

Figure S14

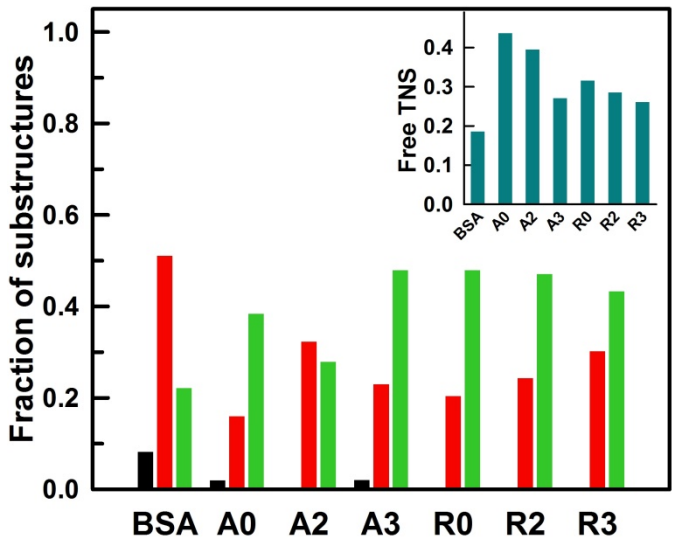


Fig. S14 The fraction of distribution of the three-substructures , monomeric TNS buried inside the protein (black), lower order aggregated on the protein surface (red) and higher order aggregates almost completely exposed to the solvent calculated from MD simulation of 50 molecules of TNS in different proteins. The inset shows the fraction of free-TNS (unbound to protein) in each case.

TABLES

Electronic Supporting Information

Table. S1 Distribution of four angles of TNS (as defined in Fig. S7) during MD simulation in the presence of different solvents and the proteins

Protein / Solvent	α1			α2			α3			α4		
	μ	σ	ƒ	μ	σ	ƒ	μ	σ	ƒ	μ	σ	ƒ
BSA	132.989	3.556	8.374	38.353	14.214	33.471	161.970	9.940	23.406	153.169	14.734	34.697
A0	133.130	3.555	8.372	38.037	15.942	37.540	162.726	10.379	24.440	152.905	14.264	33.588
A2	132.913	3.526	8.303	39.456	15.098	35.552	161.799	10.789	25.407	152.481	13.648	32.139
A3	132.937	3.513	8.273	39.573	14.539	34.235	162.676	10.199	24.016	151.649	14.936	35.171
R0	133.200	3.605	8.488	37.381	15.383	36.223	163.569	10.573	24.897	153.254	14.296	33.664
R2	132.895	3.520	8.288	39.775	15.251	35.913	161.872	10.879	25.619	153.074	14.060	33.109
R3	133.072	3.632	8.552	36.624	15.284	35.991	162.744	9.7890	23.051	154.143	14.411	33.935
Water	133.034	3.592	8.458	37.695	15.455	36.393	163.074	10.286	24.221	152.656	14.387	33.879
DMF	133.528	3.498	8.238	35.210	16.172	38.082	162.508	10.168	23.944	155.878	14.175	33.379
DMSO	133.756	3.515	8.277	34.305	15.087	35.528	163.014	9.726	22.903	157.087	14.015	33.003
Ethanol	133.069	3.513	8.272	38.237	14.722	34.668	162.884	10.34	24.349	153.255	13.964	32.884

μ - mean; σ – standard deviation; ƒ-full width at half maximum

Table. S2 Distribution of two dihedral angles of TNS (as defined in Fig. S7) during MD simulation in the presence of different solvents and the proteins

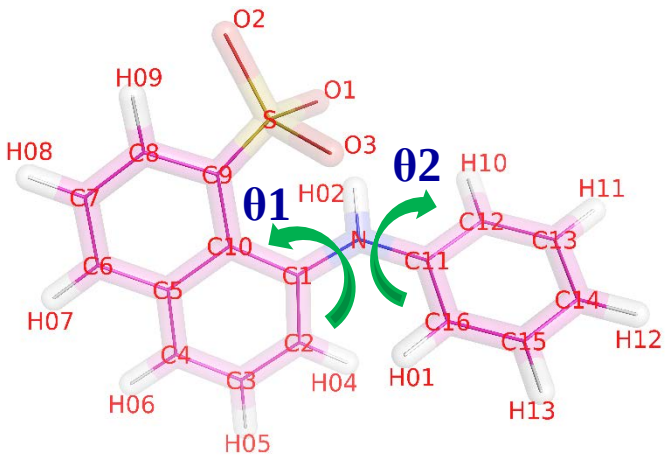
Protein / Solvent	θ1						θ2					
	μ1	σ1	ƒ1	μ2	σ2	ƒ2	μ1	σ1	ƒ1	μ2	σ2	ƒ2
BSA	163.094	13.357	31.453	195.334	11.668	27.476	171.274	8.736	20.572	192.031	7.283	17.150
A0	163.798	12.044	28.361	195.133	13.725	32.319	168.626	7.491	17.640	189.640	8.323	19.599
A2	163.120	13.171	31.016	194.916	13.706	32.275	169.105	8.040	18.932	191.006	7.944	18.708
A3	164.753	12.341	29.060	195.874	13.642	32.124	169.538	8.459	19.918	191.154	7.806	18.381
R0	164.856	13.131	30.920	193.719	13.734	32.341	169.516	7.612	17.924	190.737	8.734	20.567
R2	158.113	11.356	26.741	192.642	15.213	35.825	171.239	8.796	20.713	192.253	7.063	16.631
R3	165.260	11.437	26.932	195.023	12.228	28.796	168.896	6.990	16.460	189.415	8.424	19.837
Water	166.604	13.662	32.171	195.479	13.110	30.87	169.662	8.373	19.716	191.392	8.181	19.265
DMF	165.279	13.083	30.807	196.435	12.827	30.204	169.502	7.908	18.621	189.338	8.723	20.542
DMSO	164.920	12.397	29.192	194.521	12.070	28.424	170.557	7.810	18.391	188.852	8.521	20.065
Ethanol	163.708	12.896	30.367	195.485	12.522	29.487	168.479	7.875	18.545	190.498	8.288	19.517

μ - mean; σ – standard deviation; ƒ-full width at half maximum

Table. S3 The angle and dihedrals of protein bound ANS correspond to the angle α1 and dihedrals in TNS as obtained from PDB

PDB ids from which ANS conformations were obtained‡	Bond angle between C1-N-C11 (equivalent to α1)	Rotational angle of naphthyl ring (equivalent to θ1)	Rotational angle of phenyl ring (equivalent to θ2)
1txc-ans1	132.2	141.9	150.4
1txc-ans2	135.5	151.4	168.2
1txc-ans3	134.6	142.5	178.5
3pxq-ans1	129.2	142.2	143.7
3pxq-ans2	127.7	127.4	164.7
3pxq-ans3	127.7	169.1	124.4
4n3e-ans1	119.9	173.5	149.7
4n3e-ans2	125.2	143.3	142.8
4n3e-ans3	121.8	167.6	147.2
4n3e-ans4	121.1	157.6	145.3

‡ - ans1, ans 2, etc., represents ANS molecules bound to the same protein



stick model of ANS