Electronic Supplementary Information

Computational Design by Evolving Folds and Assemblies to Desired Structures over the Alphabet in L- and D-α-Amino Acids

Punam Ghosh, Ameeq ul Mushtaq, and Susheel Durani[a*]

[a]Department of Chemistry, Indian Institute of Technology Bombay, Mumbai-400076, India

Sequence Design and Validation

SymmDock

SymmDock[1,2] is a geometry-based docking algorithm for the prediction of cyclically symmetric complexes given the structure of the asymmetric unit. Searches are constrained to symmetric cyclic transformations of a given order n. SymmDock was implemented by using the standard protocol.

Inverse Algorithm for Sequence Design

The inverse-design package **IDeAS**[3] implements combination of deterministic pruning algorithm Dead End Elimination[4] and stochastic search method Monte Carlo.[5] Energy functions for sequence solution involve accessibility-based solvation energy, H-bond energy, coulomb energy, and entropy in frequency of side chain usage. The software uses natural database of rotamers[6,7] and their suitable symmetry transforms as the rotamers of D structure. IDeAS has been validated by characterization of the designs.[8,9]

Molecular Dynamics

MD was performed using gromos-96-43a1 force field in GROMACS package.[10,11] Simulations were under NVT with periodic boundary condition in a 4.5 nm cubic box for monomer and a 6.6 nm box for pentamer. SPC water model was used. Bond lengths were constrained with SHAKE.[12] Time step of 2 femto second was used. Non-bonded interactions were cut-off at 1.4 nm using shift function. The potential energy of the system, i.e., peptide and water, was minimized using steepest descent algorithm with a tolerance limit of 100 KJ mol^{-1} nm^{-1} in sufficient number of steps to achieve convergence. Position restrained molecular dynamics was performed for 200 ps at 298 K to equilibrate water molecules. Initial velocities were drawn from Maxwellian distribution. The peptide and solvent were separately coupled to Berendsen temperature bath using a time constant of 0.1 ps. Conformational clustering was performed to RMSD cut-off $\leq$ 0.15 nm over Cα atoms.

Experimental Section

Peptide Synthesis

Amino acids were purchased from Sigma Aldrich and Nova Biochem. The synthesis was performed by manual method on Rink Amide AM resin using standard F-moc chemistry. TBTU/HOBT/DIC or HOBT/DIC were used as the coupling reagents. Monitored with Kaiser and chloranil tests, each coupling typically required about 6 hrs. F-moc was deblocked with ~30% piperidine-DMF. N-terminus was acylated (-$NHCOCH_3$) with Ac_2O:DIPEA:DMF (1:2:20). The cleavage of the final peptides and the deprotection of side chains were accomplished together with reagent-K TFA:H2O:Phenol:Thioanisole: EDT in the ratio 82.5:5:5:5:2.5. The filtrate was precipitated with cold anhydrous ether and was lyophilized to a dry powder.

MALDI-MS

Mass spectra were recorded in MALDI-TOF (Matrix Assisted Laser Desorption Ionization-Time Of Flight) mode on a duly calibrated AXIMA-CFR Kratos instrument. Positive ions were detected in linear/reflectron mode.

Nuclear Magnetic Resonance

NMR experiments were performed on Bruker 800 MHz instrument equipped with cryoprobe, at 298 K in 9:1 H_2O and D_2O with 0.1 mM 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS) as internal reference at 5 mM concentration. Phase sensitive TOCSY and ROESY spectra were recorded and data were processed with TOPSPIN software of Bruker. The mixing times for TOCSY and ROESY experiments were 80 ms and 300 ms, respectively. The number of increments in t_1 was 512 and 450 with 32 and 48 FIDs per increment for TOCSY and ROESY, respectively. Structure calculation was performed with NMR modeling software CYANA 2.1. Representative structures were energy minimized using Gromacs software package. Molecular structures were rendered with DS Visualizer.

Circular Dichroism

CD measurements were performed on JASCO J-815 CD spectropolarimeter calibrated with d_{10}-camphoursulphonic acid. Data were collected at 298 K in 0.2 cm path-length quartz cell with 2 nm bandwidth in 195 - 240 nm range. Scanning at 100 nm / min with 1.0 s time constant in 1 nm steps, five scans were averaged after baseline correction for solvent. Peptides were prepared in the range of 5-250 µM by optical measurements in aqueous solution of water-methanol mixture at pH 7.0. The observations in millidegrees were converted to mean residue ellipticity using a reported relation.

Fluorescence

The fluorescence measurements were taken using a Varian Cary Eclipse fluorimeter. Data were collected at 298 K in 1cm cell, with excitation wavelength of 280 nm. The excitation and emission slits were kept at 5 nm with a scanning speed of 120 nm/min. The peptides were prepared in the range of 5-240 µM by optical measurements of the peptide and of standard NATA (SIGMA chemicals).

Fluorescence Anisotropy

Steady state fluorescence anisotropy measurements were recorded with Varian Cary Eclipse fluorimeter equipped with an emission-excitation polarizer. All the experiments are performed at 25^0C.

Dynamic Light Scattering

The size-distributions of the peptide samples at different concentrations were assessed by dynamic light scattering instrument (Model: 90 Plus Particle Size Analyzer); Brookhaven Instrument Co., Holtsville, NY, USA) which is a technique to detect the fluctuations of the scattering intensity due to the Brownian motion of particles in suspension using high power 35 mW diode laser source. The measurements were carried out at 90° scattering angle. Due to the random motion of the particles, the scatted light intensity fluctuates in time. Processing the fluctuating signal with a state of the art digital autocorrelator yields the particle's diffusion coefficient, from which the equivalent spherical particle size is calculated using the Stokes-Einstein equation. The resulting correlation function may be inverted to find a size-distribution for the particles using software program 90 Plus particle Sizing Software, Brookhaven Instrument Corporation. The samples were prepared in water using 0.22 μ filter with two hours of sonication.

Atomic Force Microscopy

An aliquot of peptide solution (20 μL) incubated in water was transferred onto a freshly cleaved mica surface and the sample solution was uniformly spread throughout. The sample coated mica was air dried for 12 hours at room temperature and imaged by Vecco Digital Instrument Nanoscope-IV at a scanning rate of 1.507 Hertz.

References

1. D. S. Duhovny, Y. Inbar, R. Nussinov and H. J. Wolfson, *Nucleic Acids Research*, 2005, **33**, 363.

2. D. S. Duhovny, Y. Inbar, R. Nussinov and H. J. Wolfson, *Proteins*, 2005, **60**, 224.

3. R. Ranbhor, A. Tendulkar, A. Kumar, V. Ramakrishnan, K. Patel, K. R. Srivastava and S. Durani, Submitted 2011.

4. R.F. Goldstein, *Biophys. J.*, 1994, **66**, 1335.

5. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. J. Teller, *Chem. Phys.*, 1953, **21**, 1087.

6. T. P. Creamer, *Proteins*, 2000, **40**, 443.

7. A. J. Doig and M. J. Sternberg, *Prot. Sci.*, 1995, **4**, 2247.

8. K. Patel, B. Goyal, A. Kumar, N. Kishore and S. Durani, *J. Phys. Chem. B*, 2010, **114**, 16887.

9. K. Patel, K. R. Srivastava and S. Durani, *Bioorg. Medicinal Chem.*, 2010, **18**, 8270.

10. E. Lindahl, B. Hess and D. van der Spoel, *J. Mol. Mod.*, 2001, **7**, 306.

11. W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hunenberger, P. Kruger, A. E. Mark, W. R. P. Scott and I. G. Tironi, *Biomolecular Simulation: The GROMOS96 manual and user guide*, Hochschulverlag AG an der ETH Zurich, Zurich, Switzerland, 1996.

12. J. P. Ryckaert, G. Ciccoti and H. J. C. Berendsen, *J. Comp. Phys.*, 1977, **23**, 327.

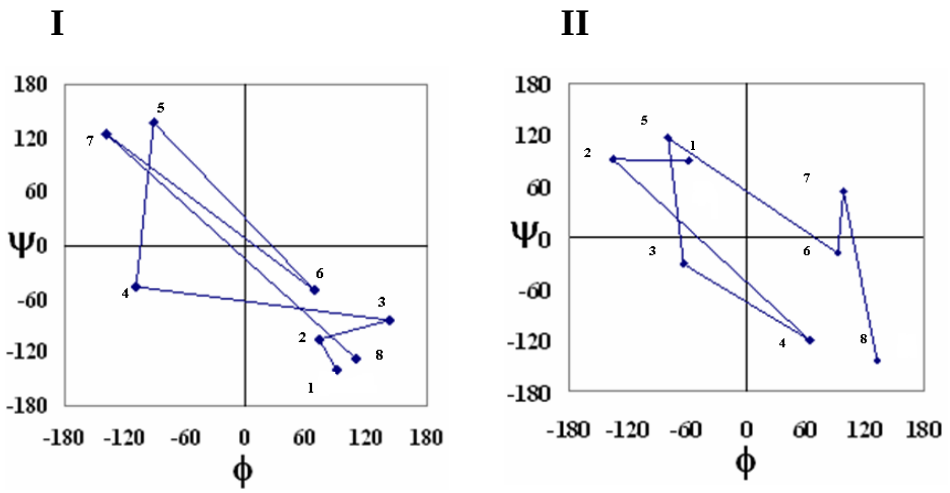


Figure S1: The folds **I** and **II** have the residues in the left or right half of φ, ψ map in accordance with the steric preferences of the L or D residues, with notable exception of ^{6}L in fold **I** being in the φ, ψs that are sterically favored for D residues.

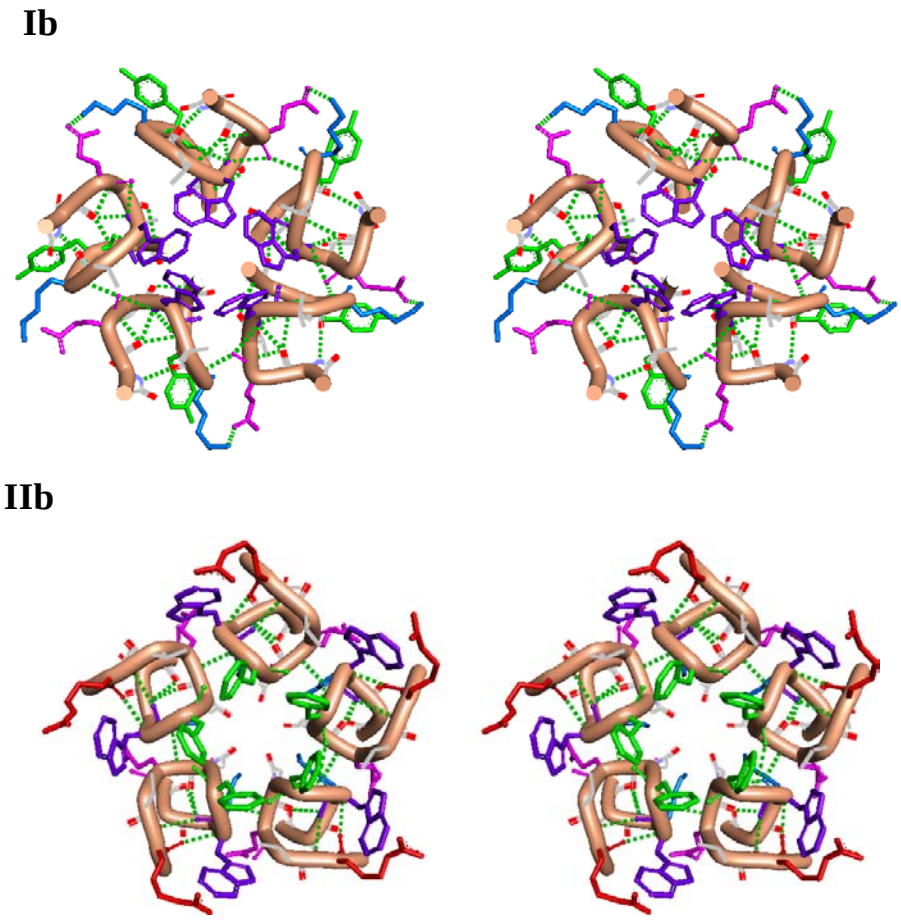


Figure S2: Stereoimages of tube and wireframe models of peptides **Ib** and **IIb** highlighting hydrogen bonds with dotted lines.

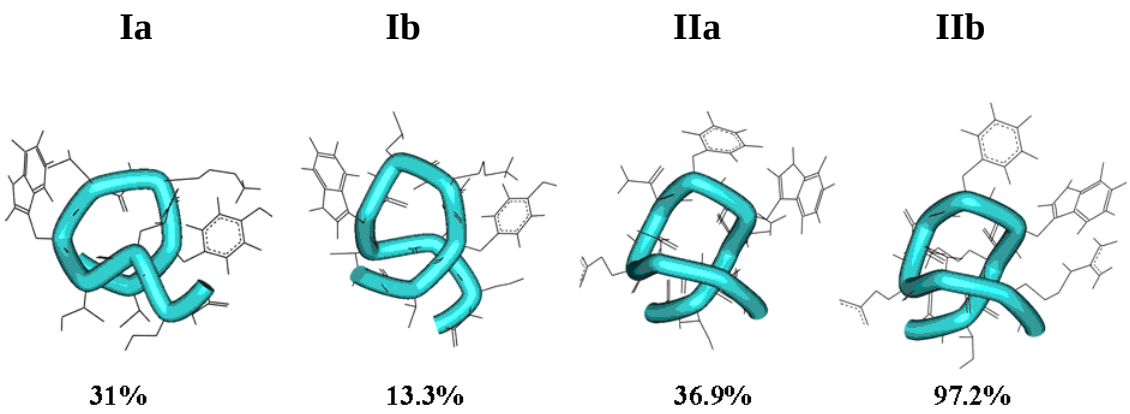


Figure S3: Tube and wireframe representation of the central member of the most populous cluster of 50 ns MD trajectory showing the percentage of the folds populating the top cluster.

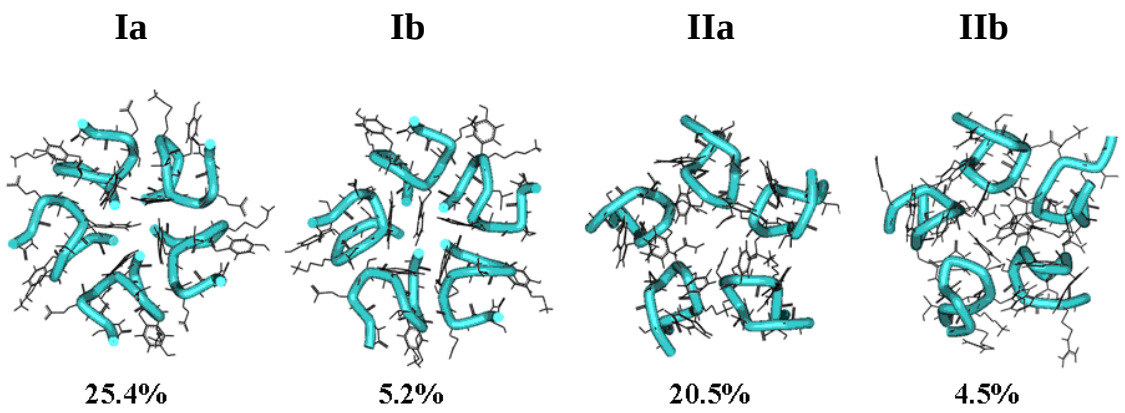
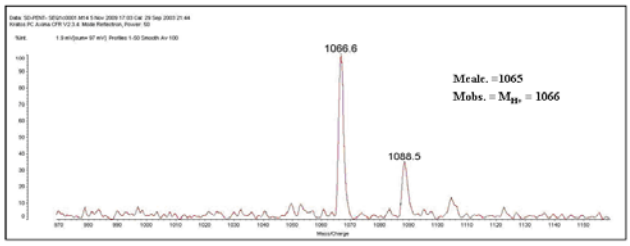


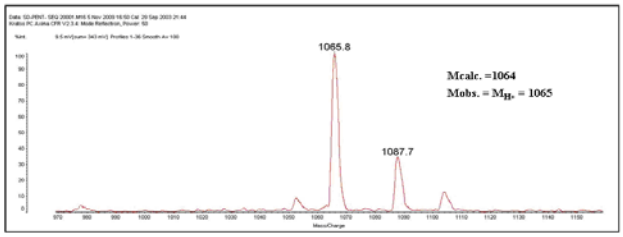
Figure S4: Tube and wireframe representation of the central member of the most populous cluster of 50 ns MD trajectory showing the percentage of the structures populating the top cluster.

Ia



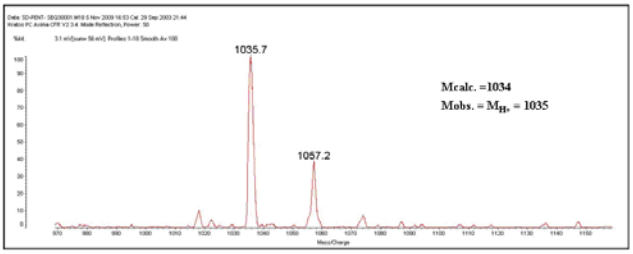
Panel A

Ib



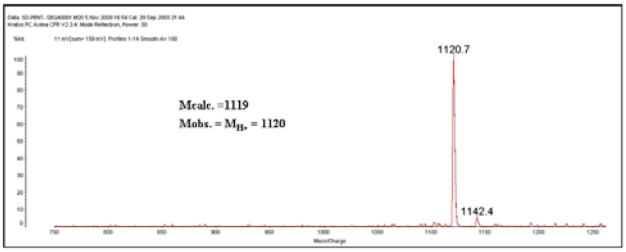
Panel B

IIa



Panel C

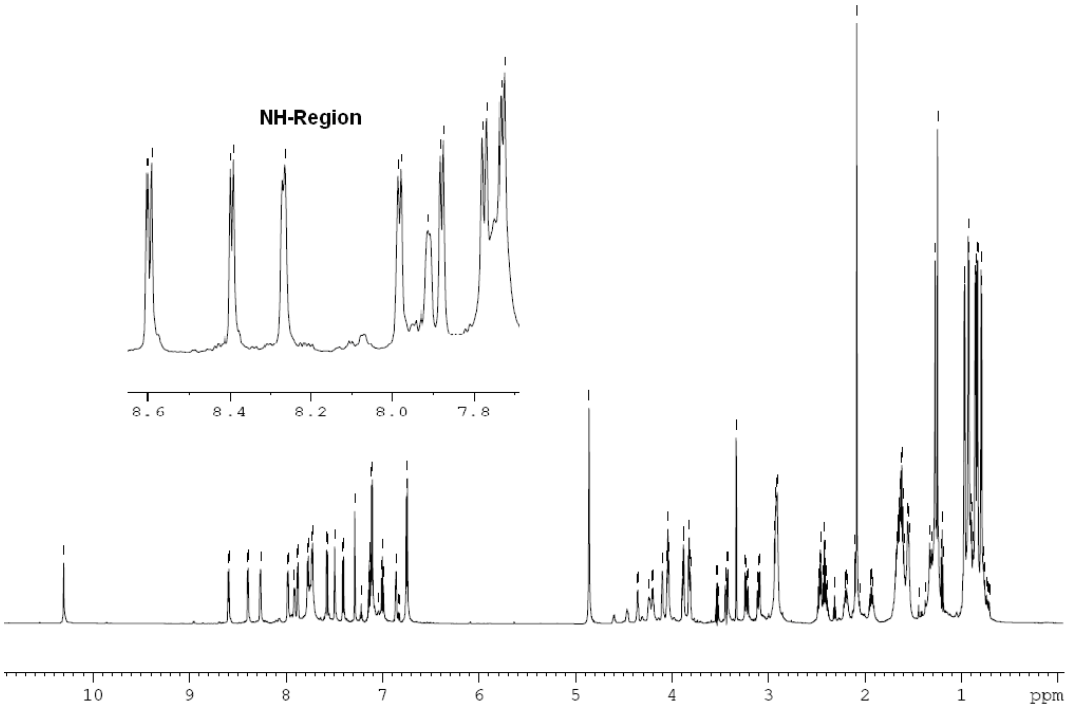
IIb



Panel D

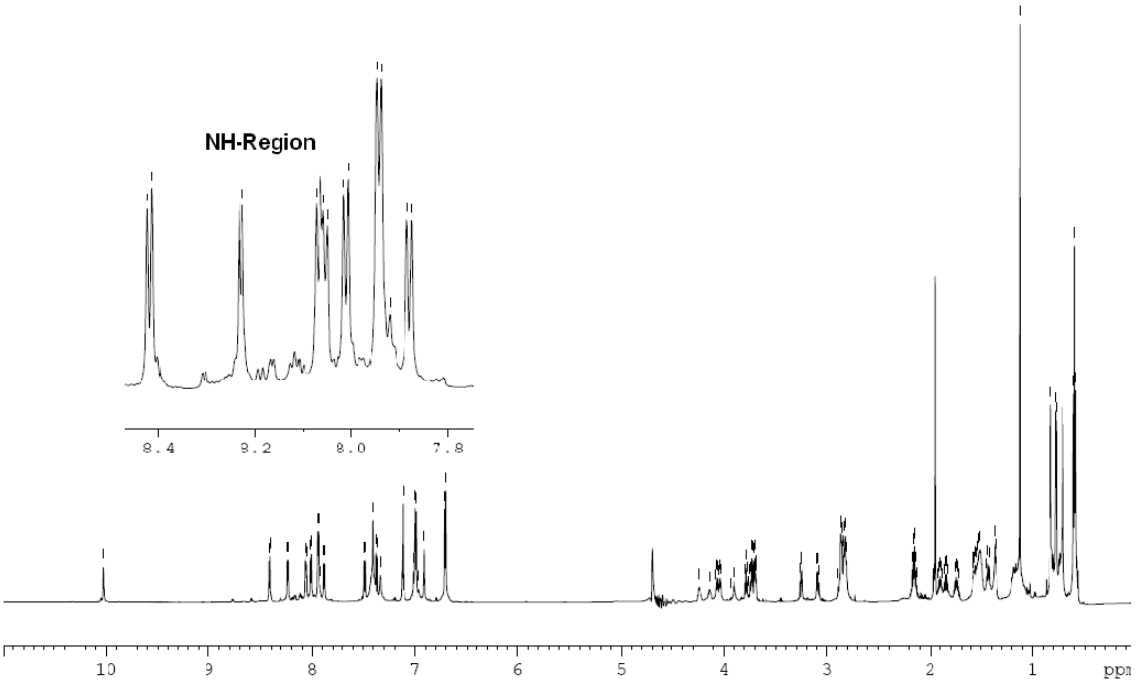
Figure S5: Maldi MS spectra of octapeptides.

Ia



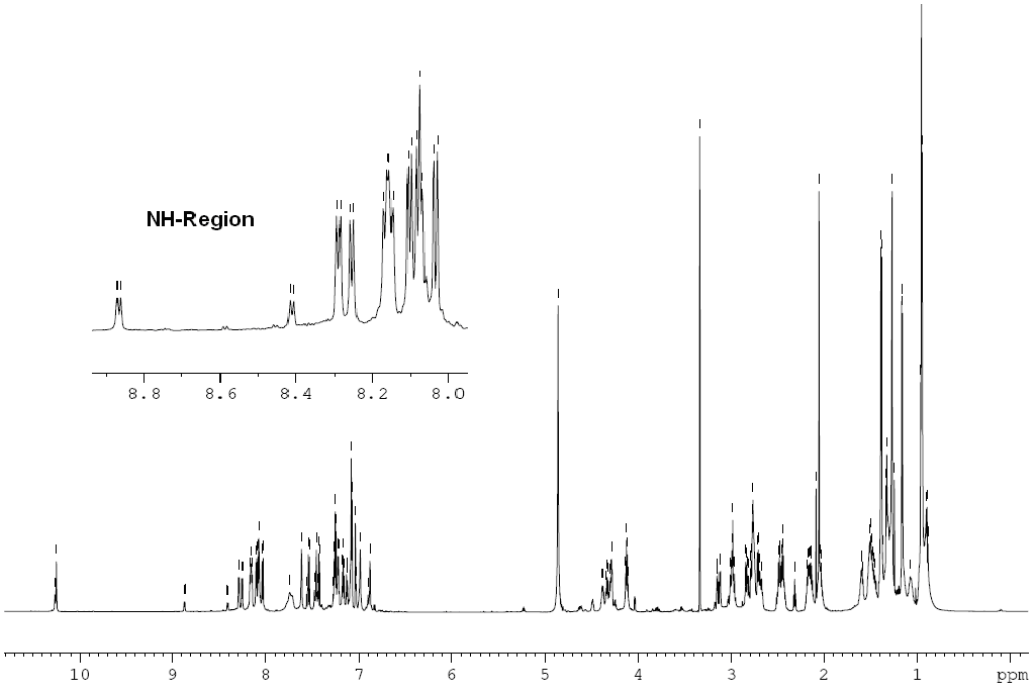
Panel A

Ib



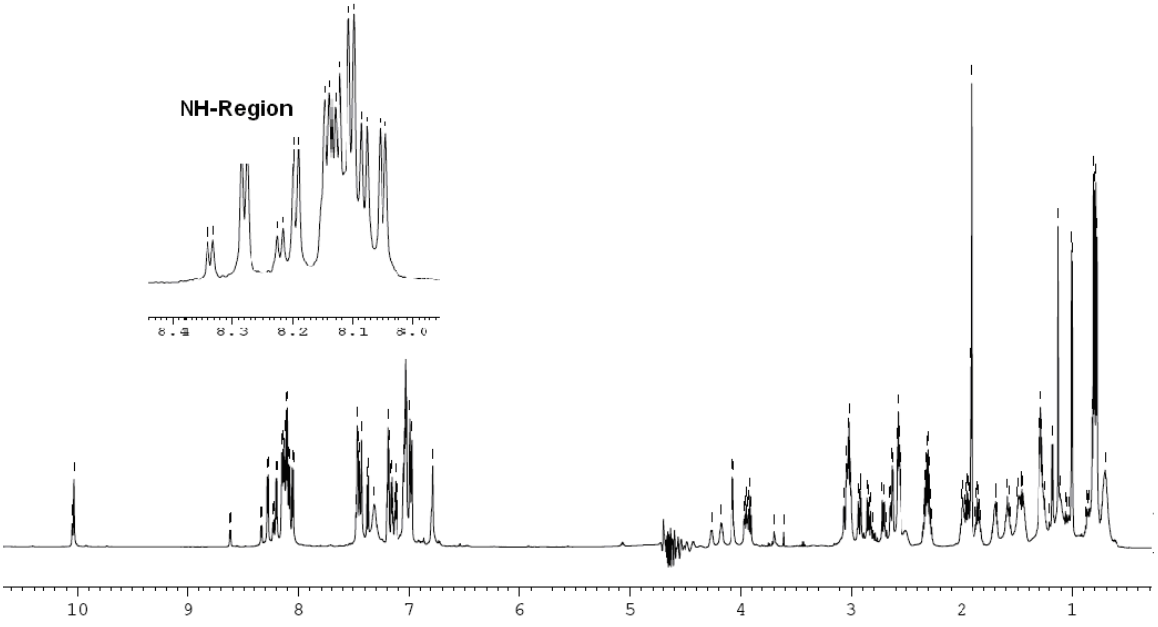
Panel B

IIa



Panel C

IIb



Panel D

Figure S6: ^{1}H-NMR spectra of peptides **Ia** (Panel A) and **Ib** (Panel B) and peptides **IIa** (Panel C) and **IIb** (Panel D) recorded at 800 Mz in 90% D_2O-H_2O mixture. Expansions of NH regions are shown as insets.

Ib

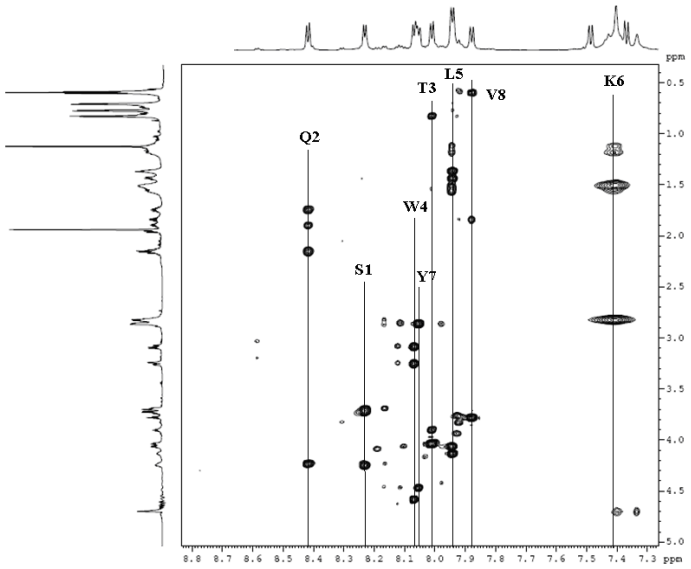


Figure S7: A portion of 800MHz TOCSY spectrum of peptide **Ib** showing cross-peak region of peptide-NH and CαH resonances highlighted.

Ib

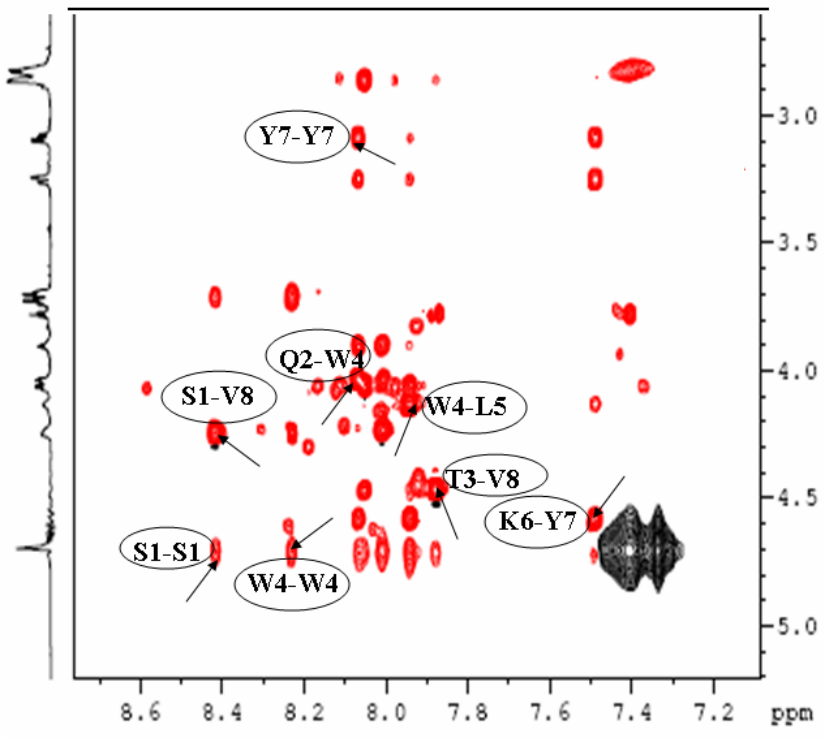


Figure S8: A portion of ROESY spectrum of peptide Ib with specific NOE cross peaks labeled.

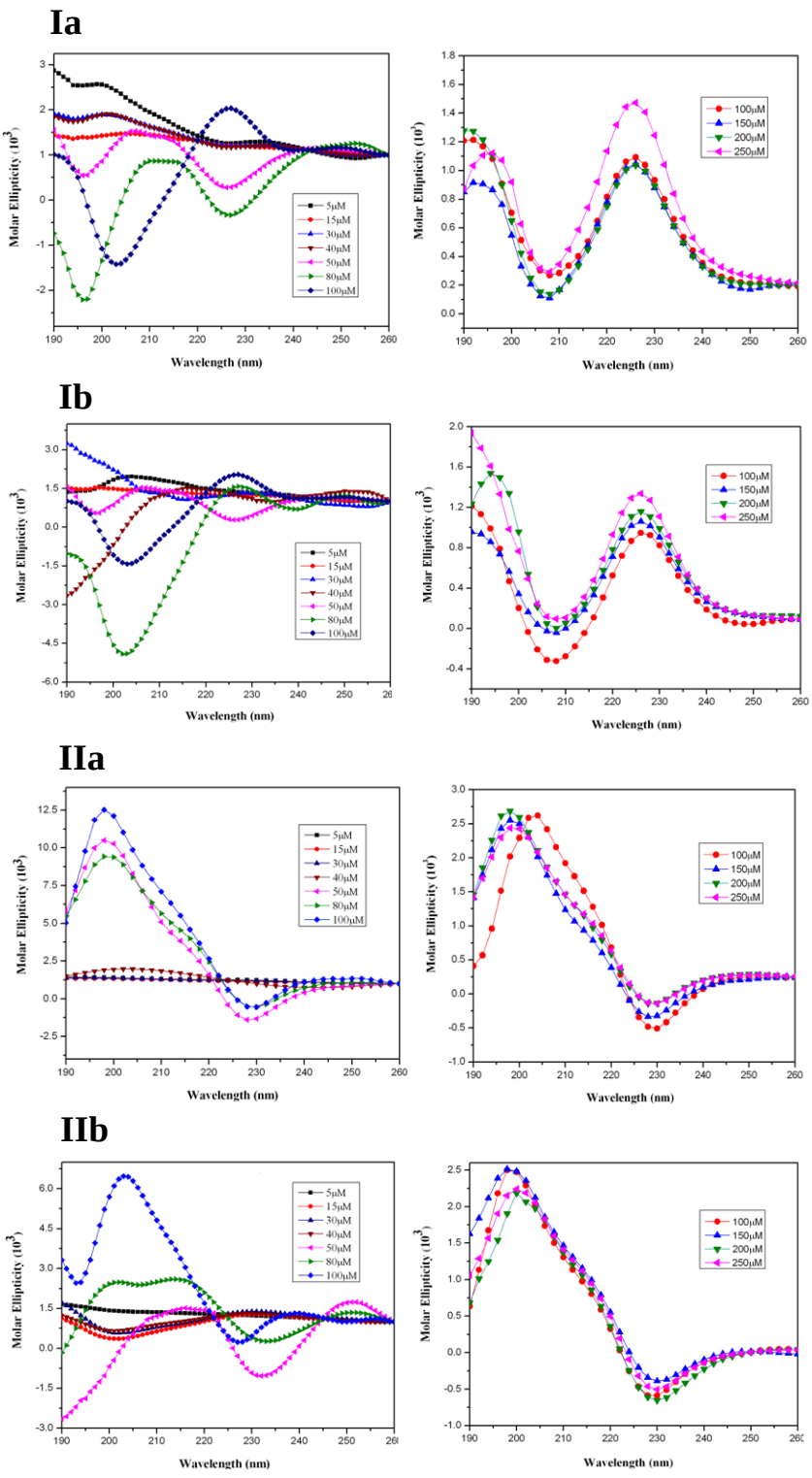
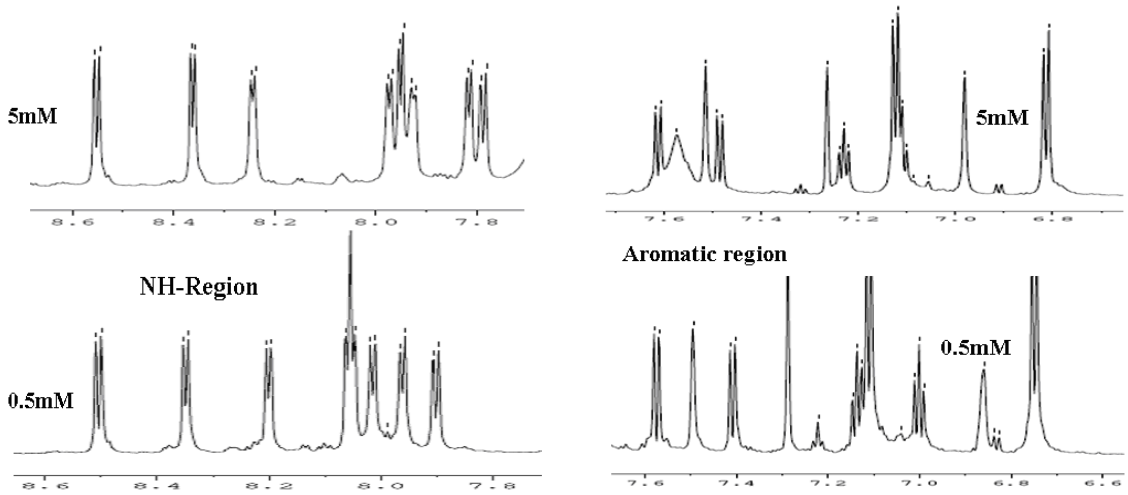


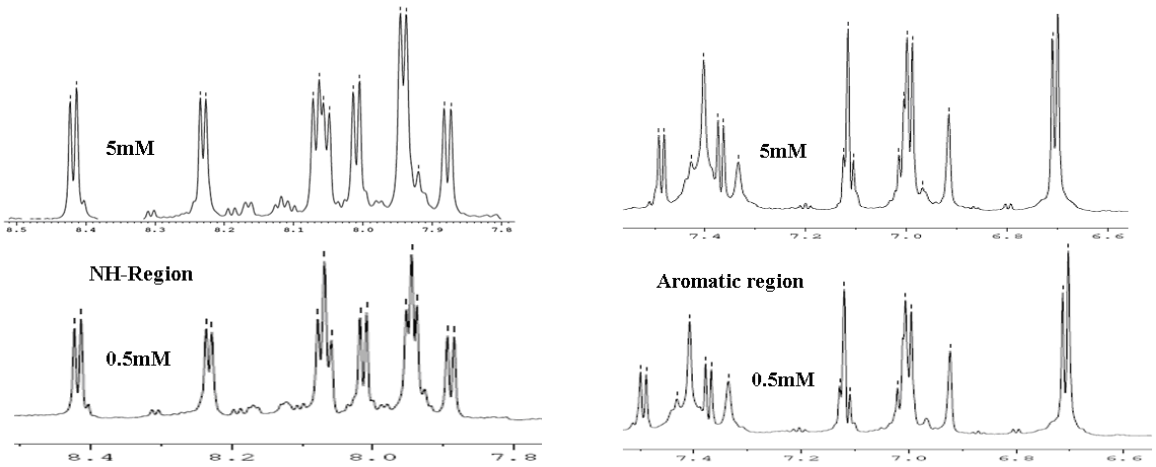
Figure S9: CD traces recorded in the concentration range 5-100 μM (Left Panels) and 100-250 μM (Right Panels). The solvent is 20% methanol-water for the neutral peptides **Ia** and **IIa** and water for the charged peptides **Ib** and **IIb**.

Ia

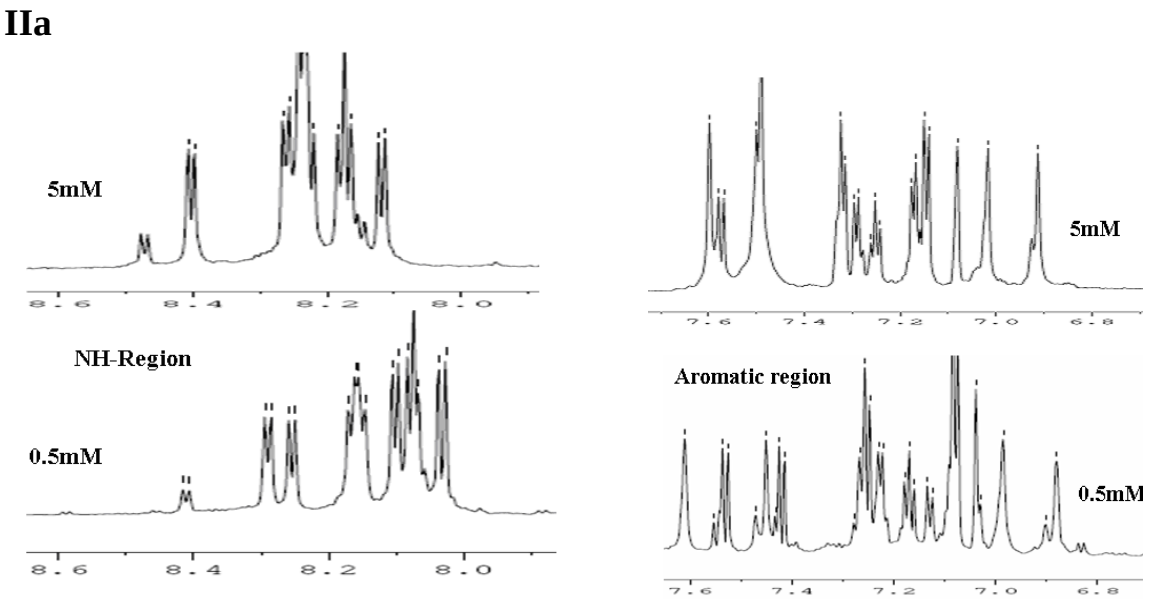


Panel A

Ib



Panel B



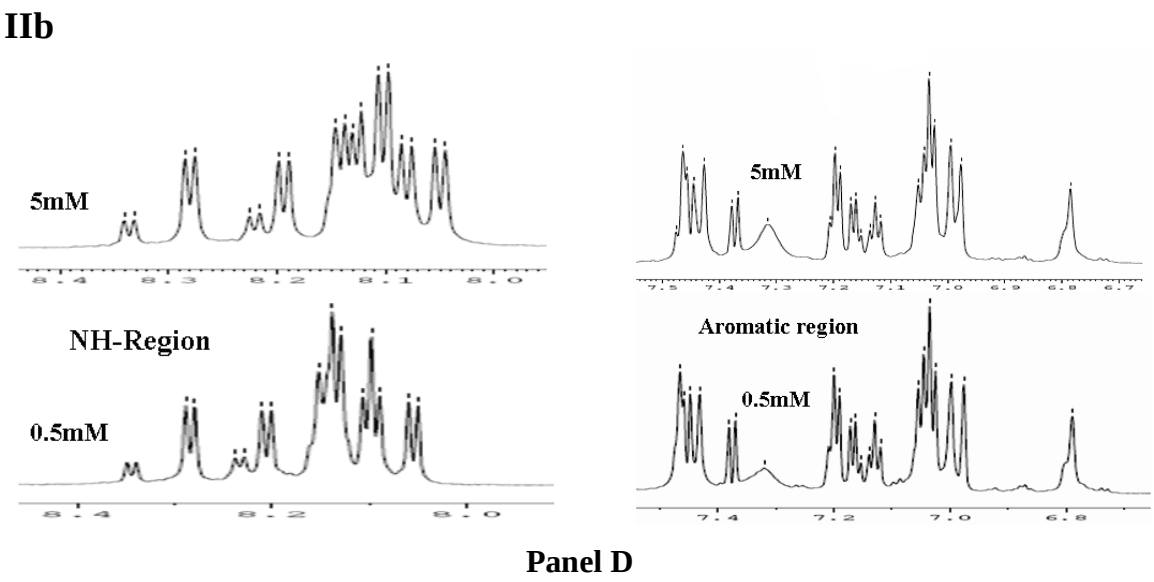


Figure S10: Aromatic-H and peptide-NH regions of ^{1}H-NMR spectra of specific peptides recorded at 0.5 mM and 5.0 mM concentrations in 90% D_2O-H_2O. Neutral peptides **Ia** and **IIa** (Panels A and C) have the chemical shifts affected by the dilution while the charged peptides **Ib** and **IIb** (Panels B and D) have the chemical shifts largely unaffected by the dilution.

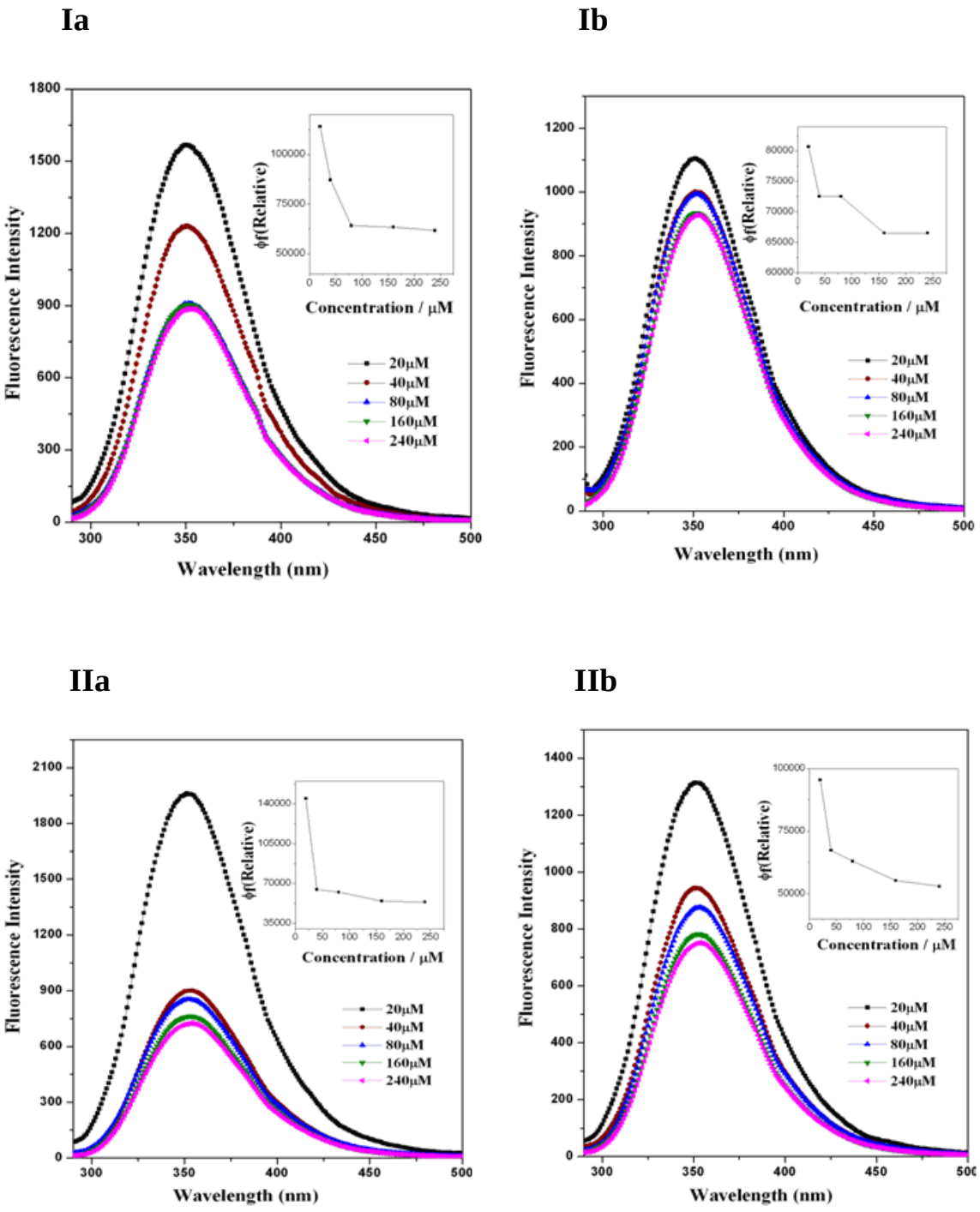
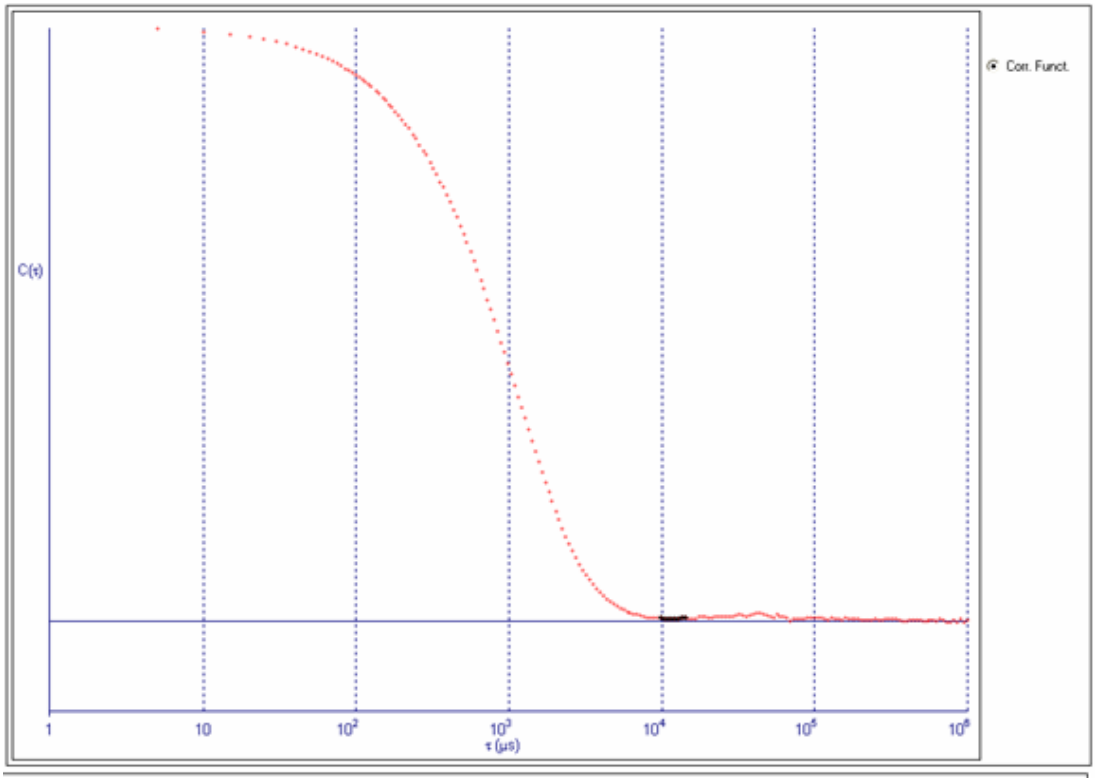
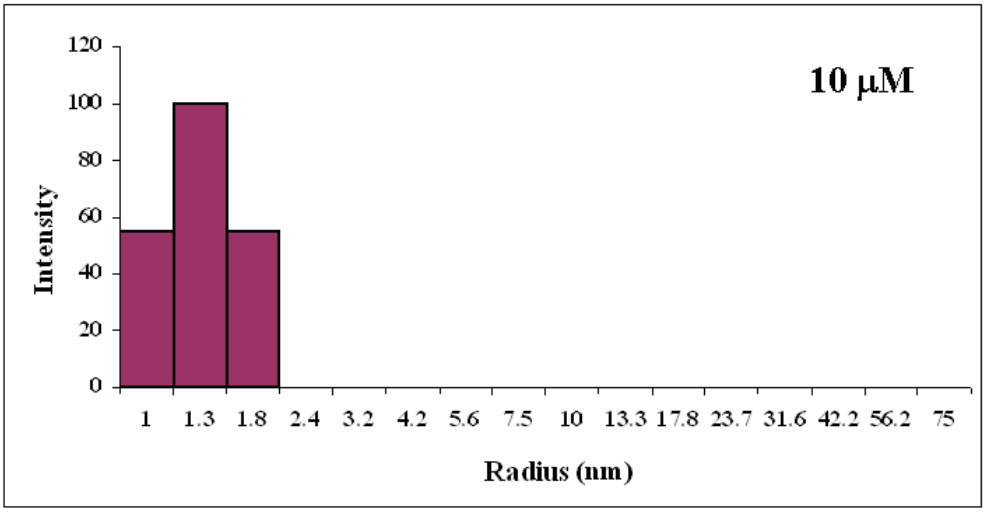


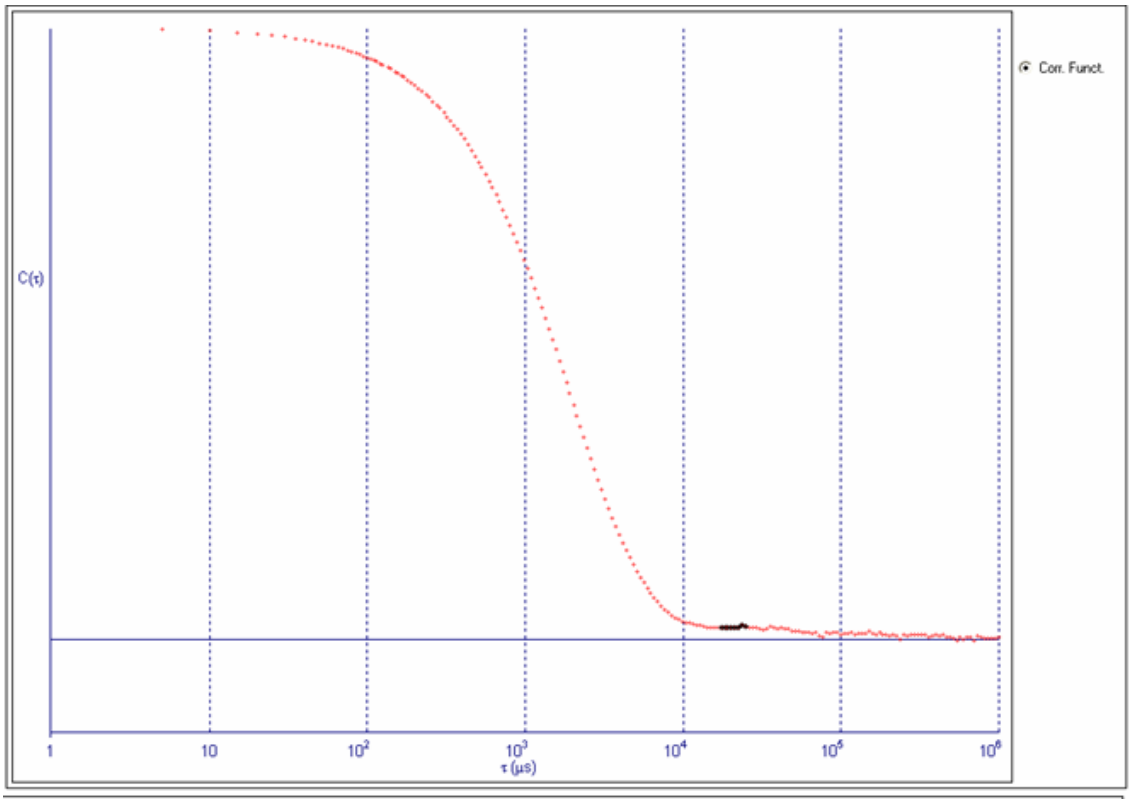
Figure S11: Fluorescence emission spectra as a function of concentration and plot of intensity against concentration (insets) with water as solvent.

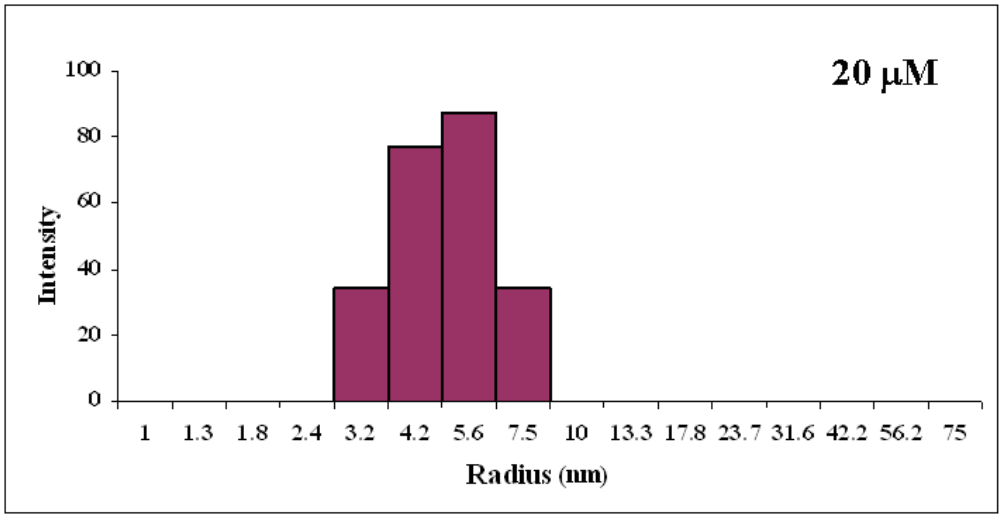
Ib (10 μM)



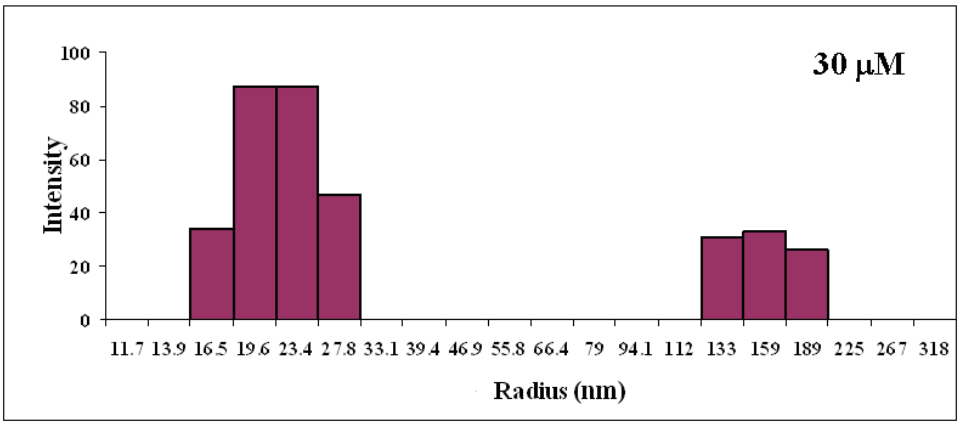


Ib (20 μM)

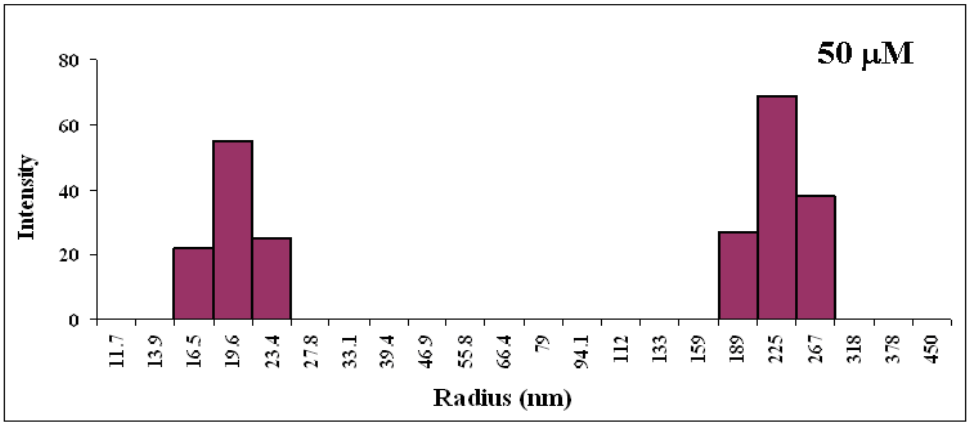




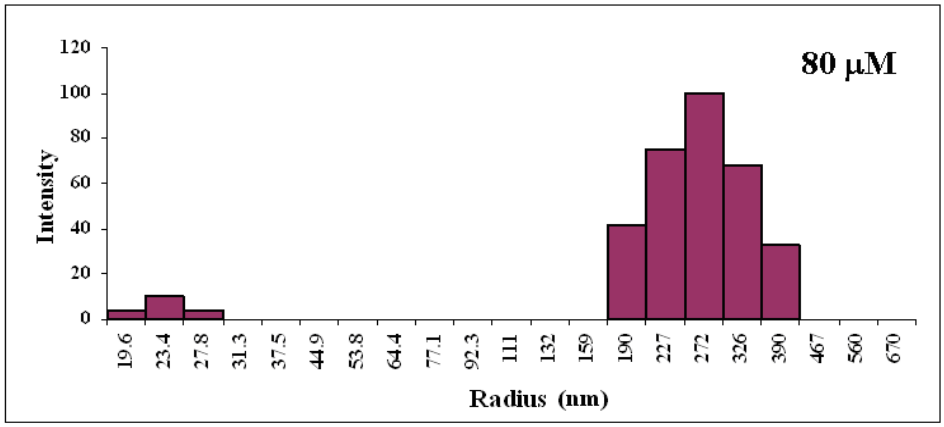
Ib (30 μM)



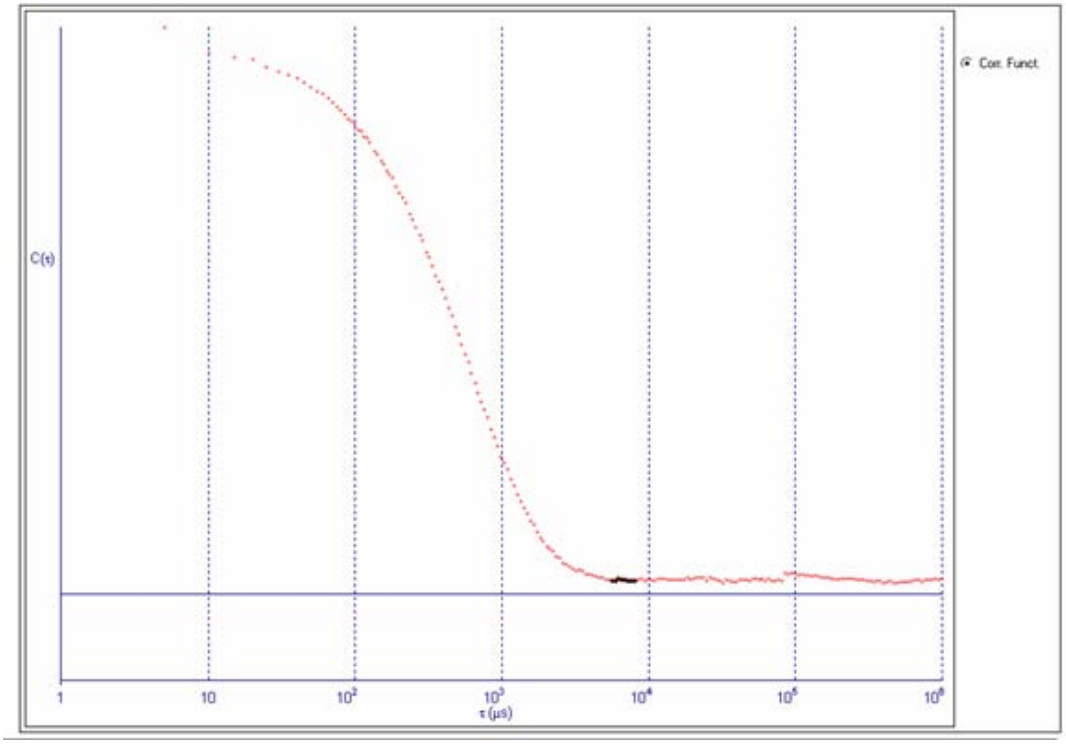
Ib (50 μM)

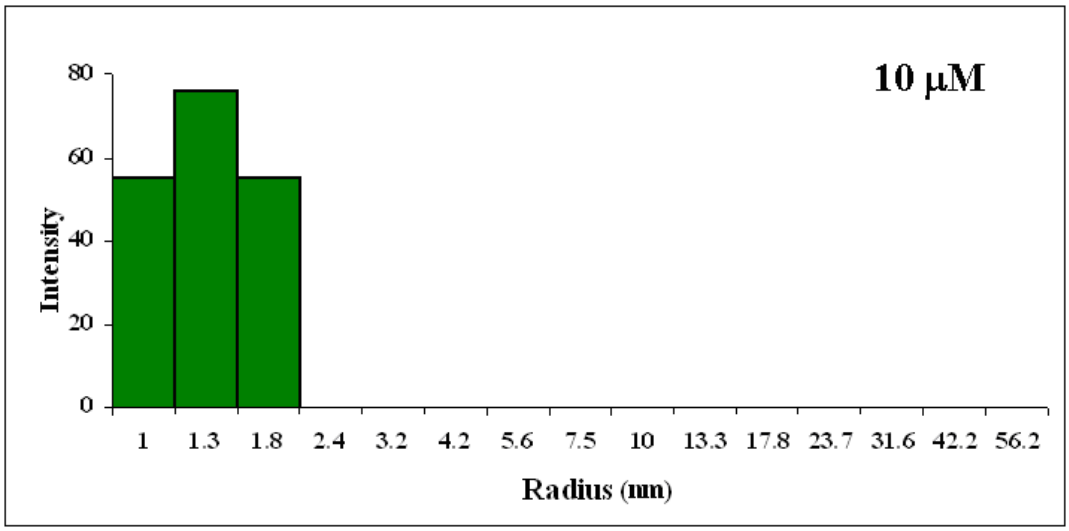


Ib (80 μM)

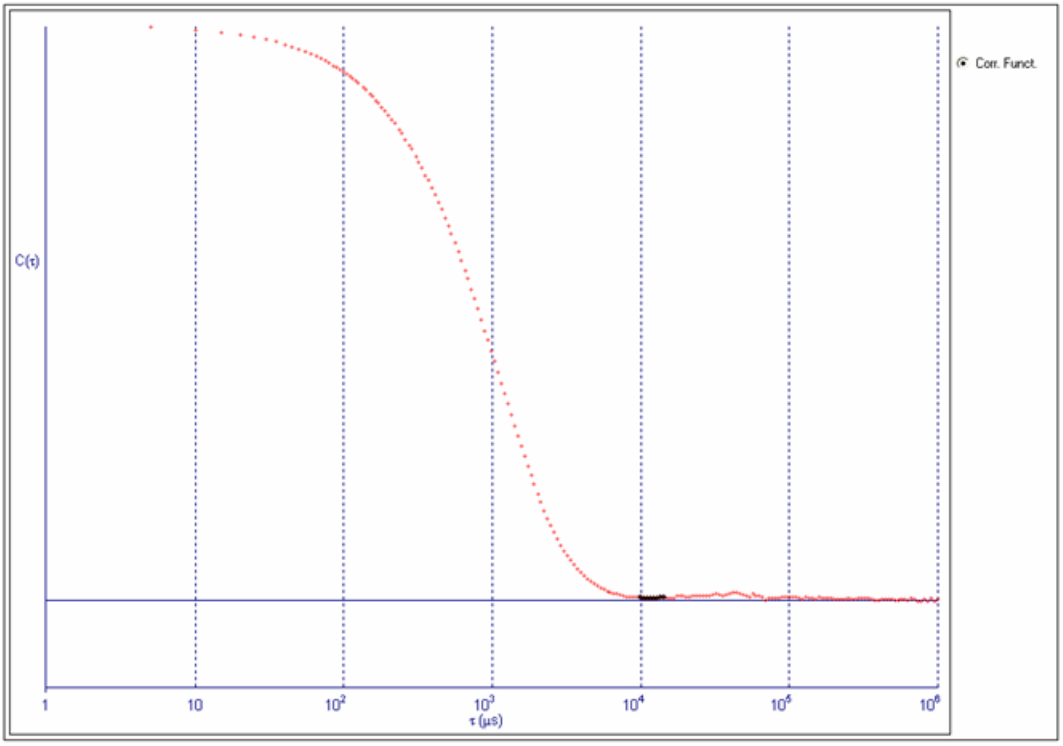


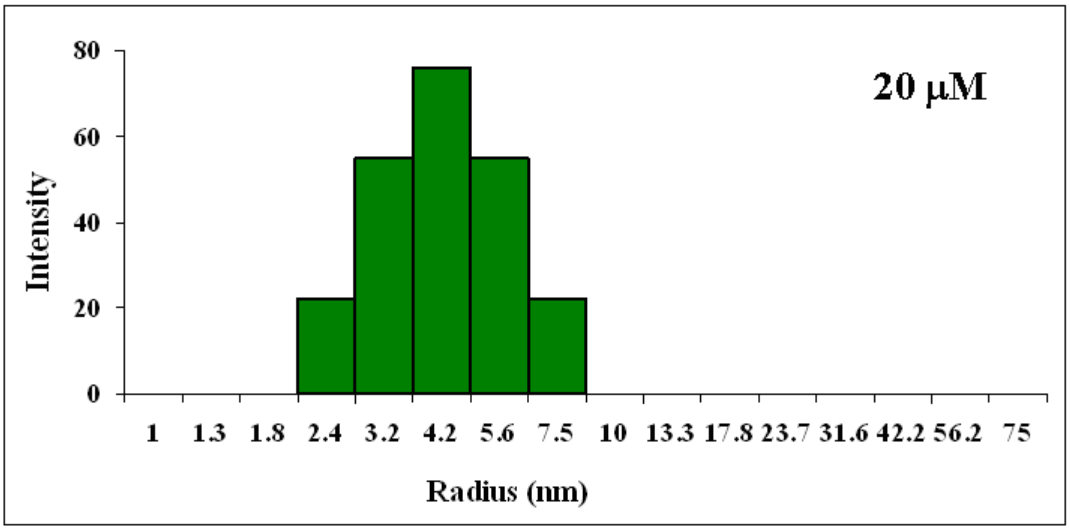
IIb (10 μM)



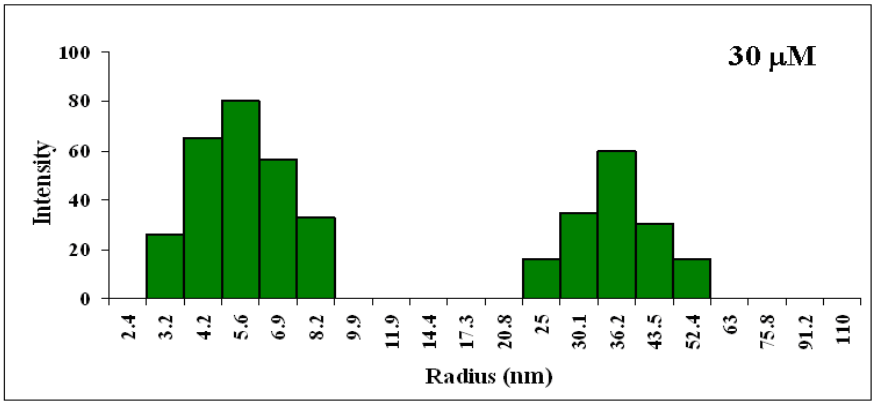


IIb (20 μM)

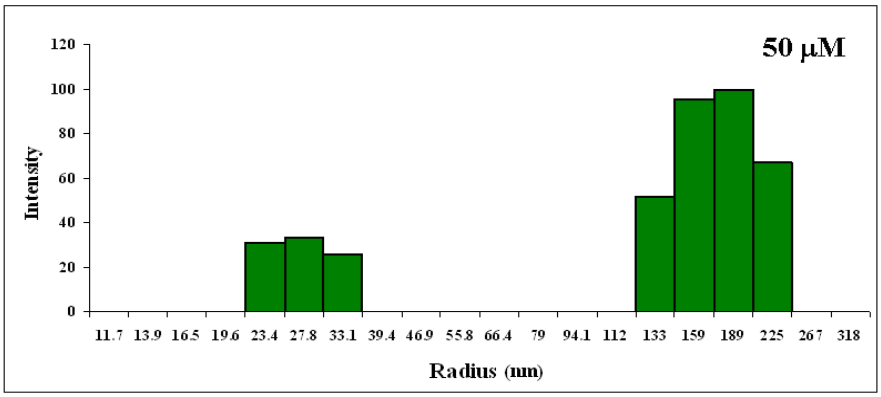




IIb (30 μM)



IIb (50 μM)



IIb (80 μM)

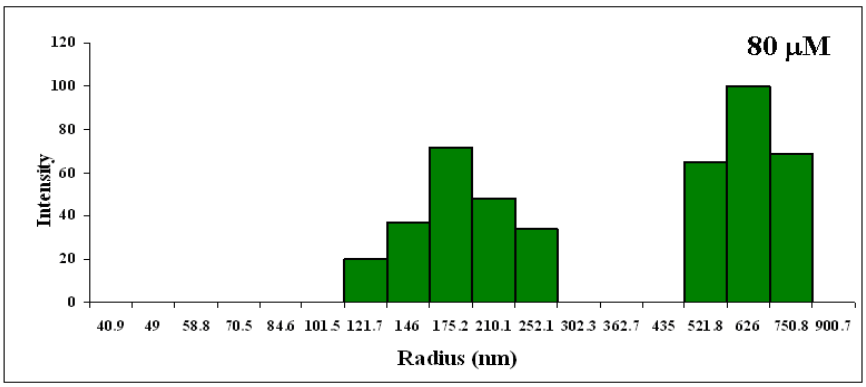


Figure S12: According to Dynamic Light Scattering (DLS) charged peptides **Ib** and **IIb** have the spread in particle size distribution at 10, 20, 30, 50 and 80 μM concentrations indicating the possible formation of higher order aggregates

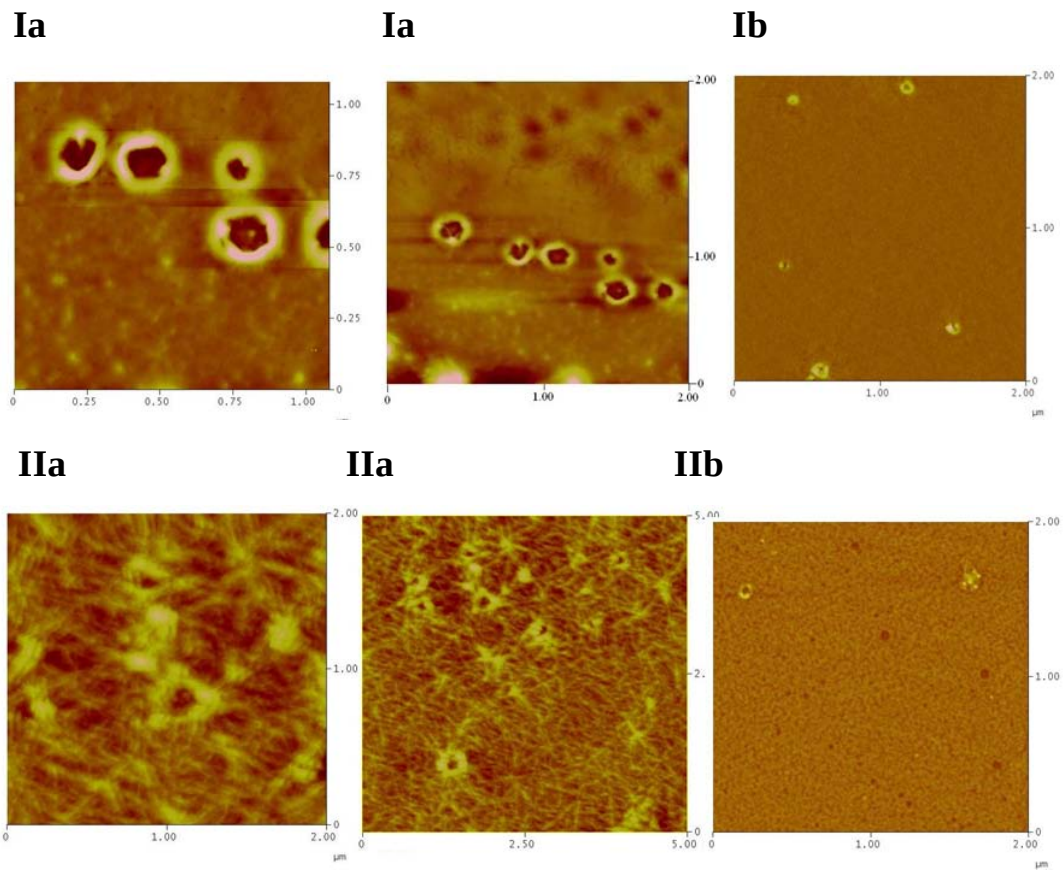


Figure S13: Low resolution AFM images show that the neutral peptides **Ia** and **IIa** tend to populate in higher order assemblies in greater proportion than the charged peptides **Ib** and **IIb**.

Table S1: Eighteen highest ranked folds showing SymmDock scores in VA and SR variations and the other specific properties of the folds.

Serial No.	Fold	SymmDock Score	Cα RMSD (nm)	Rg (nm)	Hydrogen Bonds	Number of Clusters
	VA	2556		0.50	2	
100	SR	2960	0.24	0.50	3	1
	VA	2554		0.49	3	
1004	SR	2870	0.47	0.72	1	11
	VA	2542		0.51	1	
1042	SR	2838	0.23	0.51	3	5
	VA	2536		0.51	3	
1044	SR	2810	0.38	0.59	3	4
	VA	2446		0.49	2	
1051	SR	2798	0.08	0.52	3	5
	VA	2442		0.49	4	
1053	SR	2784	0.21	0.50	3	1
	VA	2412		0.48	3	
1192	SR	2778	0.20	0.50	3	1
	VA	2388		0.50	2	
1229	SR	2774	0.48	0.56	2	7
	VA	2252		0.52	2	
1238	SR	2766	0.24	0.57	2	4
	VA	2212		0.50	5	
1243	SR	2704	0.12	0.49	2	3
	VA	2160		0.50	3	
1301	SR	2686	0.22	0.50	2	4
	VA	2136		0.49	6	
1313	SR	2402	0.37	0.52	3	3
	VA	2114		0.51	2	
1355	SR	2324	0.21	0.52	3	7
	VA	*3012*		*0.41*	*4*	
1487	*SR*	*3032*	*0.18*	*0.44*	*6*	*3*
	VA	2922		0.38	2	
1877	SR	2982	0.31	0.38	5	7
	VA	2824		0.38	6	
1890	SR	2968	0.12	0.37	3	1
	VA	2776		0.39	3	
2049	SR	2964	0.20	0.44	4	5
	VA	*2512*		*0.38*	*4*	
2161	*SR*	*3414*	*0.15*	*0.38*	*5*	*2*

Table S2: ϕ, ψ values in peptides Ia and IIa

Sequence Position	Ia	ϕ	ψ	Sequence Position	IIa	ϕ	ψ
1	**Ser**	90	-139	1	**Thr**	-58	90
2	**Glu**	72	-105	2	**Lys**	-135	90
3	**Thr**	142	-84	3	**Trp**	64	-120.
4	**Trp**	-108	-46	4	**Phe**	-64	-31
5	**Leu**	-91	136	5	**Asn**	-80	116
6	**Lys**	68	-49	6	**Glu**	91	-17
7	**Tyr**	-138	125	7	**Val**	98	53
8	**Val**	109	-126	8	**Ala**	132	-142

Table S3: Number of clusters enumerated to 1.5 nm cutoff populating 50 ns MD trajectories of specific folds and assemblies, % population in top five clusters and the top cluster in each trajectory.

Type of Fold	State of Oligomer	Total Number of Clusters	Population in Top Cluster (%)
Ia	Monomer	**60**	**31**
	Pentamer	***450***	**25**
Ib	Monomer	**198**	**13**
	Pentamer	**898**	**5**
IIa	Monomer	**145**	**37**
	Pentamer	**620**	**21**
IIb	Monomer	***5***	**97**
	Pentamer	**741**	**5**

Table S4: Chemical shift assignments for peptide Ib.

Sequence Position	Residue	NH	C $^{\alpha}$-H	C $^{\beta}$-H	C $^{\gamma}$-H, C $^{\delta}$-H	Others
	Ac					2.5
1	Ser	8.47	4.53	3.86,3.92		
2	Gln	8.29	4.33	1.96, 2.08	2.33	7.01
3	Thr	8.34	4.40	4.28	1.37	
4	Trp	8.31	4.67	3.19,3.19		2H = 7.18 4H =7.58 5H = 7.18 6H = 7.23 7H =7.40 NH=10.10
5	Leu	8.22	4.17	1.84,1.85		CH_3,CH_3=0.93,0.92
6	Lys	7.53	4.83	1.67,1.67	1.42,1.42, 1.79,1.79	C $^{\varepsilon}$ = 2.98,2.98 NH_3 = 6.86
7	Tyr	8.19	4.54	2.50,2.50		2,6 H =7.07 3,5 H = 6.79
8	Val	7.90	3.98	1.99	1.14, 0.88	
	NH_2	7.1				

Table S5: $^{3}J_{NH\text{-}C\alpha H}$ coupling constants of peptide Ib.

Sequence Position	Residue	J
1	Ser	6.4
2	Gln	7.2
3	Thr	6.4
4	Trp	7.2
5	Leu	7.2
6	Lys	8.8
7	Tyr	6.4
8	Val	8.0

Table S6: NOEs and calculated inter-proton distances in peptide Ib.

Sequence Position	Residue	Atom	Sequence Position	Residue	Atom	Calculated Distance (Å)
1	Ser	HN	1	Ser	HA	3.79
8	Val	HN	8	Val	HA	3.7
5	Leu	HN	4	Trp	HA	2.74
8	Val	HN	8	Val	HA	2.74
4	Trp	HN	4	Trp	HB2	2.8
7	Tyr	HN	7	Tyr	HB2	3.14
4	Trp	HN	4	Trp	HA	3.21
5	Leu	HA	5	Leu	HB2	3.52
4	Trp	HZ2	4	Trp	HD1	2.4
6	Lys	HE2	6	Lys	HE3	2.4
4	Trp	HH2	4	Trp	HD1	2.49
7	Tyr	HE2	7	Tyr	HE1	2.65

4	Trp	HE3	4	Trp	HD1	2.77
5	Leu	HA	4	Trp	HE3	4.6
2	Gln	HB3	2	Gln	HG2	3.42
2	Gln	HB2	2	Gln	HG2	3.55
4	Trp	HB2	4	Trp	HE3	3.92
6	Lys	HG2	6	Lys	HD2	3.92
6	Lys	HB2	6	Lys	HE2	4.17
5	Leu	HD22	4	Trp	HE1	4.26
8	Val	HB	8	Val	HG21	2.4
5	Leu	HN	5	Leu	HD12	2.83
5	Leu	HB2	5	Leu	HD22	2.86
6	Lys	HE2	15	Tyr	HN	4.42