Electronic Supplementary Information

A Small Peptide Stereochemically Customized as a Globular Fold with a Molecular Cleft†

Soumendra Rana,[a] Bijoy Kundu,[b] and Susheel Durani[a*]

Department of Chemistry,[a] Indian Institute of Technology Bombay, Mumbai-400076 and Medicinal Chemistry Division,[b] Central Drug Research Institute, Lucknow-226001, India.

Sequence Design and Validation

The sequence design was approached rationally selecting β-sheet favoring amino acid residues to obtain a heterogeneous sequence for the ease of experimental validation by spectroscopic techniques like NMR and Fluorescence but of amphipathic character to achieve a fold with a H_2O exposed hydrophilic exterior and a non-polar interior topology. The core of the design approach is guided by the fundamental stereospecificity of folding in polypeptides, as captured in the Ramachandran diagram. The positions for stereochemical inversions from L-to-D in the sequence were so chosen to obtain an amphipathic kind of molecular fold with a hydrophobic core and a surface largely composed of hydrophilic residues, illustrated in Figure S-1, C. The residues Thr(8)-Ser(13), Asn(15)-Ser(6), and Lys(17)-Glu(4) were positioned as the cross-strand neighbors to obtain inter-strand locking required for a stable β-hairpin kind of motif. The conformational ordering of our designed sequence was validated by MD against "Trp-cage" (PDB code = 1L2Y, Figure S-2, B), a recently reported globular mini-protein, achieved by rational design of comparable polypeptide length.[5] Parallel unrestrained MD simulations in explicit H_2O for 2ns at 300K performed under identical conditions, reflected favorably on the promise of our designed heterotactic peptide (Figure S-2, A) as a protein compared with the simulated properties of the "Trp-cage", providing the confidence for experimental validation of the design based on synthesis and spectroscopic characterization.

Experimental Section

All amino acids with suitable protecting groups were purchased from Novobiochem (Switzerland). The 20-residue peptide was synthesized manually in a 10ml syringe by standard solid phase Fmoc-chemistry with TBTU / HOBT or HOBT / DIC method with Rink Amide AM resin as the solid support, at a 0.74mMol/gm scale. The resin was deblocked at the N-terminus by 30% Piperidine–DMF with continuous shaking for 0.5 hours. Kaiser and ninhydrin tests were used for monitoring the completion of coupling or de-blocking of amino groups. The N-terminus was capped

by acetylation (-$NCOCH_3$) using Ac_2O: DIPEA: DMF in a ratio of 1:2:20 for 5 hours. The final peptide was obtained as a primary amide group (-$CONH_2$) at C-terminus. Simultaneous cleavage of the peptide from the resin bed and deprotection of side-chains were achieved with reagent-K [TFA: H_2O: Phenol: Thioanisole: EDT in a ratio of 82.5:5:5:5:2.5] over 1.5-4.0 hours at room temperature with continuous shaking. After removal of TFA by passing N_2 gas, the peptide was precipitated with cold diethyl ether (anhydrous), lyophilized and then purified by HPLC over RP-18 silica column (10μm, 10mm × 250mm; Merck) using ACN \ H_2O \ 0.1%TFA \ 0-100% gradient**.** The homogeneity was checked on a RP-18e silica analytical column (5μm, 4mm × 250mm; Merck) using the above eluents. The presence of the desired peptide was confirmed by MALDI-TOF Mass spectrometry (Mcalc: 2157 Da; M_{obs}: 2158 Da; and M_{Na+}: 2180 Da, which confirmed that methionine sulphur is not oxidized) and subsequently by ^{1}H NMR analysis. Two sets of ^{1}H NMR spectra recorded in H_2O at 2.5mM and 0.25mM concentration did not show any significant changes in chemical shifts or in sharpness of the resonances, evidencing absence of aggregation. Concentration dependent CD measurements were recorded at 298 K in 0.5mM CH_3CO_2Na buffer at pH ≅ 5. No significant change in molar ellipticity was observed for the peptide in the 40μM to 100μM concentration range, suggesting the absence of aggregation. The peptide concentration was determined using molar extinction coefficient of Trp (~5600 at 278nm)**.** Spectra were recorded in Far UV region on a JASCO J-810 CD spectroploarimeter equipped with a peltier controller [0.2cm quartz cell, 700μl volume, scanning rate of 100nm/min, time constant 1.0s] over an average of 5 accumulations with step resolution of 0.5 nm, and bandwidth of 2nm. Phase sensitive TOCSY (mixing time 80ms, 32scans, 480 experiments) and NOESY (mixing time 200ms, 48 scans, 480 experiments) spectra were recorded at 298K in pure ^2H_2O or 90%H_2O / 10% ^2H_2O for the peptide ≅ 2-3mM concentration at pH ≅ 5 on a VARIAN INOVA 600 MHz NMR. A 0.1mM concentration of 3-(trimethyl-silyl)-1-propanesulfonicacid (DSS) was used for an internal chemical shift reference. Solvent suppression in H_2O was achieved either by presaturation or by WATERGATE pulse sequence provided by VARIAN. The data processing was done on a SGI workstation equipped with VNMR and XWIN-NMR or on a windows workstation equipped with FELIX software. Typically a sine-squared window function, phase shifted by 70° was applied in both dimensions with the data zero filled to 2K × 1K or else 4K × 4K points before Fourier transformation. The fluorescence measurements were on a Perkin Elmer LS-55 spectrofluorimeter equipped with a standard PMT, properly stabilized and after base line corrections over a single accumulation in 20mM CH_3CO_2Na buffer, pH ≅ 5 [1ml cell, $\lambda_{excitation}$ = 295 nm, $\lambda_{emission}$ = 297-450, excitation and emission slits = 3nm, scan speed = 100nm / min] at room

temperature. The working concentration (20μM) of the HPLC purified peptide and NATA (SIGMA chemicals) was calibrated by OD measurements using molar extinction coefficient of Trp (~5600 at 278nm). All the quenching experiments involving KI and Methionine were performed in the above said manner with suitable quencher concentration adjusted by dilution of stock solution of molar concentration. The initial coordinates of the peptide were energy minimized both in vacuum and in H_2O for 2000 steps using steepest descent algorithm. The molecule was kept at 300K in the center of a cubic box in explicit solvent for 100ps for achievement of thermal equilibration followed by production MD run. All MD simulations were performed on P-IV linux platform using gromos-96 force field in the GROMACS package.

Figure S-1. A) The 20-residue isotactic model sequence. The six-long arm hydrophobic residues (side-chains shown in CPK) are far apart from each other, featuring a planar canonical β-hairpin morphology. The residues for stereochemical inversions from L-to-D (side-chains shown in ball and stick) are in blue. B) The β-hairpin after the first pair (7, 14) of stereochemical inversion. C) The β-hairpin after the second pair (3, 18) of stereochemical inversion, resulting in a 20-residue quasi-globular mini-protein with six-long arm hydrophobic residues (side-chains shown in CPK) forming a hydrophobic core and a largely hydrophilic surface. The stereochemical inversions from L-to-D (side-chains shown in blue ball and stick) has resulted in engineering the flat β-hairpin to a potential boat shaped molecular cleft, creating a scope for diversified functional design.

Figure S-2. A) Snapshots of molecular cleft taken from MD trajectory at an interval of 500ps. B) Snapshots of "Trp-cage" taken from MD trajectory at an interval of 500ps.

Figure S-3. Summary of the results obtained from the 2ns MD simulations at 300K. A) Trp-cage. B) Molecular Cleft. C) Total number of hydrogen bonds populated over the trajectory. $R_{peptide}$ = radius of gyration of the peptide, R_{core} = radius of gyration defined by residues forming the hydrophobic core and C-alpha = RMSD of C_α atoms.

Figure S-4. MALDI-MS and analytical HPLC profile of boat shaped peptide molecular cleft.

Figure S-5. Concentration dependent mean residue ellipticity of the peptide in H_2O at 298K.

Figure S-6. The blurrogram of 10 best DYANA generated NMR structures are with mean global backbone root mean square deviation (**RMSD**) over residue 2-19 of 0.39 ± 0.24 Å and mean global all heavy atoms RMSD of 1.11 ± 0.28 Å in H_2O.

Figure S-7. The average number of hydrogen bonds populated in H_2O over the MD trajectory of 16 ns is 8 compared with 9 hydrogen bonds expected in the ideal boat shaped peptide motif. The

calculation was done in explicit solvent using gromos-96 force field in GROMACS 3-1 package in a cubic box at 300K.

Figure S-8. Plot of potential energy of the boat shaped peptide molecular cleft against time over the MD trajectory in H_2O.

Figure S-9. Time dependent proximity plot of C^{β} atoms of the core forming residues [Rcore4 = Leu(16), Trp(5), Leu(20) and Met(1) and Rcore6 = Val(12), Ile(9), Leu(16), Trp(5), Leu(20) and Met(1)] against the radius of gyration of the peptide over the entire MD trajectory. The purple line is the cut-off distance for the core residues in hydrophobic contact with each other.

Figure S-10. Fluorescence spectrum of the peptide against NATA displaying 4nm blue shift and a two-fold lower intensity of fluorescence.

Figure S-11. Quenching of the peptide fluorescence by the external quencher KI over the concentration range 0-0.2M.

Figure S-12. Quenching of the NATA fluorescence by the external quencher KI over the concentration range 0-0.2M.

Figure S-13. Stern-Volmer plots of the KI quenching experiments for the peptide and NATA at room temperature. Since the Stern-Volmer plots were deviating from their ideal behavior of linearity a statistical linear fit was generated by Microsoft excel sheet (Windows XP) over all the points for NATA with R^2 value = 0.97 and Peptide with R^2 value = 0.91 to satisfy the bimolecular quenching equation $I_0 / I = 1 + K_{sv}[Q]$, whereI_0 = Fluorescence intensity in the absence of external quencher, I = Fluorescence intensity in the presence of external quencher, Q = Concentration of the quencher and K_{sv} = Stern-Volmer constant, calculated from the slope of the line.

Figure S-14. Reduction of fluorescence intensity of NATA by free methionine over the concentration ranges 0-0.3M establishing it as a quencher of tryptophan fluorescence.

Figure S-15. Stern-Volmer plots of the methionine quenching experiments for NATA, illustrating its quenching efficiency as a non-standard quencher against the standard quencher KI.

Figure S-16. The % burial of Trp(5) side-chain in the peptide against Gly-Trp-Gly model calculated by NACCESS over the MD trajectory in H_2O at 300K.

Table S-1. The chemical shift assignments of the peptide in H_2O achieved by the combined use of TOCSY and NOESY. The observed sequential cross peaks in NOESY helped in unambiguous assignments of the spin systems

Table S-2. The distance restraints used for structure calculations by DYANA generated from unambiguous NOEs (checked by Distance check module in DYANA) from adjacent and non-adjacent residues in H_2O. 5 van der Waals inter-strand distance constraints were also used for the structure calculations (not shown).

Table S-3. Average values of backbone dihedral angles (ϕ, ψ) with standard deviation over 11 structures collected from 50 random structures generated by DYANA simulation and 50 structures collected at an interval of 300ps over the entire molecular dynamics trajectories in H_2O. The program MolMol was used for superposition and calculation of the dihedrals.

(A) **(B)** **(C)**

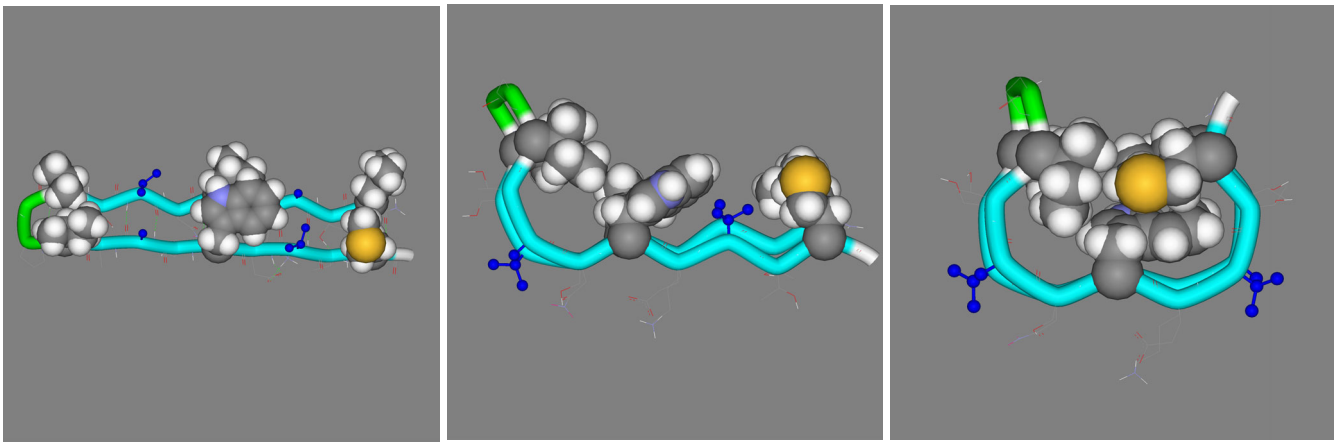


Figure S-1

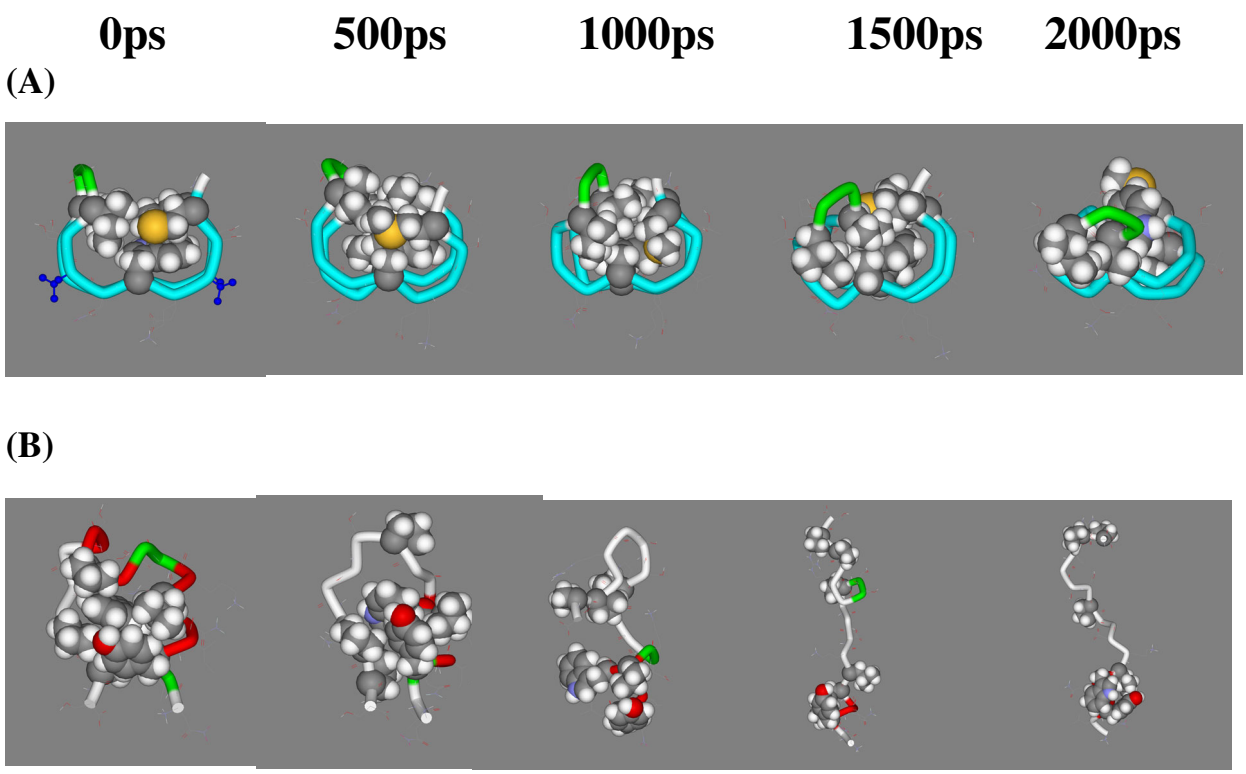
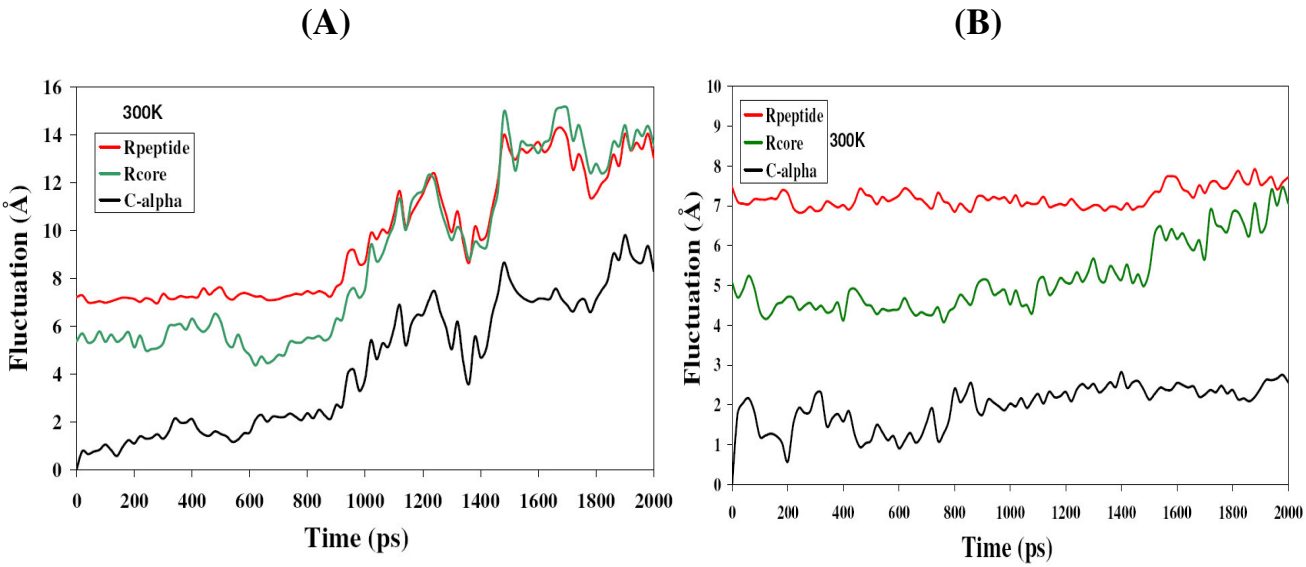

0ps
500ps
1000ps
1500ps
2000ps
(A)
(B)

Figure S-2


(A)
(B)
300K
Rpeptide
Rcore
C-alpha
Fluctuation (Å)
Time (ps)
0
200
400
600
800
1000
1200
1400
1600
1800
2000

(C)

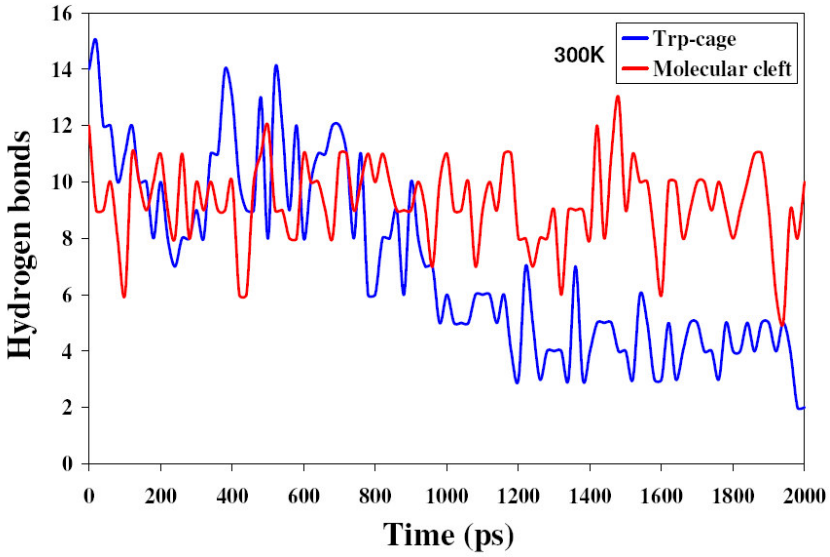


Figure S-3

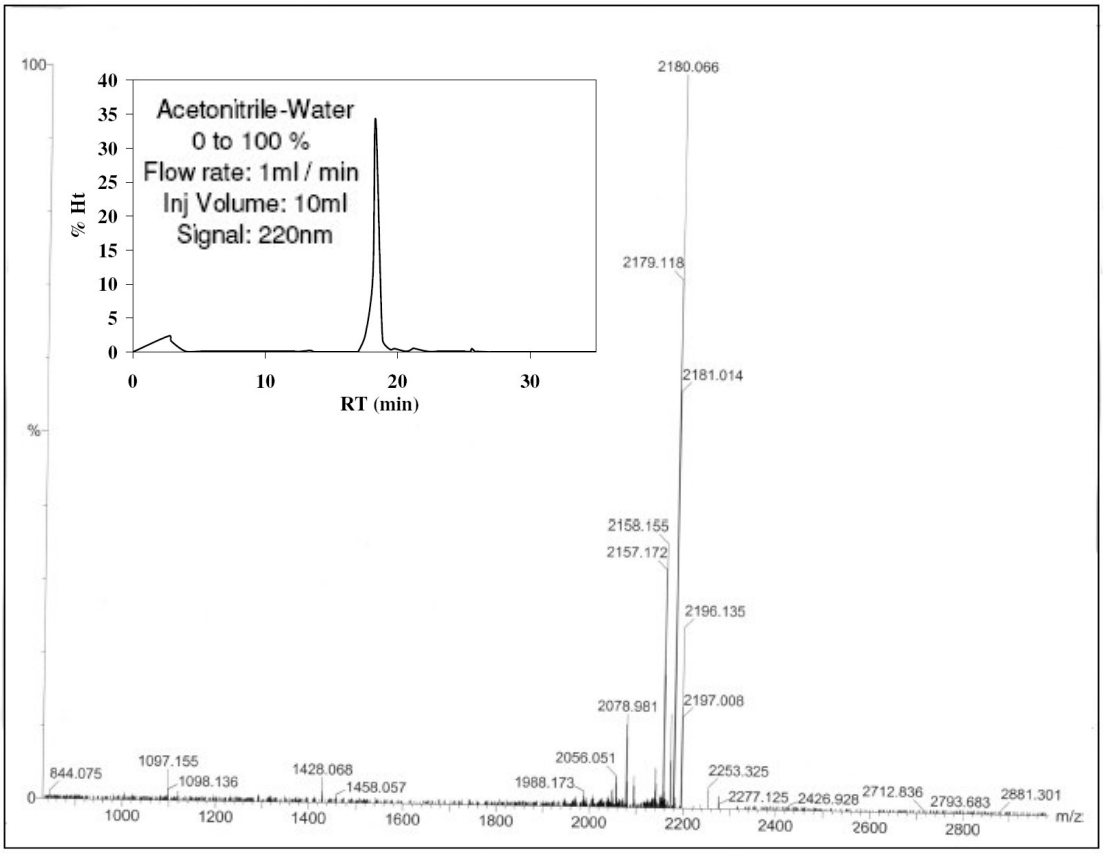


Figure S-4

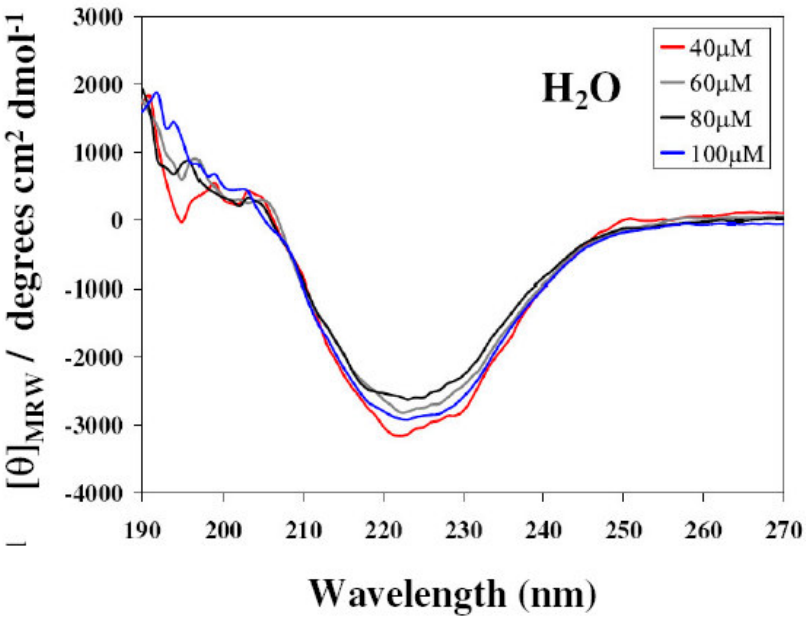


Figure S-5

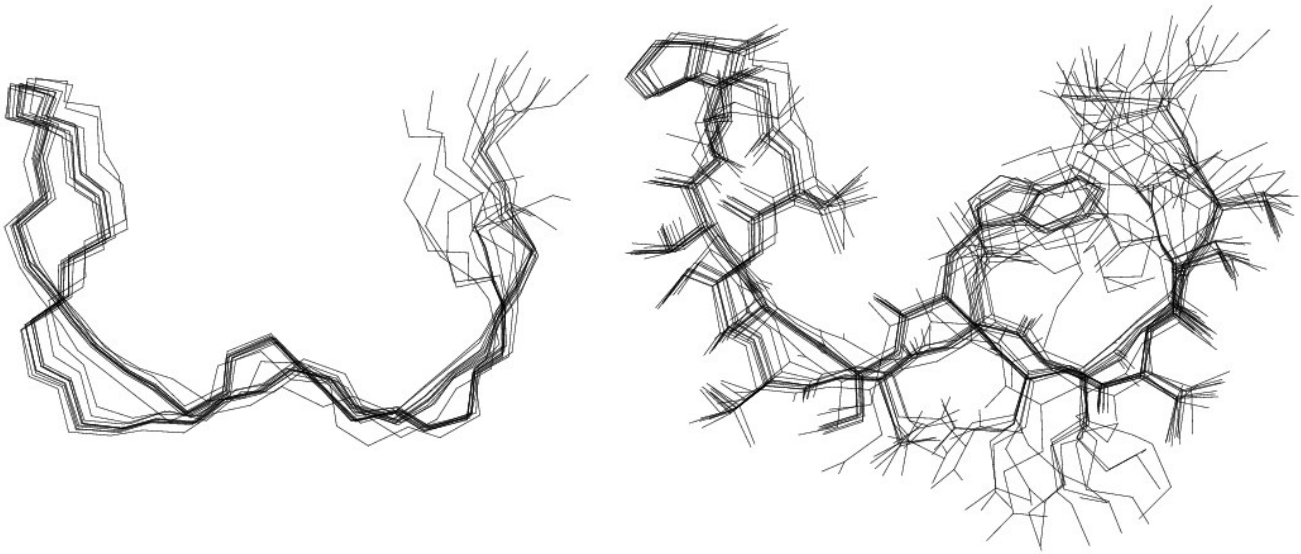


Figure S-6

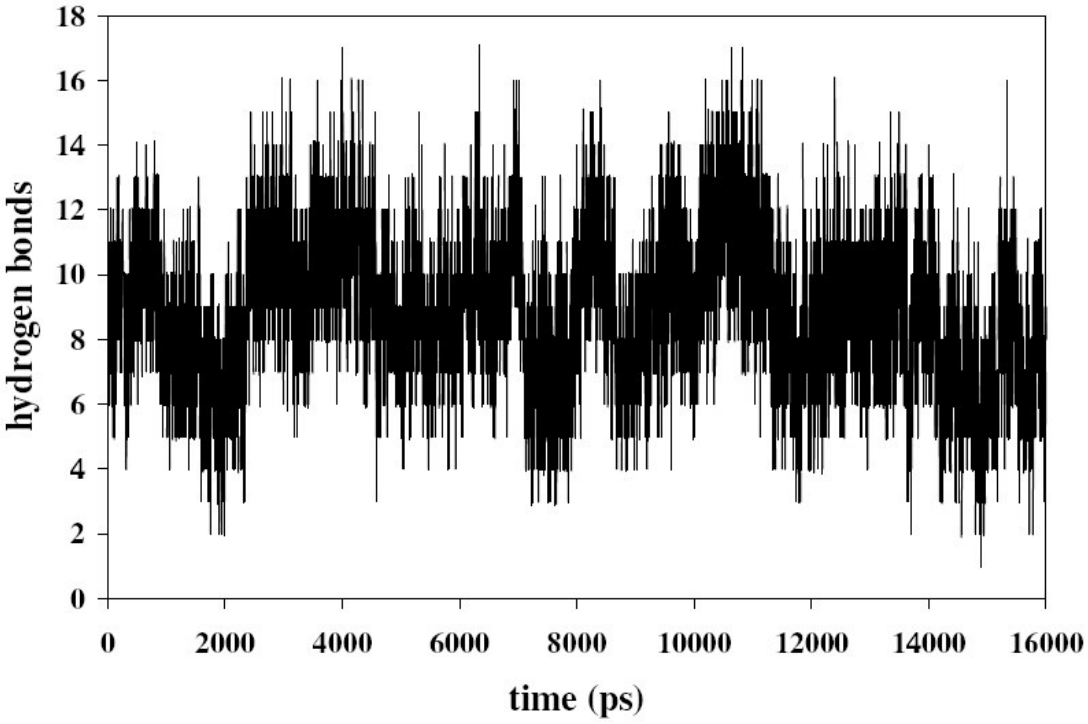


Figure S-7

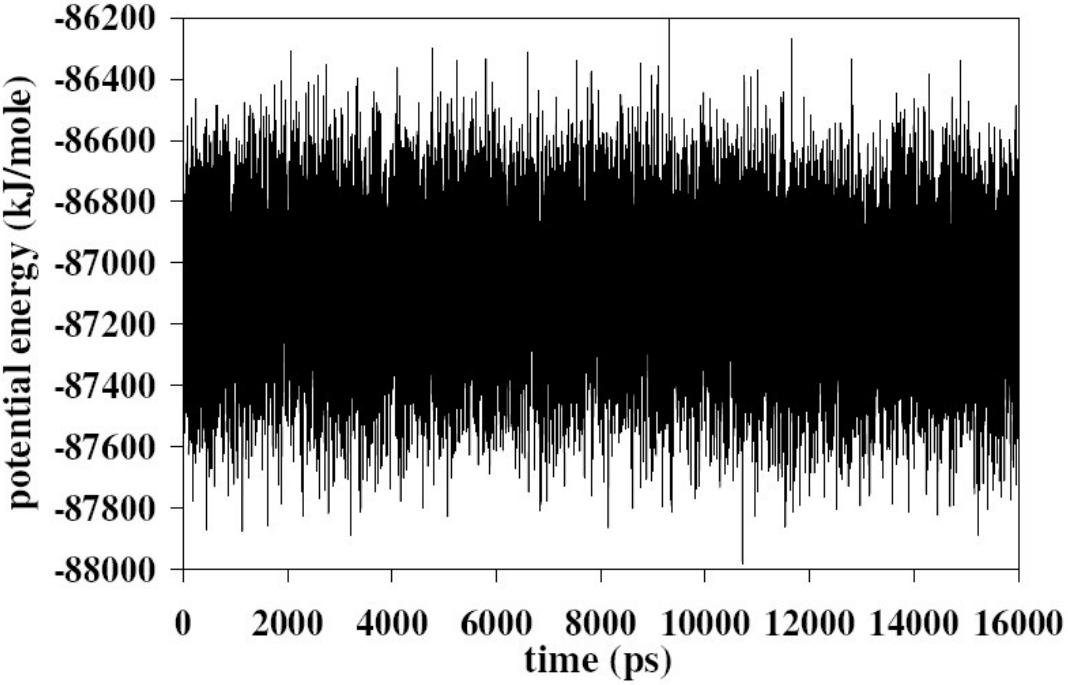


Figure S-8

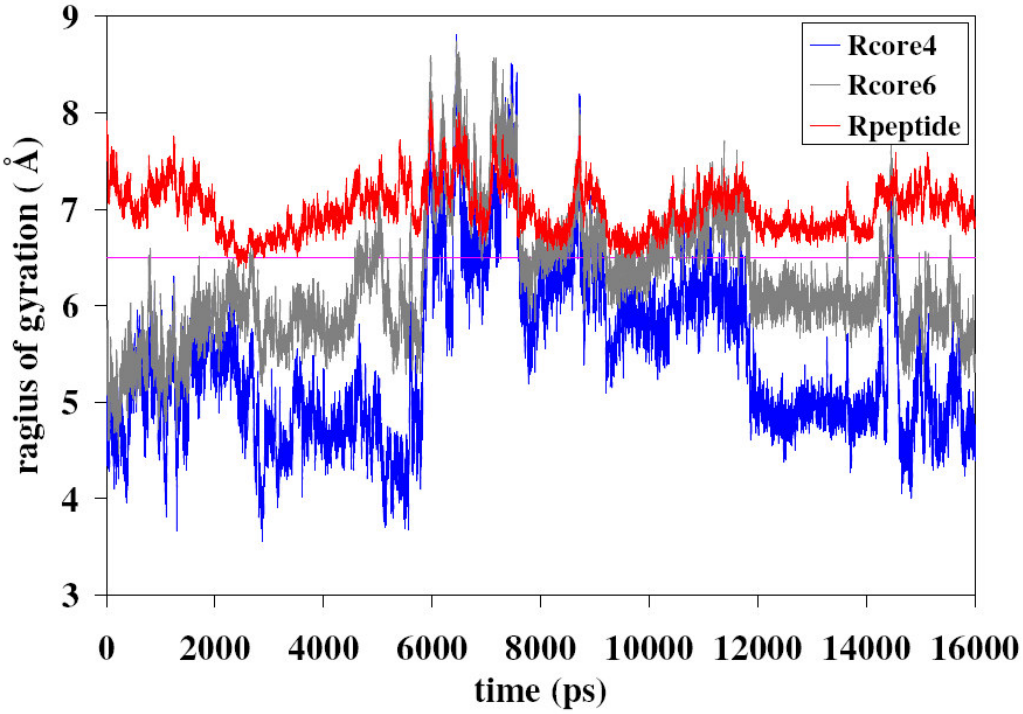


Figure S-9

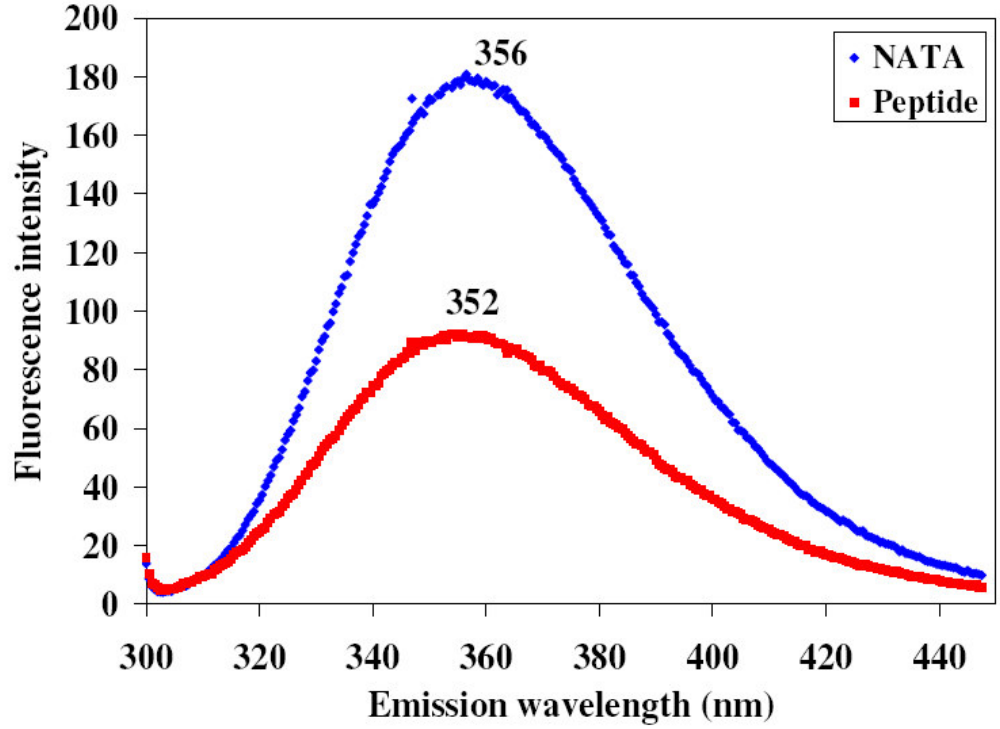


Figure S-10

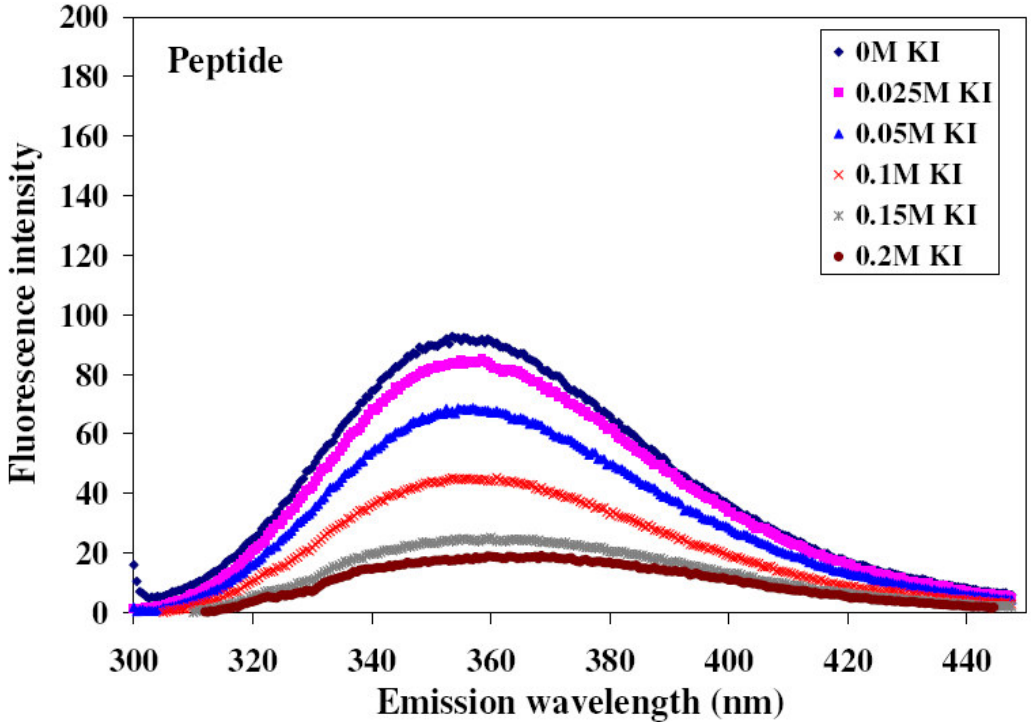


Figure S-11

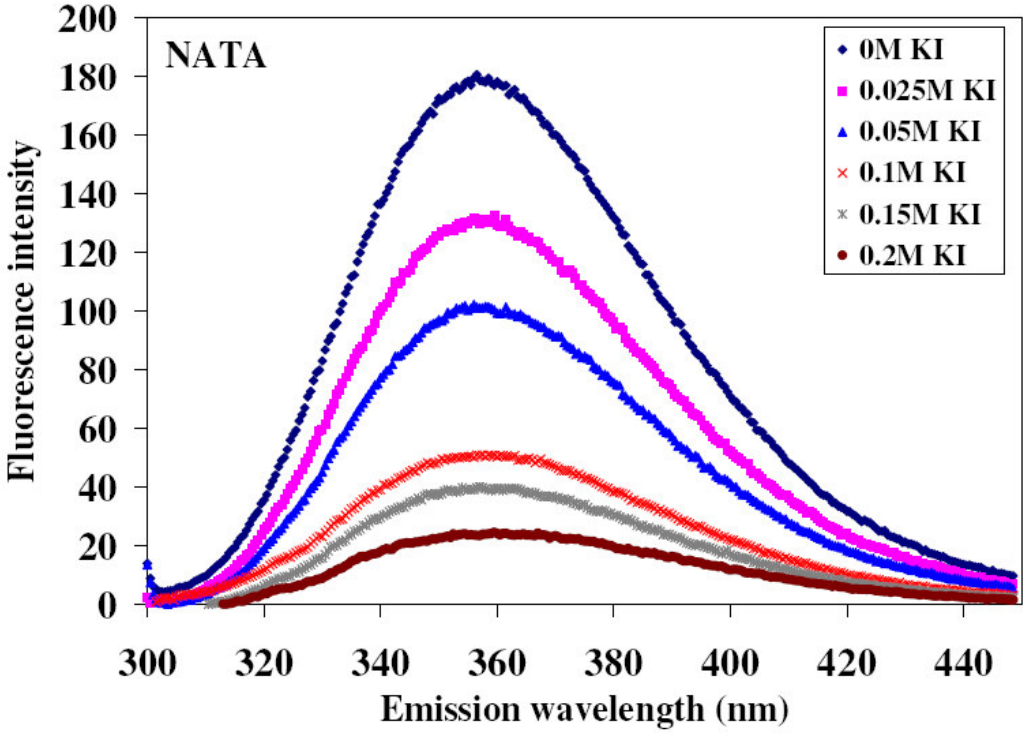


Figure S-12

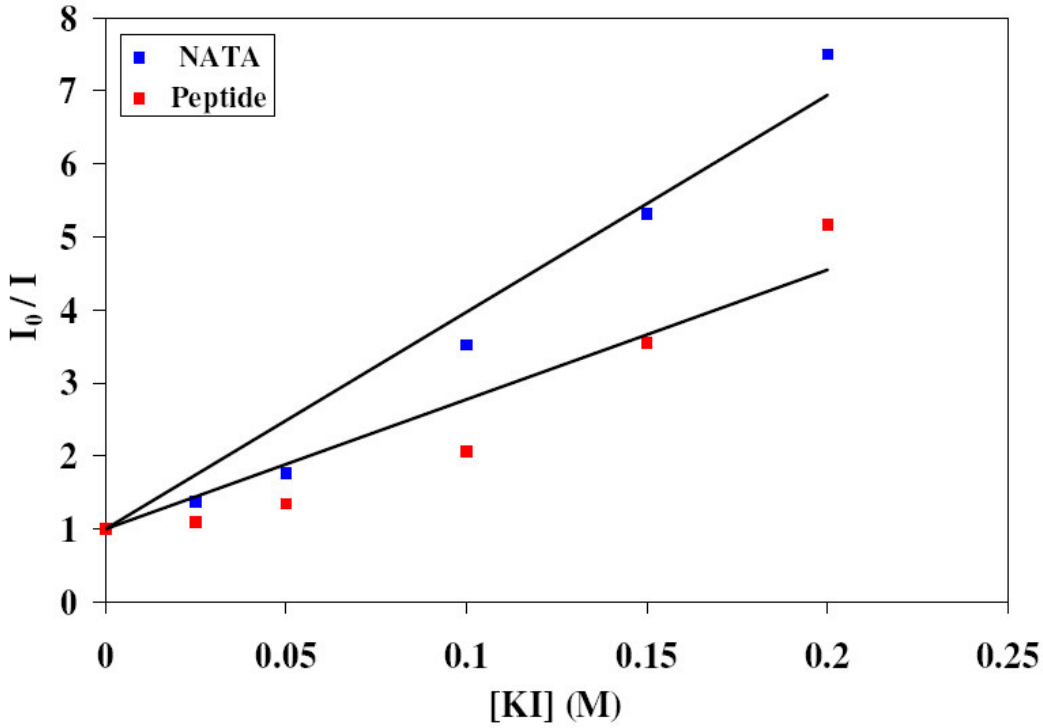


Figure S-13

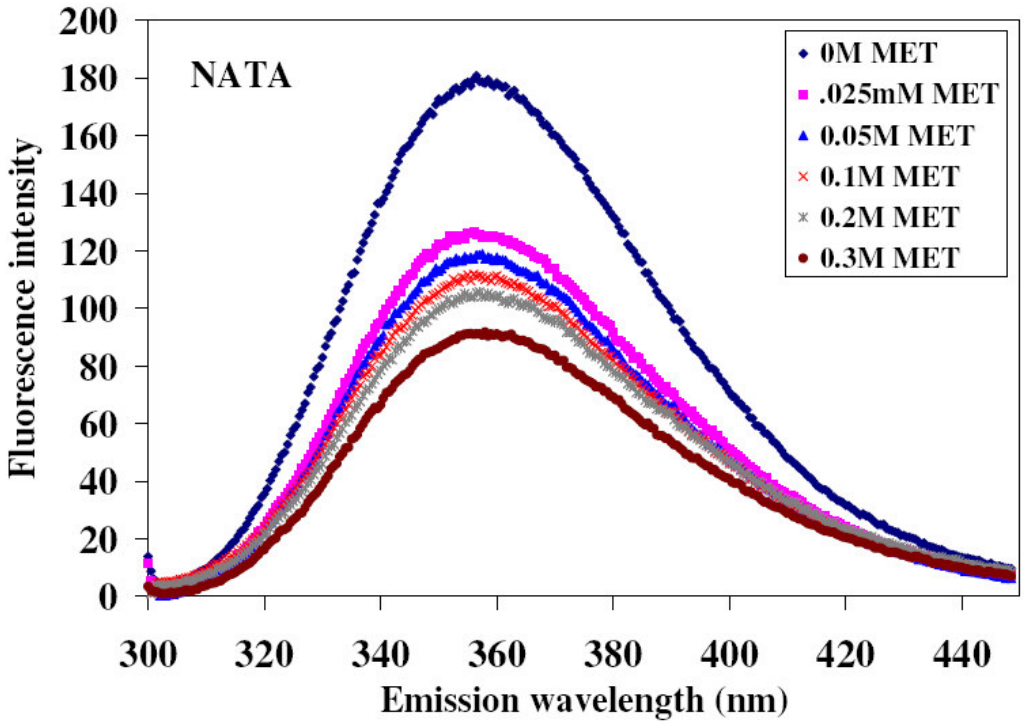


Figure S-14

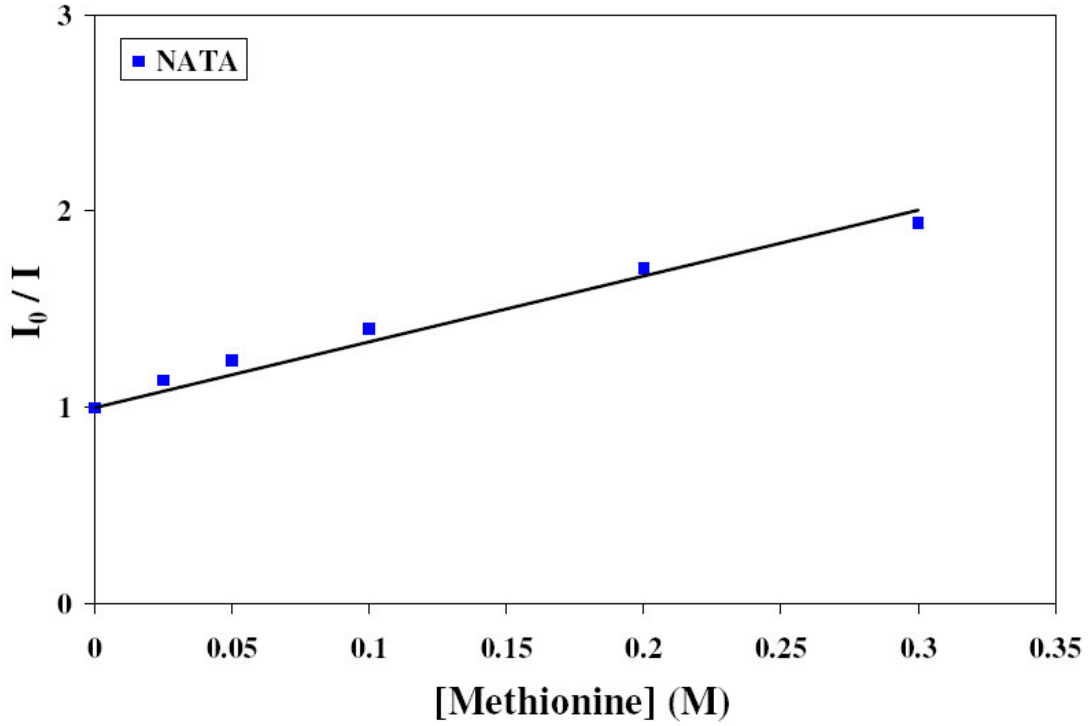


Figure S-15

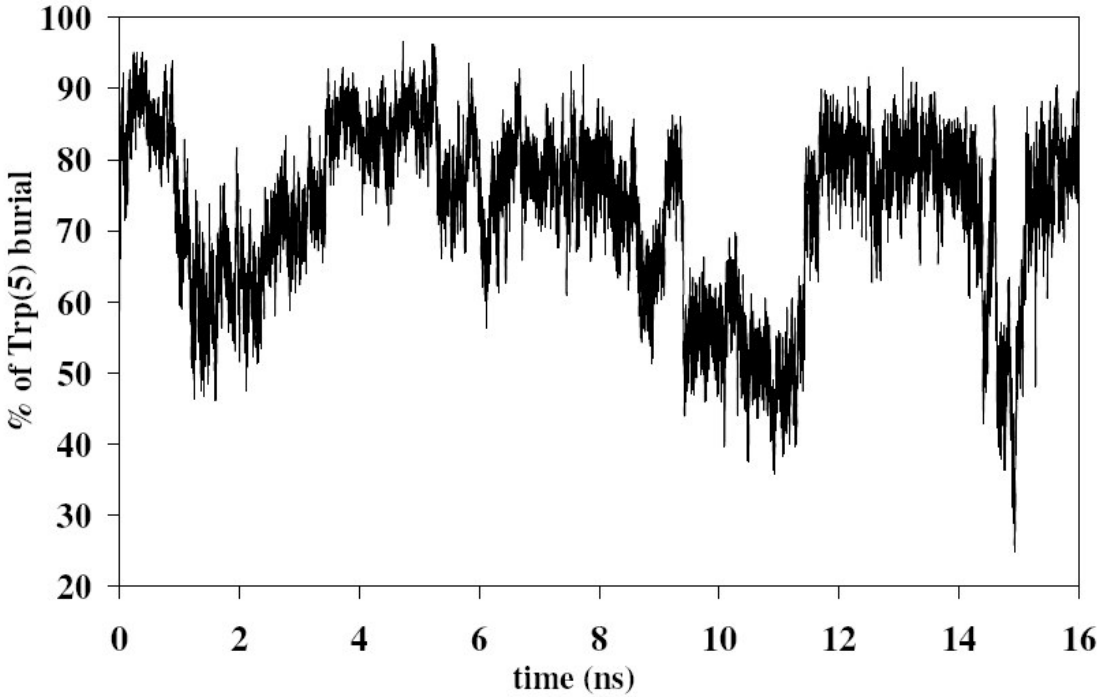


Figure S-16

Amino-acids	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$ / $C^{\delta}H$	Others
Met(1)	8.26	4.40	1.88, 1.39	2.43, 1.96	2.50
Thr(2)	8.21	4.35	4.27	1.81	
D-Val(3)	8.72	4.36	1.61		0.88
Glu(4)	8.60	4.24	1.77	2.21	
Trp(5)	8.11	4.67	3.31, 3.20	7.24	10.10, 7.58, 7.04, 7.18, 7.40
Ser(6)	7.98	4.38	3.72		
D-Ala(7)	8.17	4.44	1.11		
Thr(8)	8.34	4.40	4.28	1.37	
Ile(9)	8.35	4.44	2.17	1.80, 1.66	0.82, 0.81
D-Pro(10)		4.25	1.81, 1.65	1.43, 1.36	2.96
Gly(11)	7.96	4.21, 4.00			
Val(12)	8.08	4.10	2.04		0.90
Ser(13)	8.33	4.60	3.81		
D-Val(14)	7.90	4.21	2.10		0.90
Asn(15)	8.58	4.36	2.84, 2.68		7.54, 6.88
Leu(16)	8.22	4.38	1.73, 1.64	1.54	0.89, 0.82
Lys(17)	8.43	4.50	2.60, 2.52	1.96, 2.10	3.80, 7.50
D-Ala(18)	8.10	4.46	1.10		
Thr(19)	8.14	4.31	4.07	1.40	
Leu(20)	8.15	4.43	1.70	1.60	0.90

Table S-1

NOEs between the residues in H2O	Calculated NOE distance
1 MET HN 1 MET QG	4.24
1 MET HN 19 THR HA	3.39
2 THR HA 5 TRP HE3	4.26
2 THR HA 5 TRP HZ3	4.42
2 THR HA 18 ALAD HN	2.93
2 THR HB 20 LEU HN	2.40
2 THR HB 5 TRP HD1	5.07
3 VALD HN 4 GLU- HA	3.45
3 VALD HA 5 TRP- HZ2	4.38
4 GLU- HA 16 LEU HN	2.40
4 GLU- HA 18 ALAD HN	3.64
4 GLU- QG 17 LYS+ QG	5.98
5 TRP HN 6 SER QB	4.92
5 TRP HA 16 LEU HN	4.60
5 TRP HD1 9 SER HB	5.50
5 TRP HZ2 20 LEU HN	4.97
5 TRP HH2 19 THR HA	3.89
5 TRP HZ2 19 THR HA	4.23
5 TRP HH2 20 LEU HN	4.63
6 SER QB 16 LEU HN	5.35
7 ALAD HN 15 ASN HA	2.46
7 ALAD HA 8 THR HN	3.14
7 ALAD QB 9 ILE HN	5.78
7 ALAD QB 14 VALD HN	5.10
8 THR HN 9 ILE QG1	4.80
8 THR HA 14 VALD HN	3.73
8 THR HB 14 VALD HN	3.30
8 THR QG2 14 VALD HN	5.75
11 GLY HN 13 SER QB	4.95
12 VAL HB 13 SER HN	3.58
13 SER HN 14 VALD HN	3.30
13 SER HN 14 VALD HA	2.90
13 SER HN 14 VALD QQG	6.94
14 VALD HN 15 ASN HN	3.48
14 VALD HB 15 ASN HN	4.17
15 ASN QB 16 LEU HN	4.42
16 LEU QQD 20 LEU HG	5.46
18 ALAD HA 19 THR HN	2.83
19 THR HB 20 LEU HN	3.45
19 THR QG2 20 LEU HN	3.92

Table S-2

Residues	Minimum energy DYANA structure (ϕ, ψ)	Average DYANA structure (ϕ, ψ)	Average MD structure (ϕ, ψ)
Met (1)	ψ = -177.7	ψ = 160 ± 48	ψ = 136 ± 42
THR (2)	-60.2, 122.5	-141 ± 49, 116 ± 7	-82 ± 60, 138 ± 40
(D) VAL (3)	161.2, -58.3	162 ± 3, -65 ± 14	120 ± 30, -122 ± 26
GLU (4)	-172.8, 111.3	-170 ± 3, 124 ± 25	-78 ± 29, 89 ± 82
TRP (5)	-100.9, 56.0	-110 ± 21, 57 ± 3	-81 ± 92, 148 ± 94
SER (6)	-63.8, 106.4	-59 ± 4, 108 ± 20	-117 ± 19, 131 ± 23
(D) ALA (7)	163.9, -165.5	160 ± 9, -162 ± 6	97 ± 38, -119 ± 19
THR (8)	-162.6, 72.5	-152 ± 21, 68 ± 7	-99 ± 26, -43 ± 44
ILE (9)	-112.1, 83.3	-104 ± 17, 98 ± 29	-91 ± 30, 122 ± 16
(D) PRO (10)	ψ = 81.4	ψ = 62 ± 67	ψ = -137 ± 51
GLY (11)	131.8, -76.4	152 ± 39, -75 ± 78	-90 ± 47, 151 ± 64
VAL (12)	-69.4, 160.0	-71 ± 34, 150 ± 17	-100 ± 63, 37 ± 73
SER (13)	-163.1, 54.8	-165 ± 2, 48 ± 13	-128 ± 34, 119 ± 32
(D) VAL (14)	162.1, -82.0	140 ± 50, -77 ± 11	107 ± 28, -127 ± 22
ASN (15)	-164.5, 165.0	-159 ± 8, 165 ± 0.9	-104 ± 27, 132 ± 30
LEU (16)	-154.2, 89.4	-149 ± 10, 106 ± 40	-105 ± 55, 100 ± 87
LYS (17)	-100.1, 101.3	-112 ± 33, 109 ± 14	-101 ± 82, 128 ± 35
(D) ALA (18)	157.1, -161.8	162 ± 2, -157 ± 11	99 ± 27, -126 ± 24
THR (19)	-74.3, 138.1	-100 ± 22, 142 ± 7	-86 ± 30, 92 ± 77
LEU (20)	ϕ = -60.1	ϕ = -140 ± 31	ϕ = -112 ± 48

Table S-3