Supplementary information

## The 'CαNN' motif: an intrinsic lover of sulfate and phosphate ions

Tridip Sheet, Raja Banerjee*

Maulana Abul Kalam Azad University of Technology, West Bengal
(Formerly Known as West Bengal University of Technology)
BF-142,Sector-1, Salt Lake, Kolkata- 700064, INDIA

Telephone No: 91-33-23341021/1031; Fax: 91-33-2334-1030
Email: ban_raja@yahoo.com, banraja10@gmail.com

Short title: Anion-recognition peptide motif

* Correspondence to: Raja Banerjee, Maulana Abul Kalam Azad University of Technology, West Bengal, (Formerly known as West Bengal University of Technology) BF-142, Sector-1, Salt Lake, Kolkata-700064, India. E-mail: ban_raja@yahoo.com, banraja10@gmail.com
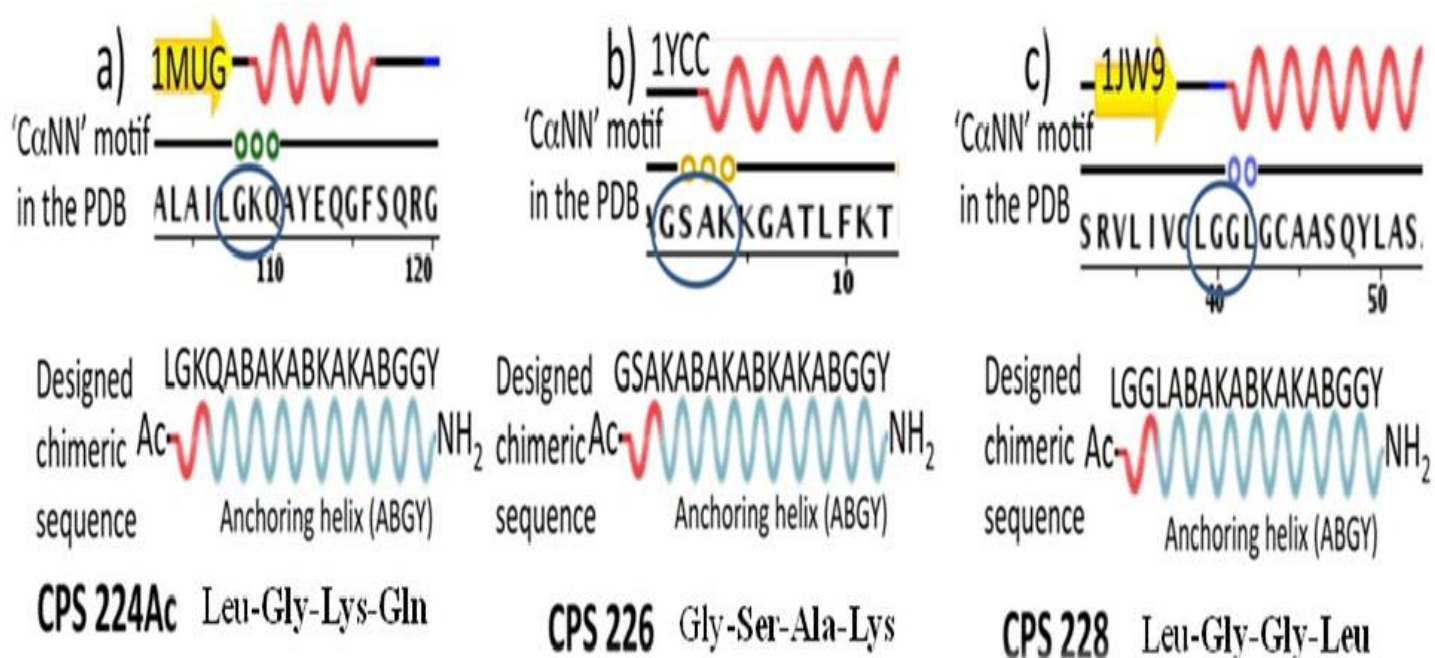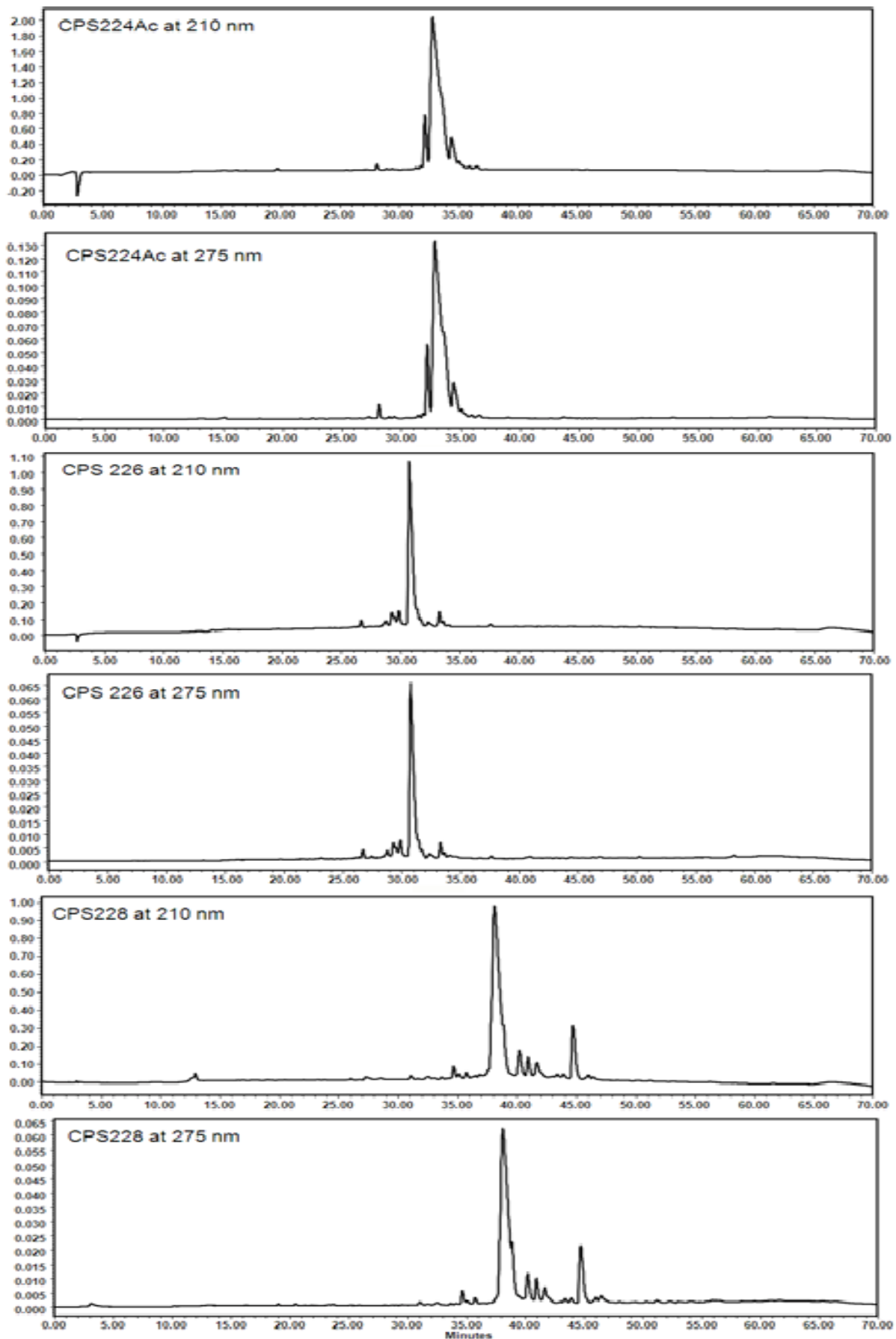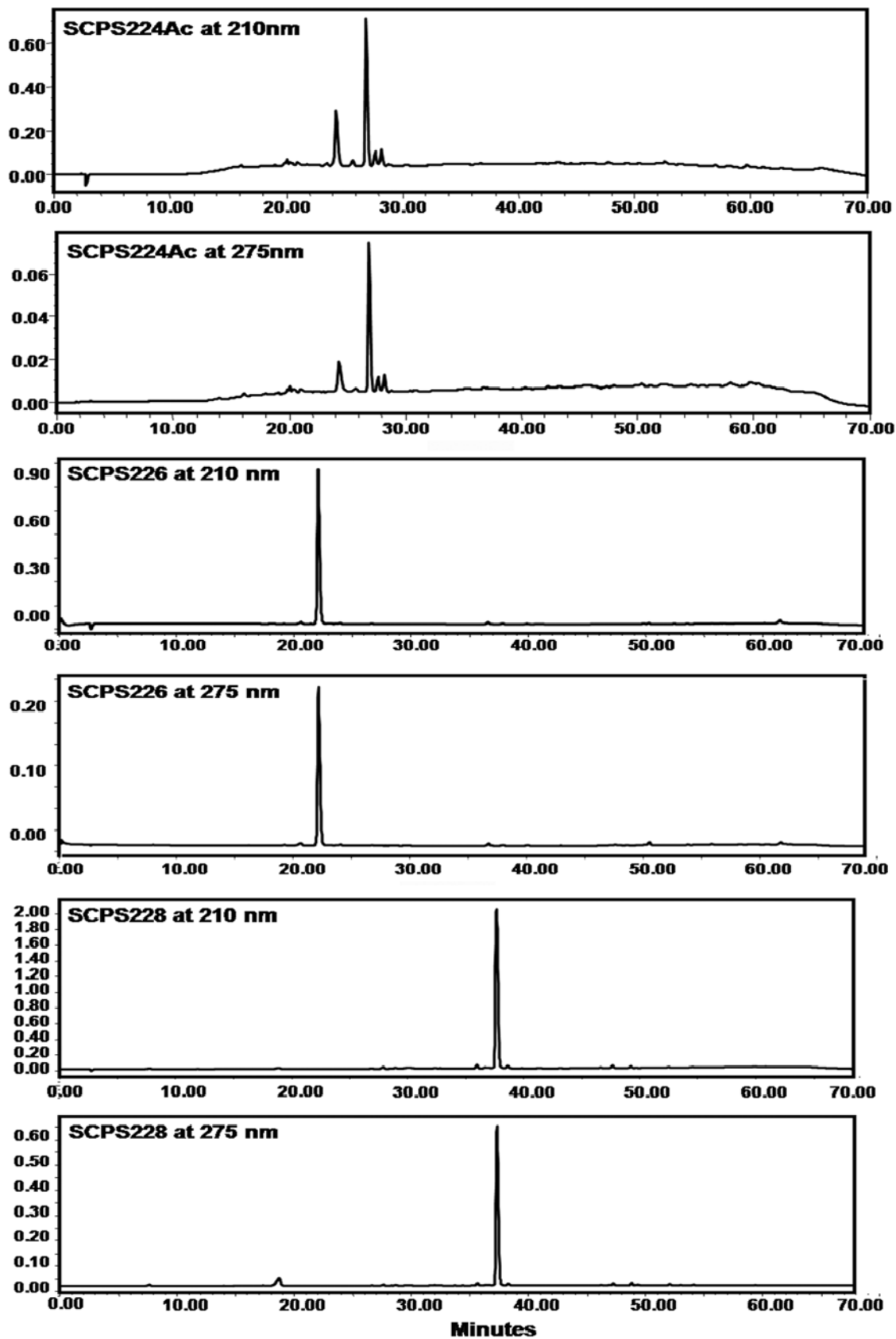
Figure S1: Design of peptides a) CPS224Ac, b) CPS226 and c) CPS228 by appending the 'CαNN' sequences from the respective PDB (1MUG for CPS224Ac; 1YCC for CPS226 and 1JW9 for CPS228)
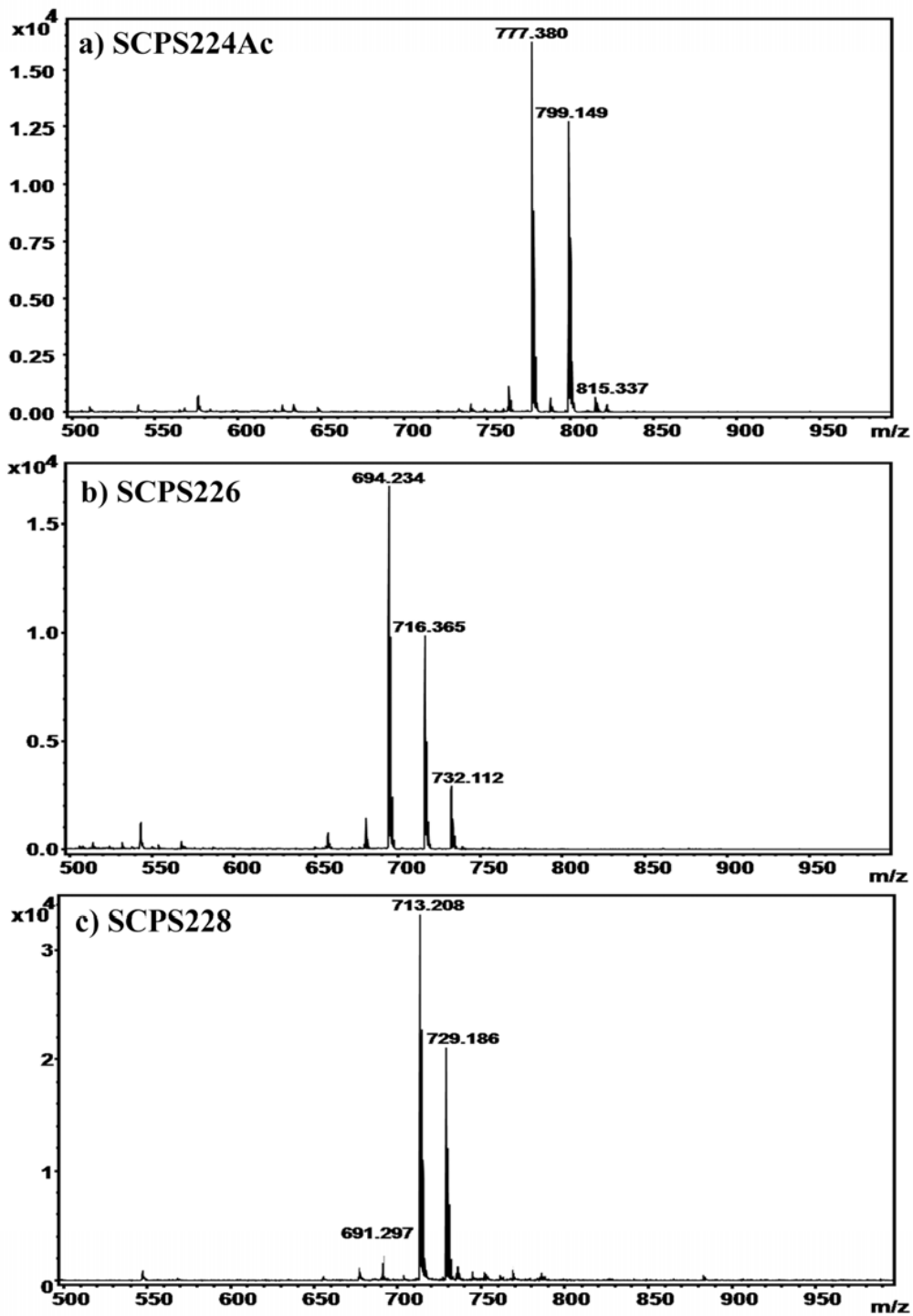
Figure S2: HPLC traces of the 18-residue peptides (CPS224Ac, CPS226 and CPS228) containing the naturally occurring 'CᵃNN' motif sequences using gradient of 0-60% acetonitrile in water with 0.1% TFA in dual λ absorbance detector (at 210 and 275 nm).
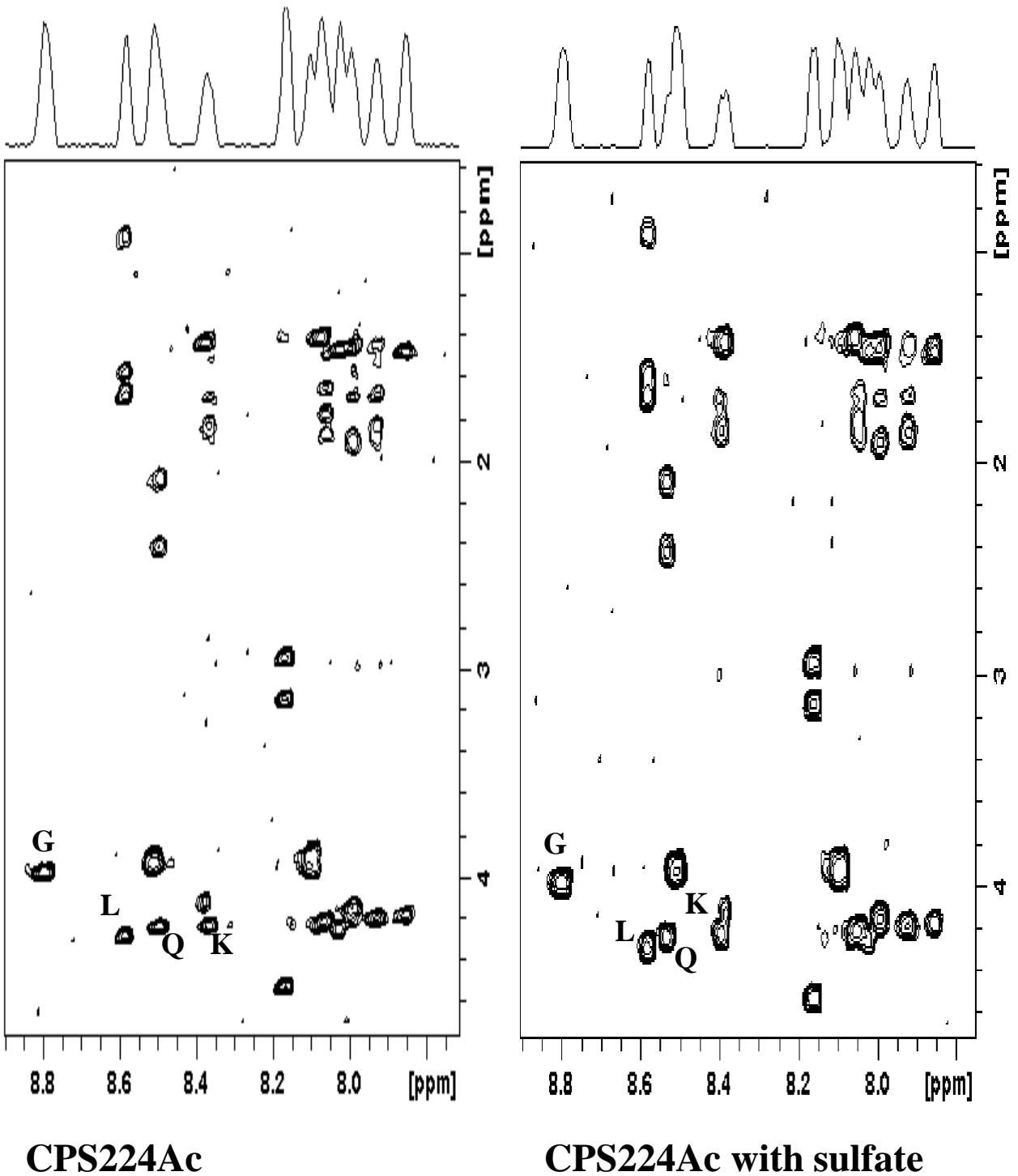
Figure S3: HPLC traces of the short 7-residue peptides (SCPS224Ac, SCPS226 and SCPS228) containing the naturally occurring 'CαNN' motif sequences using gradient of 0-60% acetonitrile in water with 0.1% TFA in dual λ absorbance detector (at 210 and 275 nm).
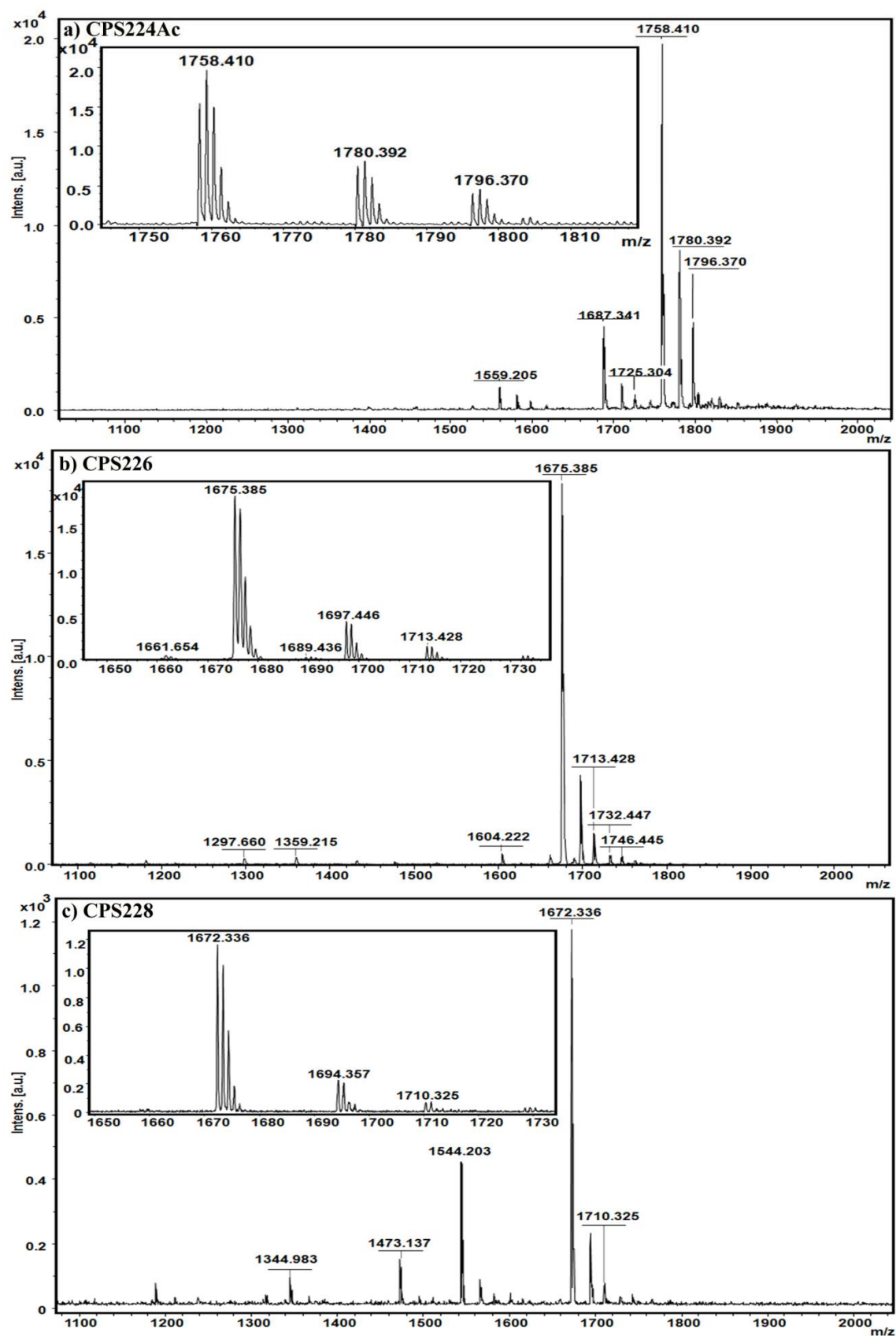
Figure S4: MALDI-MS (in the positive mode) traces of the 18-residue peptides [CPS224Ac (MW: 1757), CPS226 (MW: 1674) and CPS228 (MW: 1671)] containing the naturally occurring 'CᵅNN' motif sequences showing the m/z peaks (z=1) corresponding to respective M+H⁺; M+Na⁺ and M+K⁺ ions

Figure S5: MALDI-MS (in the positive mode) traces of the short 7-residue peptides [SCPS224Ac (MW: 776), SCPS226 (MW: 693) and SCPS228 (MW:690)] containing the naturally occurring 'CᵅNN' motif sequences showing the m/z peaks (z=1) corresponding to respective M+H⁺; M+Na⁺ and M+K⁺ ions

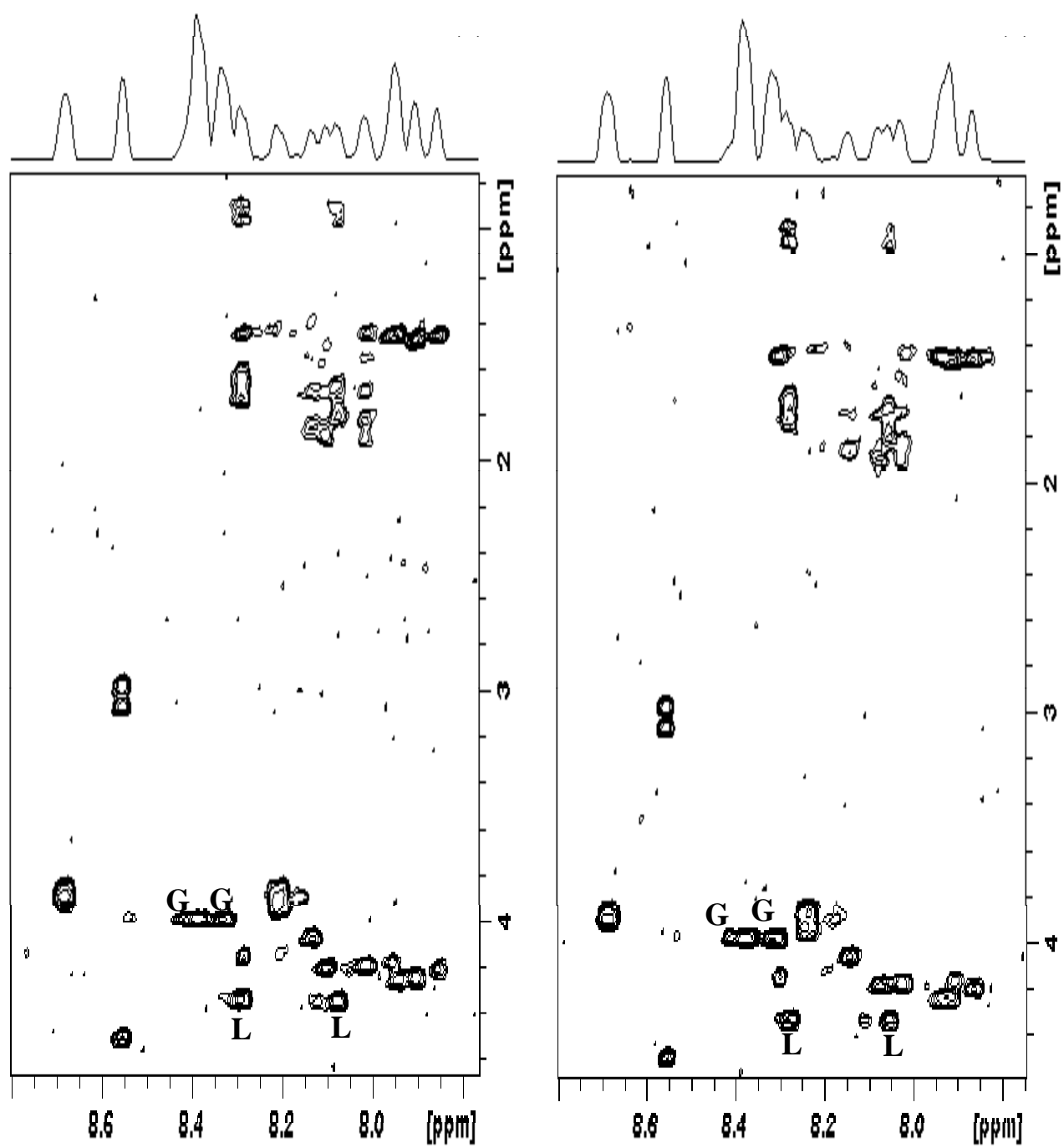**Figure S6:** TOCSY experiments of CPS224Ac and sulfate added species of CPS224Ac (1:3 M peptide : anion). The 'C$^{\alpha}$NN' motif sequence (-Leu-Gly-Lys-Gln-) present at N-terminus is labeled.
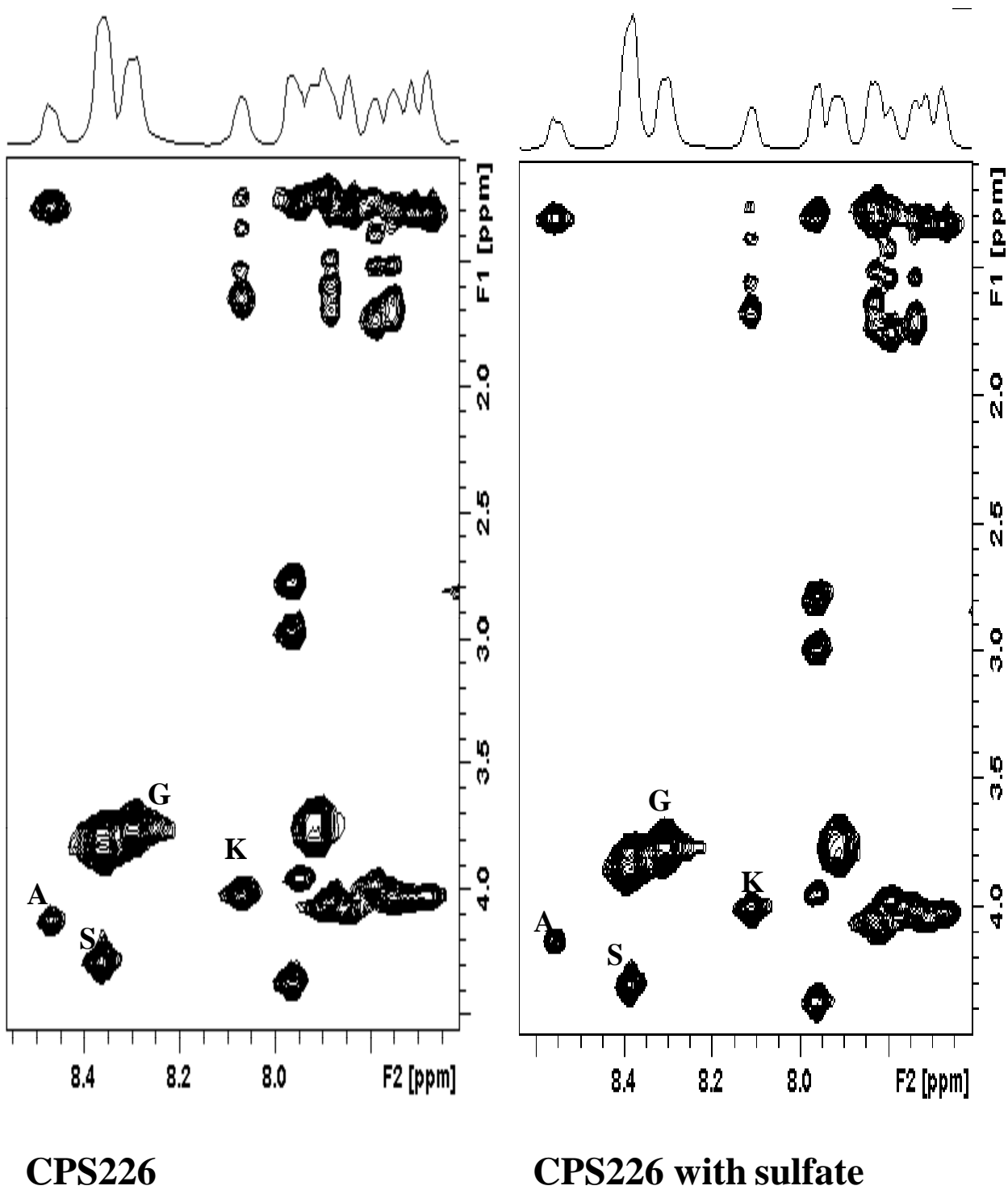
Figure S7: TOCSY experiments of CPS226 and sulfate added species of CPS226 (1:3 M peptide : anion). The 'C$^\alpha$NN' motif sequence (–Gly-Ser-Ala-Lys-) present at the N-terminus is labeled.
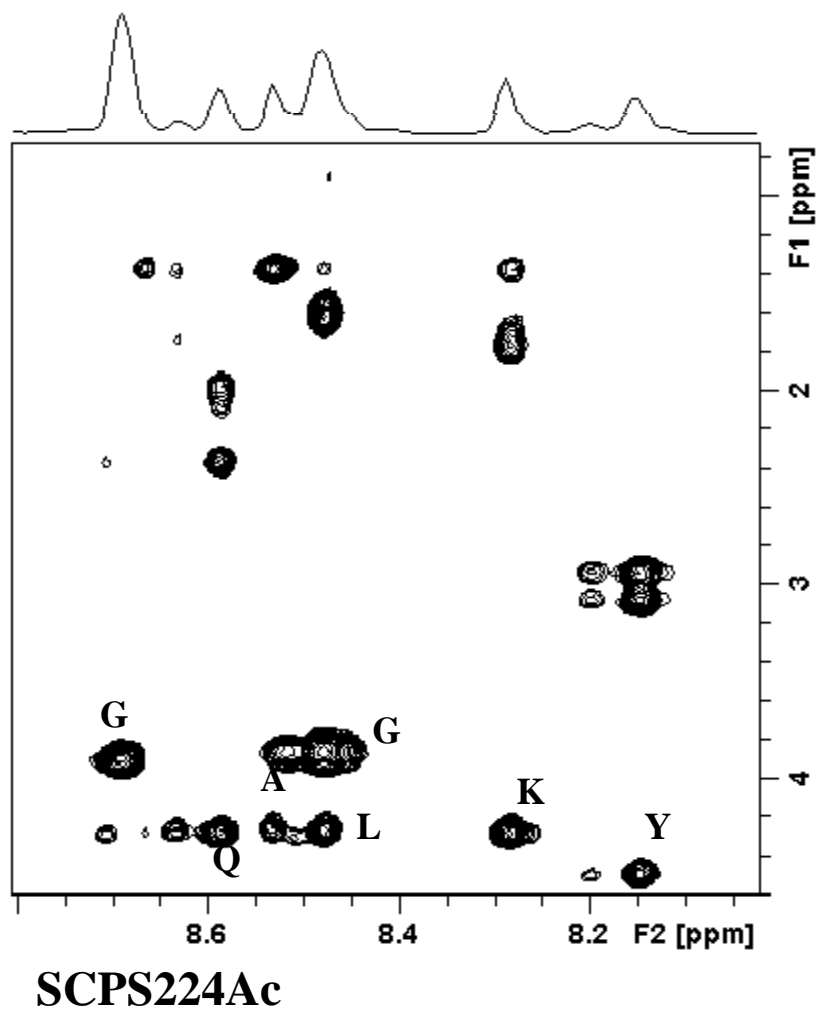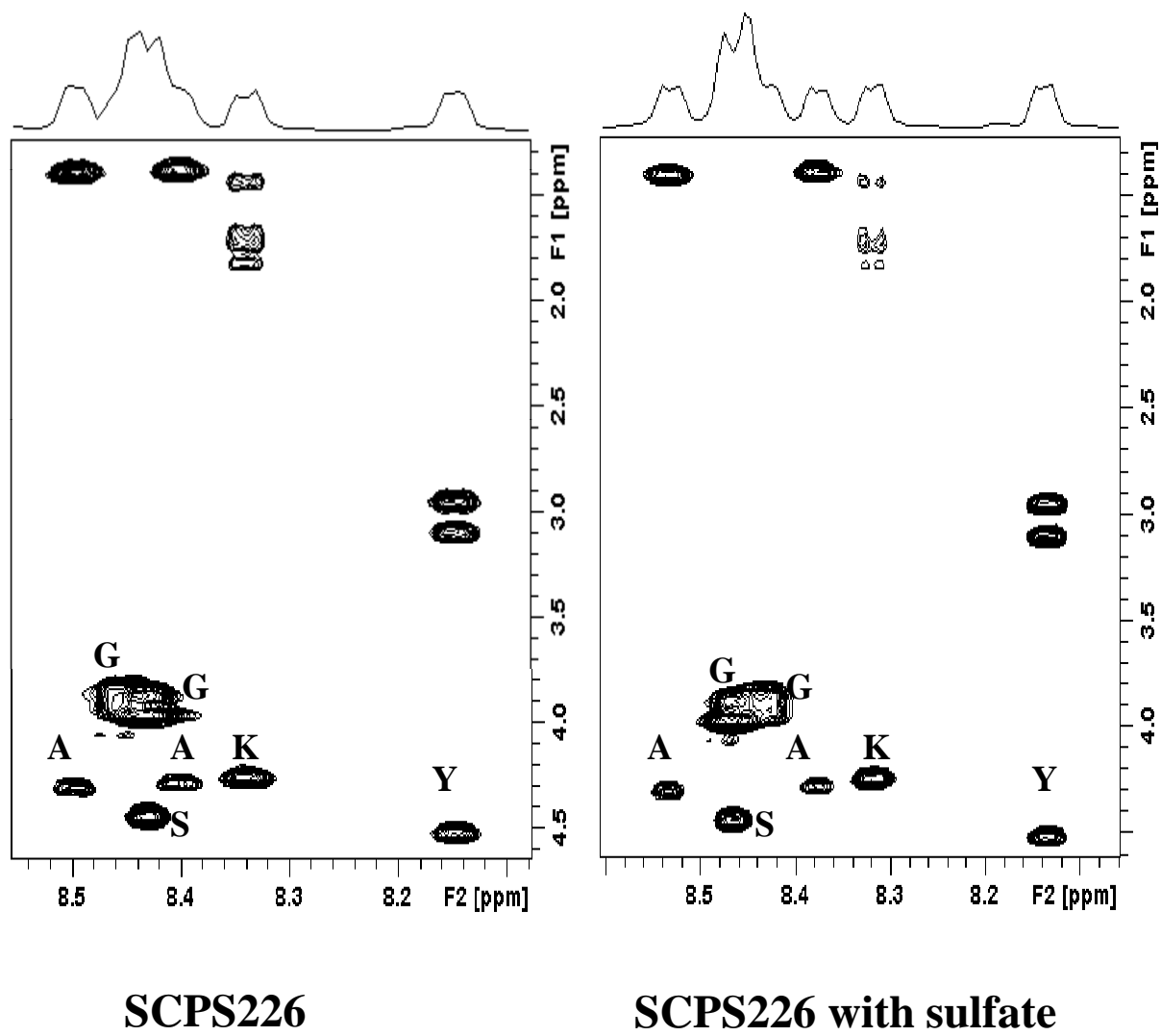
CPS228　　　　　　**CPS228 with sulfate**

Figure S8: TOCSY experiments of CPS228 and sulfate ion added species of CPS228 (1:3 M peptide : anion). The 'CαNN' motif sequence (-Leu-Gly-Gly-Leu-) present at N-terminus is labeled.

SCPS224Ac

Figure S9: TOCSY experiments of SCPS224Ac. The residues of the sequence (-Leu-Gly-Lys-Gln-Ala-Gly-Tyr) are labeled with one letter symbol.
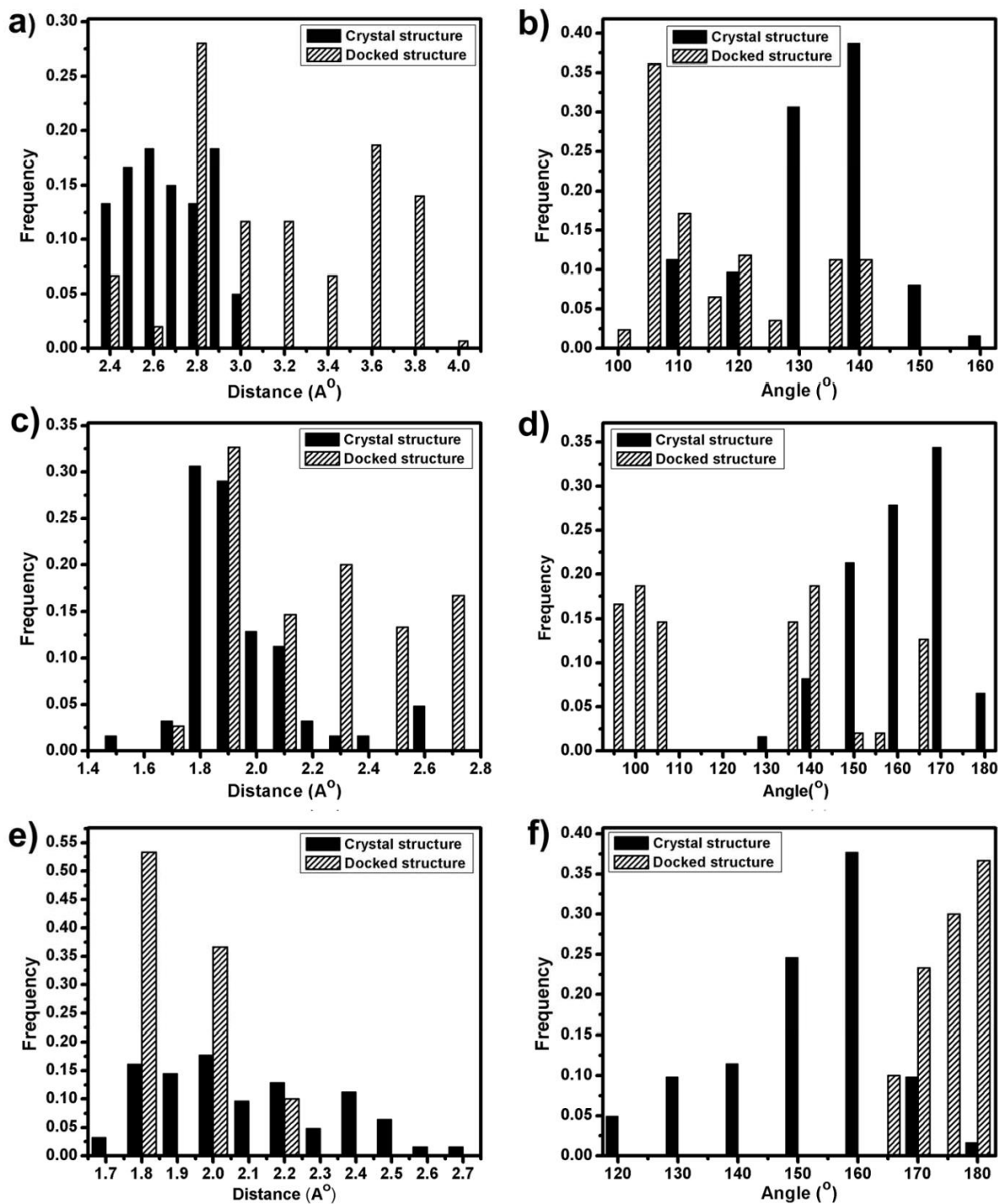
Figure S10: TOCSY experiments of SCPS226 and sulfate added species of SCPS226 (1:3 M peptide : anion). The residues of the sequence (-Gly-Ser-Ala-Lys-Ala-Gly-Tyr) are labeled with one letter symbol.
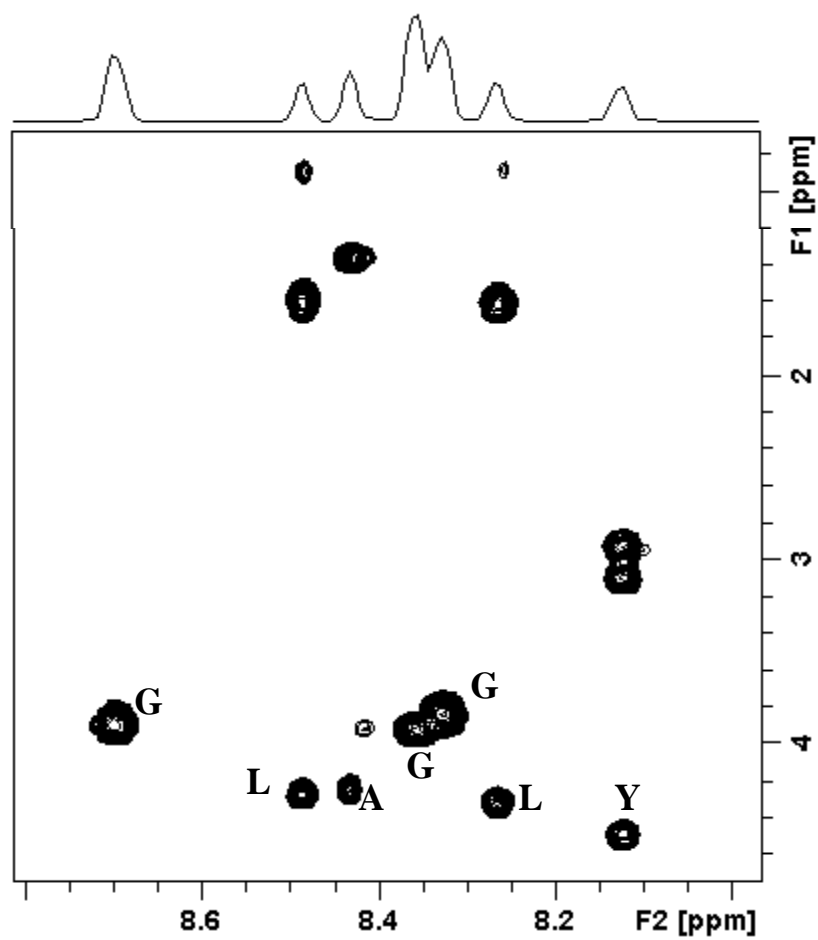
Figure S11: TOCSY experiments of SCPS228. The residues of the sequence (-Leu-Gly-Gly-Leu-Ala-Gly-Tyr-) are labeled with one letter symbol.

Figure S12: Comparison of X-H---O (where X= $C^{\alpha}_{-1}/N_0/N_{+1}$) distance and angle constraints of the anion added 'CαNN' sequences between the respective crystal structures (obtained from pdb) and the structures obtained from Molecular docking experiment for the designed chimeric peptide sequences (CPS224Ac, CPS226 and CPS228). a) C^{α}_{-1}-H---O distance; b) C^{α}_{-1}-H---O angle; c) N_0-H---O distance; d) N_0-H---O angle; e) N_{+1}-H---O distance, f) N_{+1}-H---O angle.
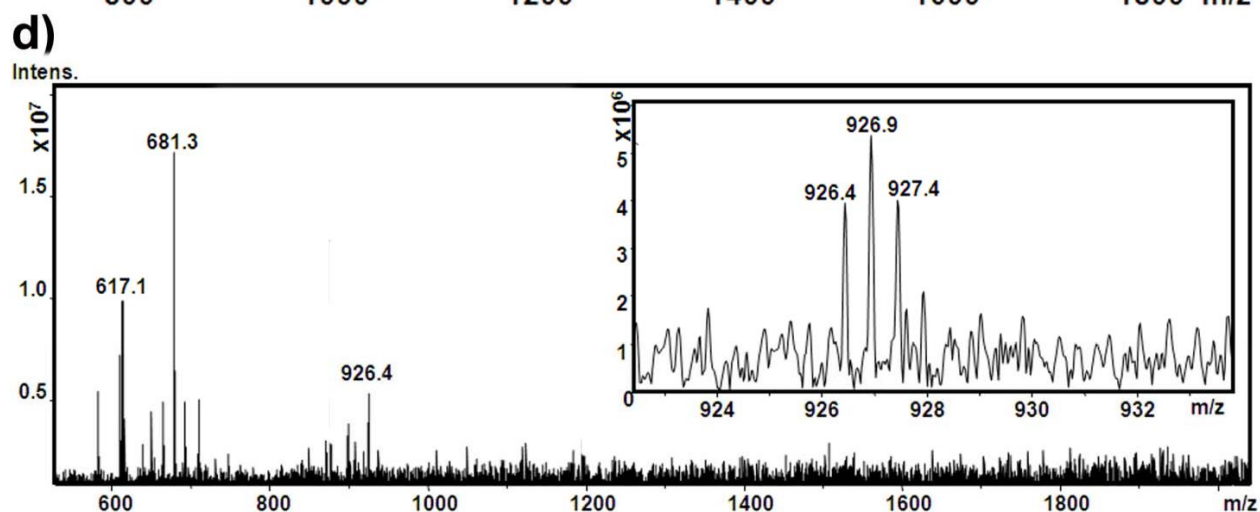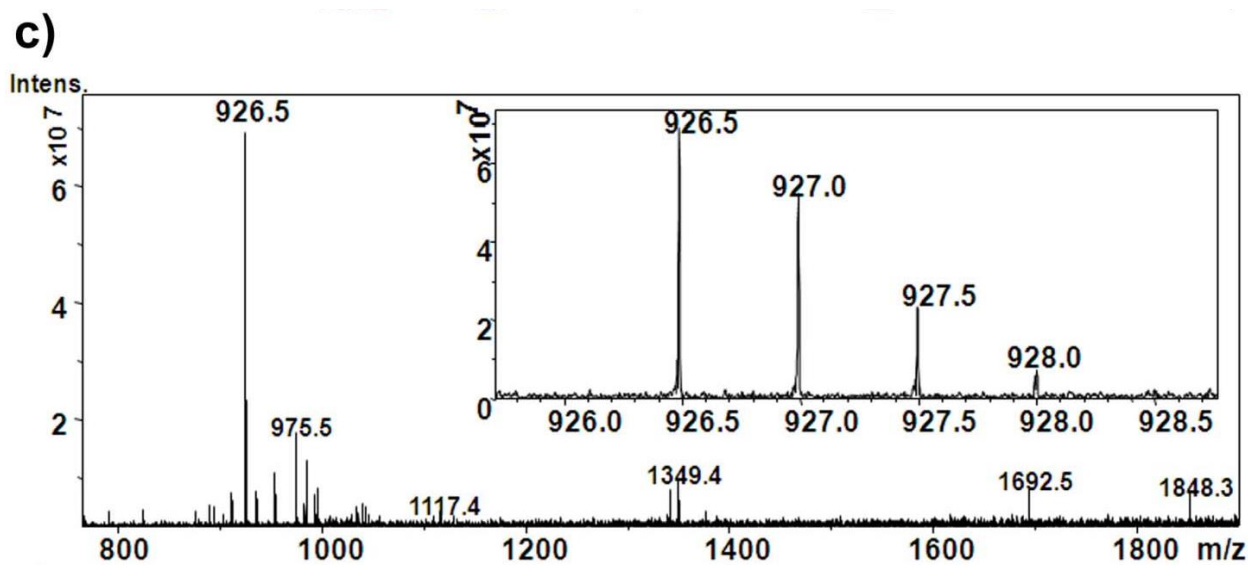
Figure S13 c) ESI-MS result showing the binding of the sulfate ion with CPS224Ac (Inset: isotopic distribution of m/z 926.5 showing the difference of 0.5, indicating doubly charged species); d) ESI-MS result showing the binding of phosphate ion with CPS224Ac (Inset: isotopic distribution of m/z 926.4 showing the difference of 0.5, indicating doubly charged species) [Reproduced from Figure 1 of Plos One 8(3), (**2013**) e57366 by Tridip Sheet, Subhrangsh Supakar & **Raja Banerjee*** (http://www.ncbi.nlm.nih.gov/pubmed/23516403) under the terms of the Creative Commons Attribution License, which permits, unrestricted use, distribution, and reproduction in any medium]
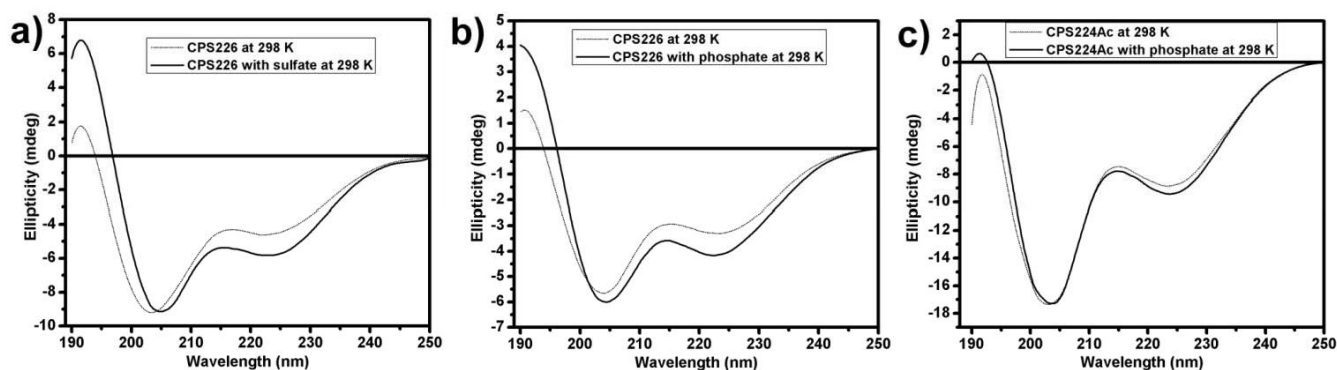
Figure S14: CD spectra of peptides in fully aqueous condition at 298 K in absence as well as in presence of anions. a) CPS226 with sulfate; b) CPS226 with phosphate and c) CPS224Ac with phosphate.
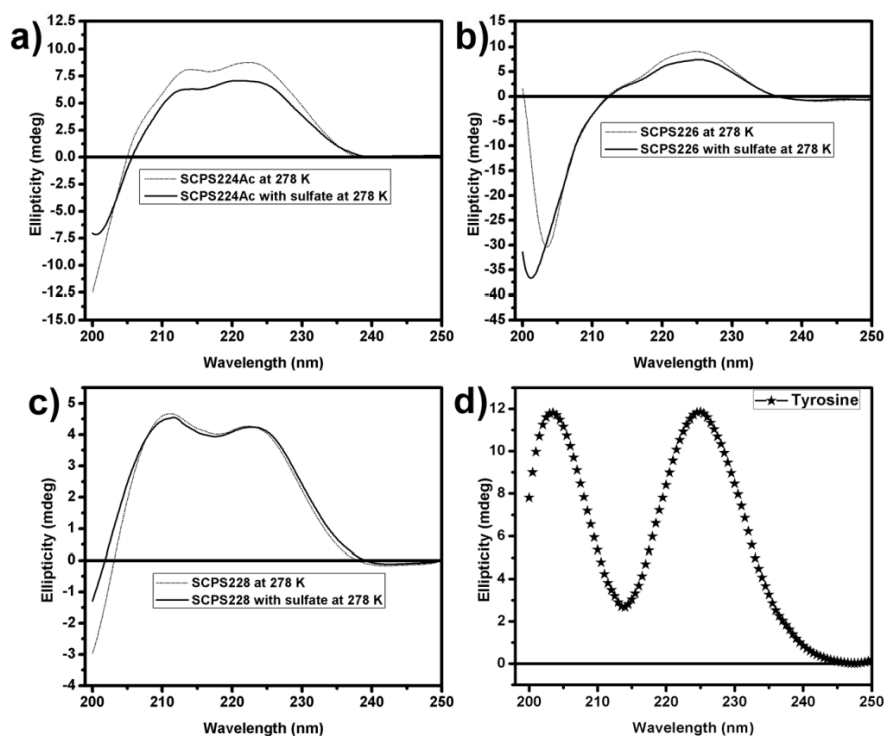
Figure S15: CD spectra of short peptides along with their sulfate added species and free tyrosine recorded at 278 K in fully aqueous condition; a) SCPS224Ac; b) SCPS226; c) SCPS228; d) tyrosine in aqueous solution.
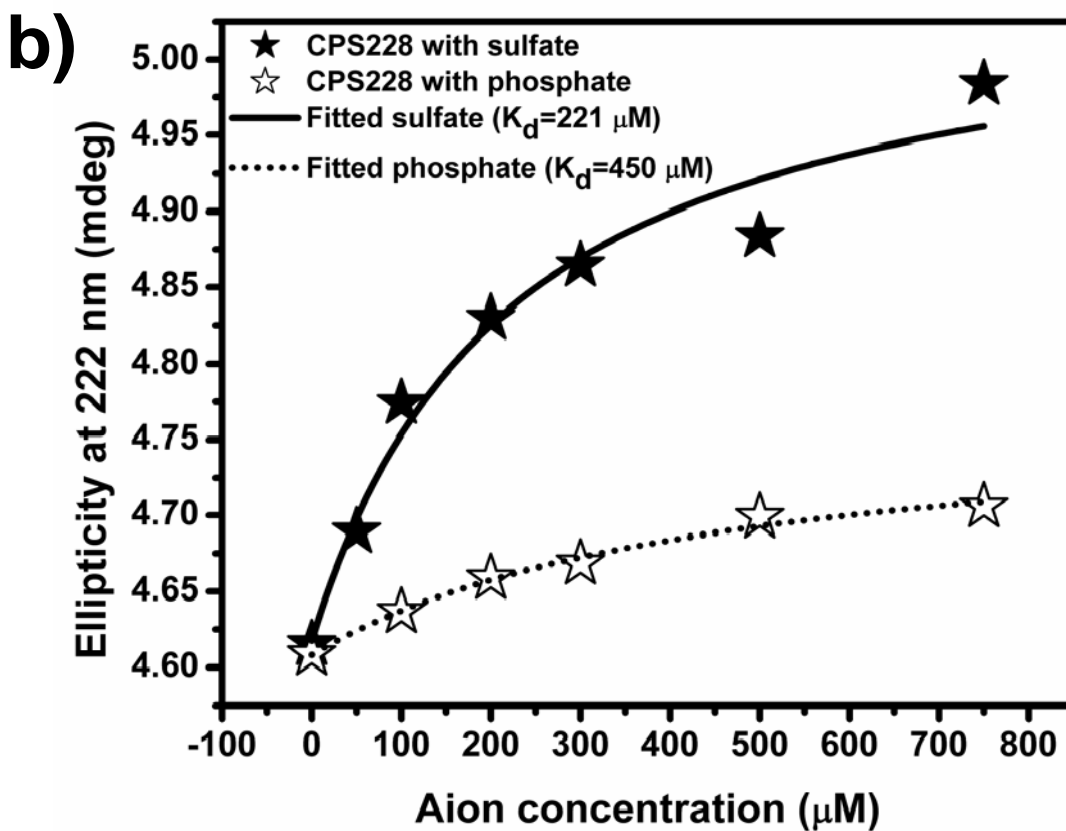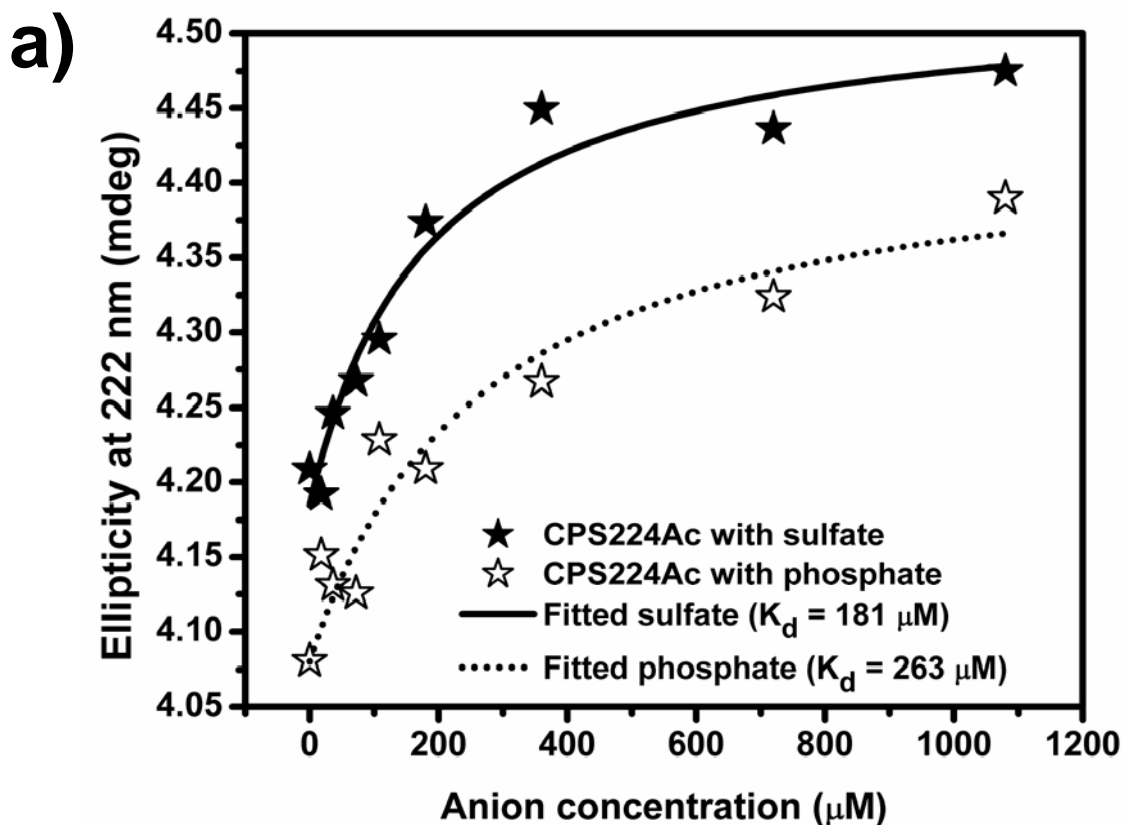
Figure S16: Change of ellipticity at 222nm (θ_{222}) with respect to added anion (sulfate and phosphate) in CD measurement of peptides a) CPS224Ac and b) CPS228 along with their respective anion dependent K_d
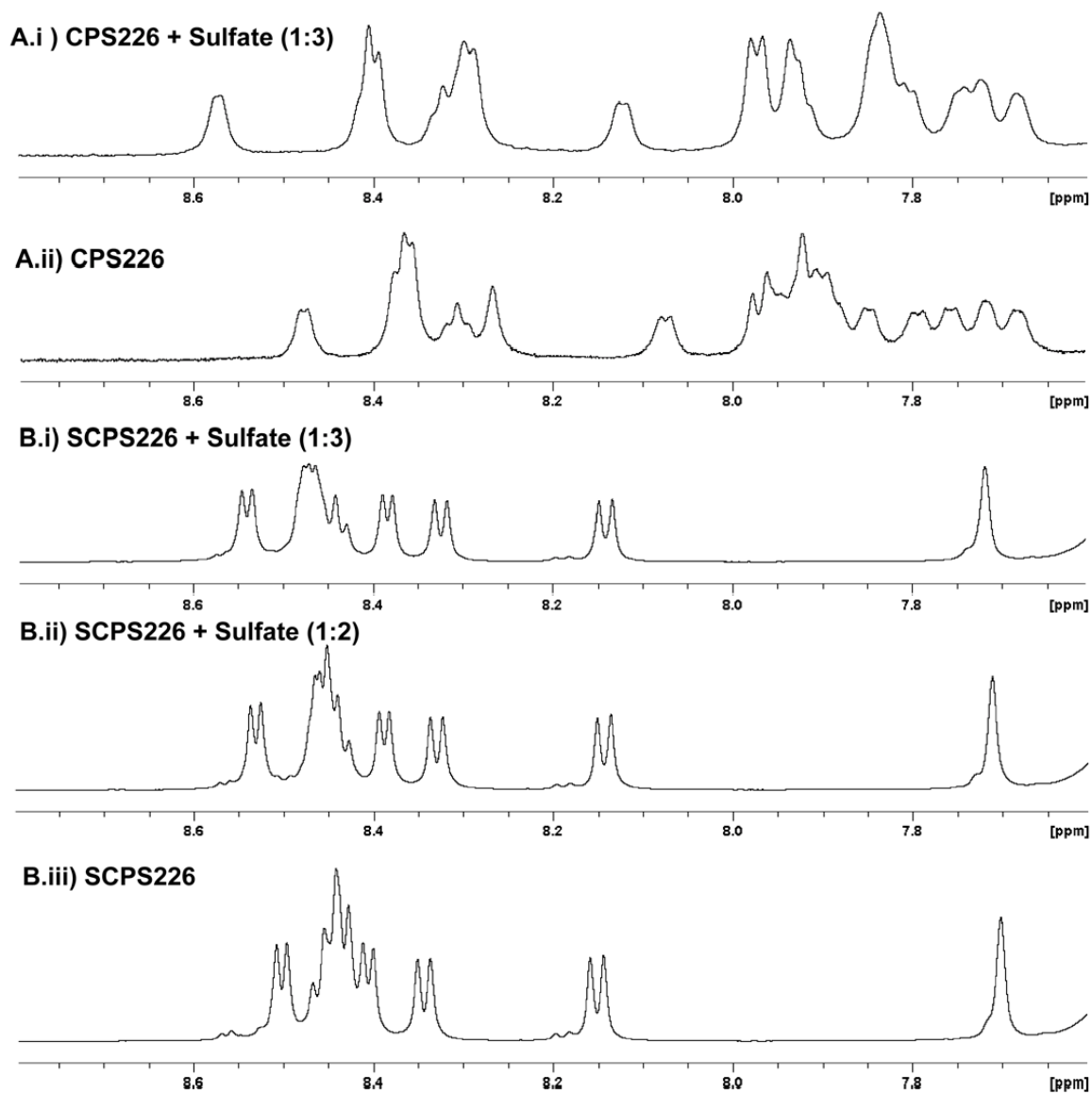
Figure S17: 1D spectra of main-chain HN region A)CPS226 and B) SCPS226 showing the effect of addition of sulfate ion to it in fully aqueous condition.

Table-S1. Comparison of sulfate ion bound 'G-S-A-K' motif structure obtained from NMR experiments of CPS226 with that of crystal structure in 1YCC deposited in pdb. Propagation of H-bonds in the crystal structure was calculated using DSSP Program (residue, secondary structure, propagation of H-bonds).

GSAK motif	Structures of the motif deposited in pdb**									Experimental study of the motif in solution		
	Distance (Å)		Dihedral angle			H-bond observed (DSSP)[4][6]			Expected NN noe in NMR	Observed NMR parameters		Proposed H-bond
	Residue	NH-NH (i,i+1)	Residue	ϕ	ψ	CO→NH (i→i+4) (i→i+3) a	b	c		Residue	NH-NH noe (i,i+1)	CO→NH (i→i+3) & (i→i+4)
** 1YCC (SO$_4^{2-}$ bound)	G/S (1/2)	3.86	G1	-131.8	-146.2	G			weak (w)			
	S/A (2/3)	4.18	S2	-97.3	115.6	S	>		weak (w)			
	A/K (3/4)	2.95	A3	-67.7	-35.0	A H	>		strong (s)			
						K H	>					
	K/ K (4/5)	2.27	K4	-68.1	-47.8	K H	>		strong (s)			
						G H		X				
			K5	-57.8	-39.9	A H		X				
CPS226	G/S (1/2)									G(1)	(1/2) weak	O$_1$(SO$_4^{2-}$) →A3
	S/A (2/3)									S(2)	(2/3) weak	O$_4$(SO$_4^{2-}$) → K4
	A/K (3/4)									A(3)	(3/4) strong	G1 → A5
	K/A (4/5)									K(4)	(4/5) strong	S2→ B6