Supplementary Information:
Predicting secondary structure propensities in IDPs using simple statistics from three-residue fragments

Alejandro Estaña[a,b], Amélie Barozet[a], Assia Mouhand[b], Marc Vaisset[a], Christophe Zanon[a], Pierre Fauret[a], Nathalie Sibille[b], Pau Bernadó[b,*], Juan Cortés[a,*]

[a] *LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France*
[b] *Centre de Biochimie Structurale. INSERM, CNRS, Université de Montpellier, France*

[*] Corresponding authors
Email addresses: pau.bernado@cbs.cnrs.fr, juan.cortes@laas.fr

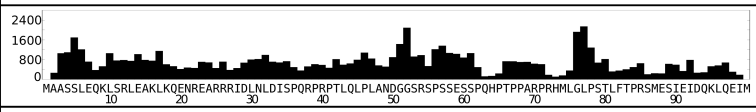
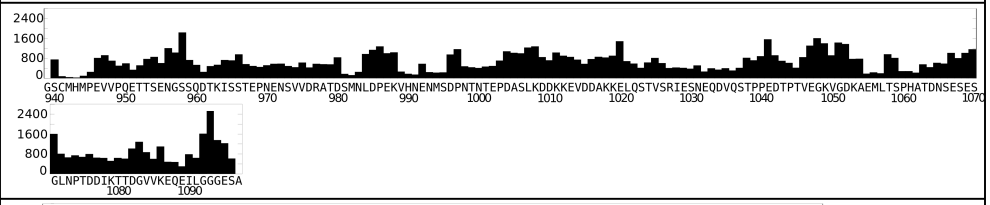
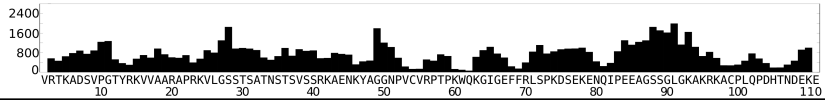
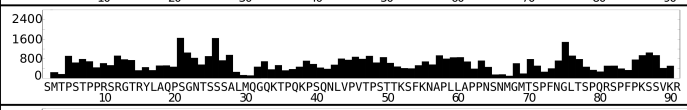
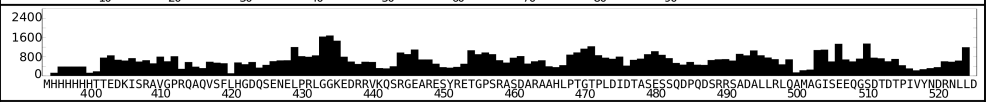
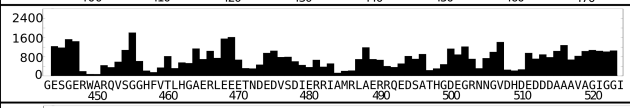
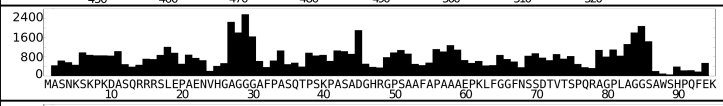
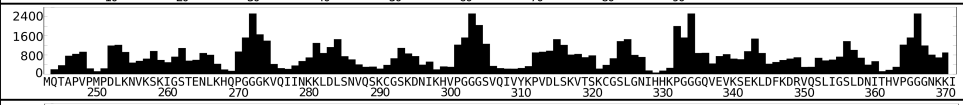
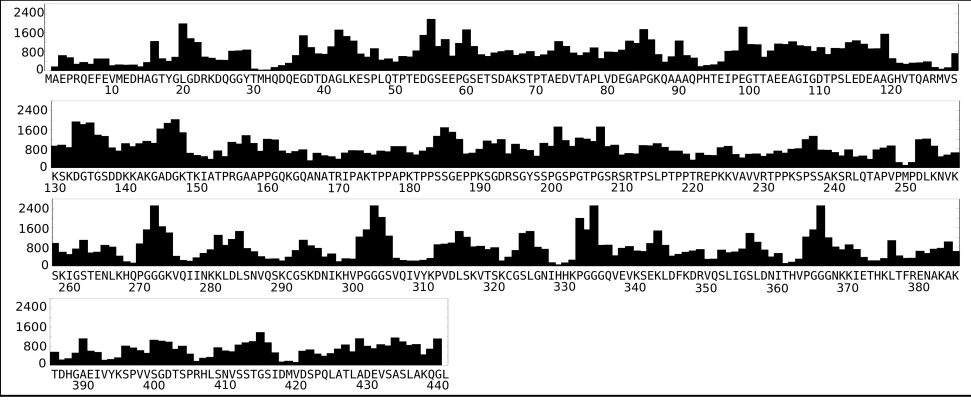
IDP	Number of Tripeptides			Number of Tripeptides along the sequence
	Average	Min.	Max.	
MKK7	694.5	109	2126	
EBA-181	714.8	22	2560	
p15	749.4	59	1998	
Sic1	637.3	104	1732	
ntailMV	658.3	58	1679	
ntailSV	719.2	72	1795	
USrc	782.2	58	2560	
K18	785.0	65	2560	
Tau	792.3	42	2560	

Figure S 1: Number of structures in our database for the tripeptides in the nine studied IDPs. The value for each tripeptide is shown on the central residue.

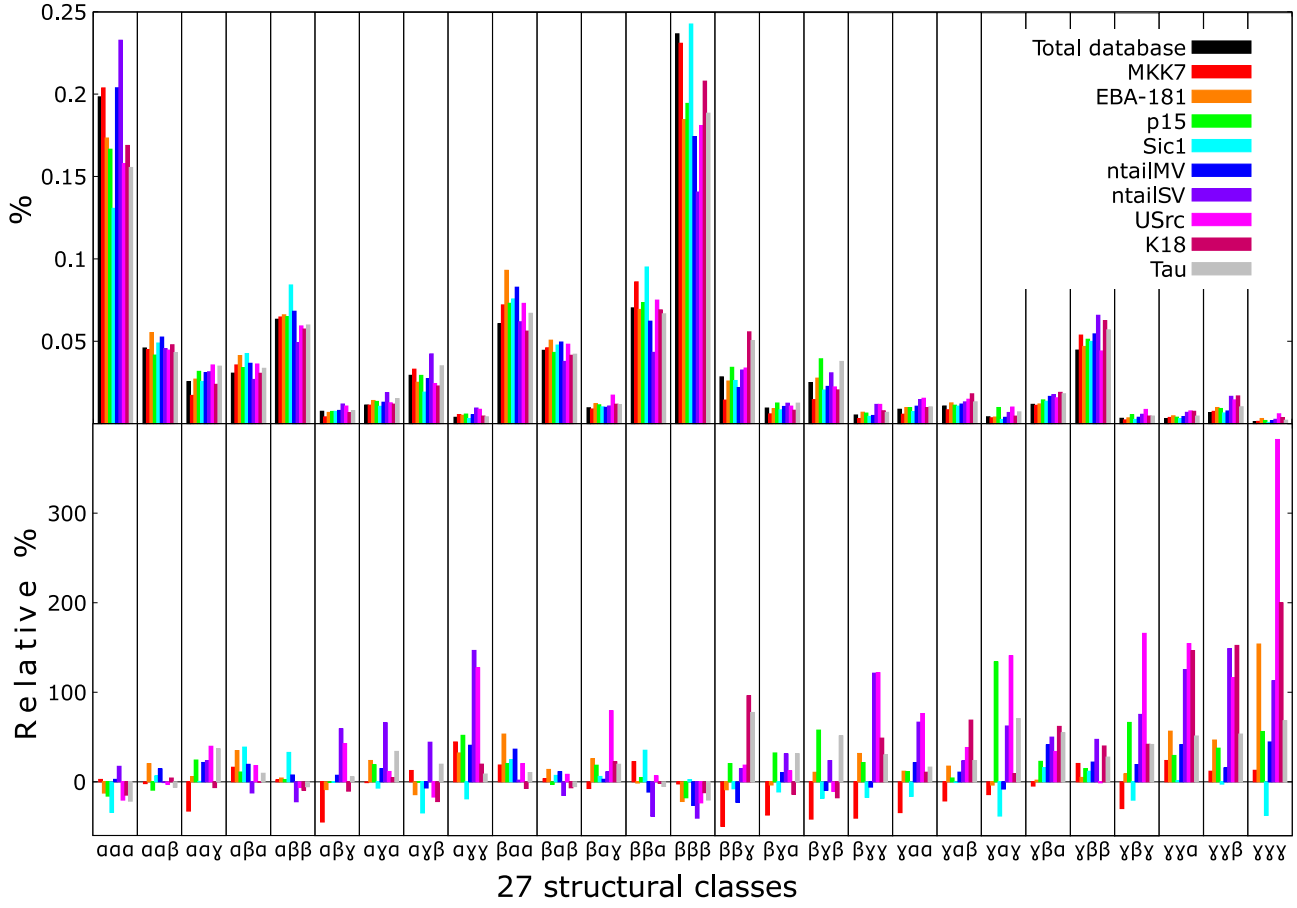


Figure S 2: (Top) Percentage of each secondary structure class for the tripeptides in the entire database and for the subset of tripeptides of the nine IDPs considered in this study. (Bottom) Relative percentage for each each structural class in all IDPs compared with the entire database.

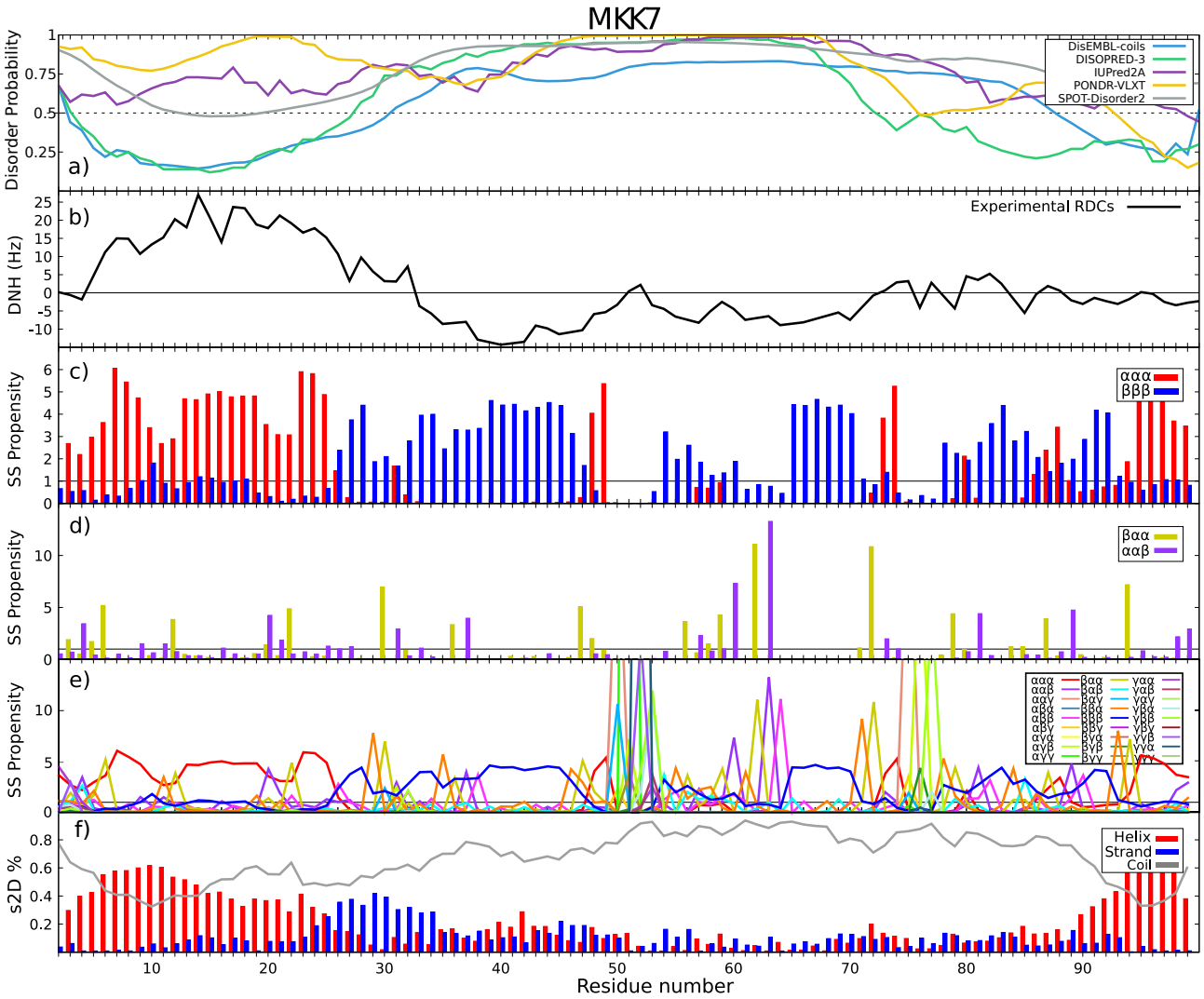


Figure S 3: LS2P secondary structure prediction for MKK7. (a) Disorder probability predictions. (b) Experimental RDC profile. (c) Helical ($\alpha\alpha\alpha$) and extended ($\beta\beta\beta$) propensities predicted by LS2P. (d) $\beta\alpha\alpha$ and $\alpha\alpha\beta$ propensities predicted by LS2P, allowing to identify β-turns. (e) Propensities predicted by LS2P for the 27 structural classes. For clarity purposes, values are truncated at 15. (f) Secondary structure populations predicted by s2D.

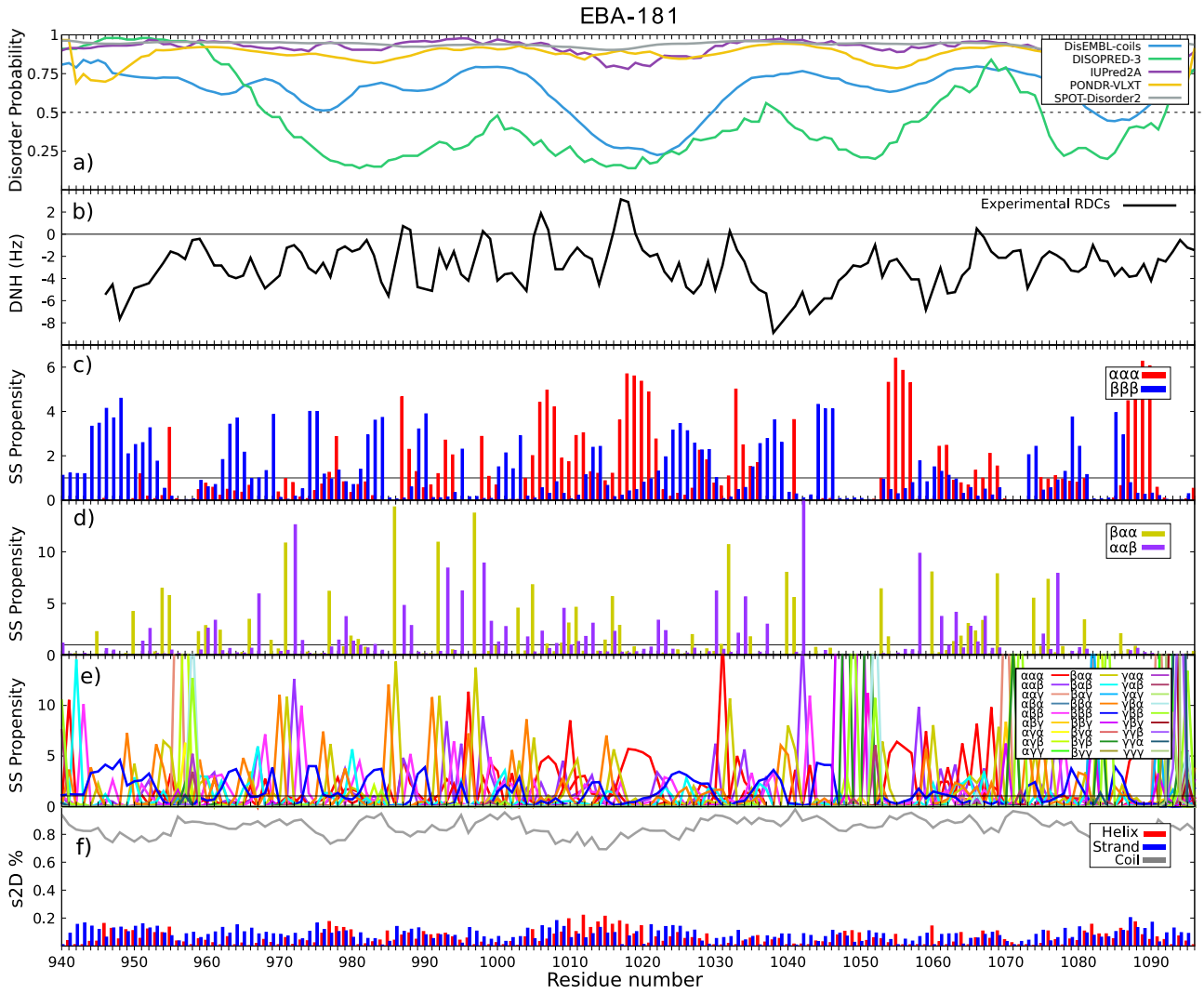


Figure S 4: LS2P secondary structure prediction for EBA-181. (a) Disorder probability predictions. (b) Experimental RDC profile. (c) Helical ($\alpha\alpha\alpha$) and extended ($\beta\beta\beta$) propensities predicted by LS2P. (d) $\beta\alpha\alpha$ and $\alpha\alpha\beta$ propensities predicted by LS2P, allowing to identify β-turns. (e) Propensities predicted by LS2P for the 27 structural classes. For clarity purposes, values are truncated at 15. (f) Secondary structure populations predicted by s2D.

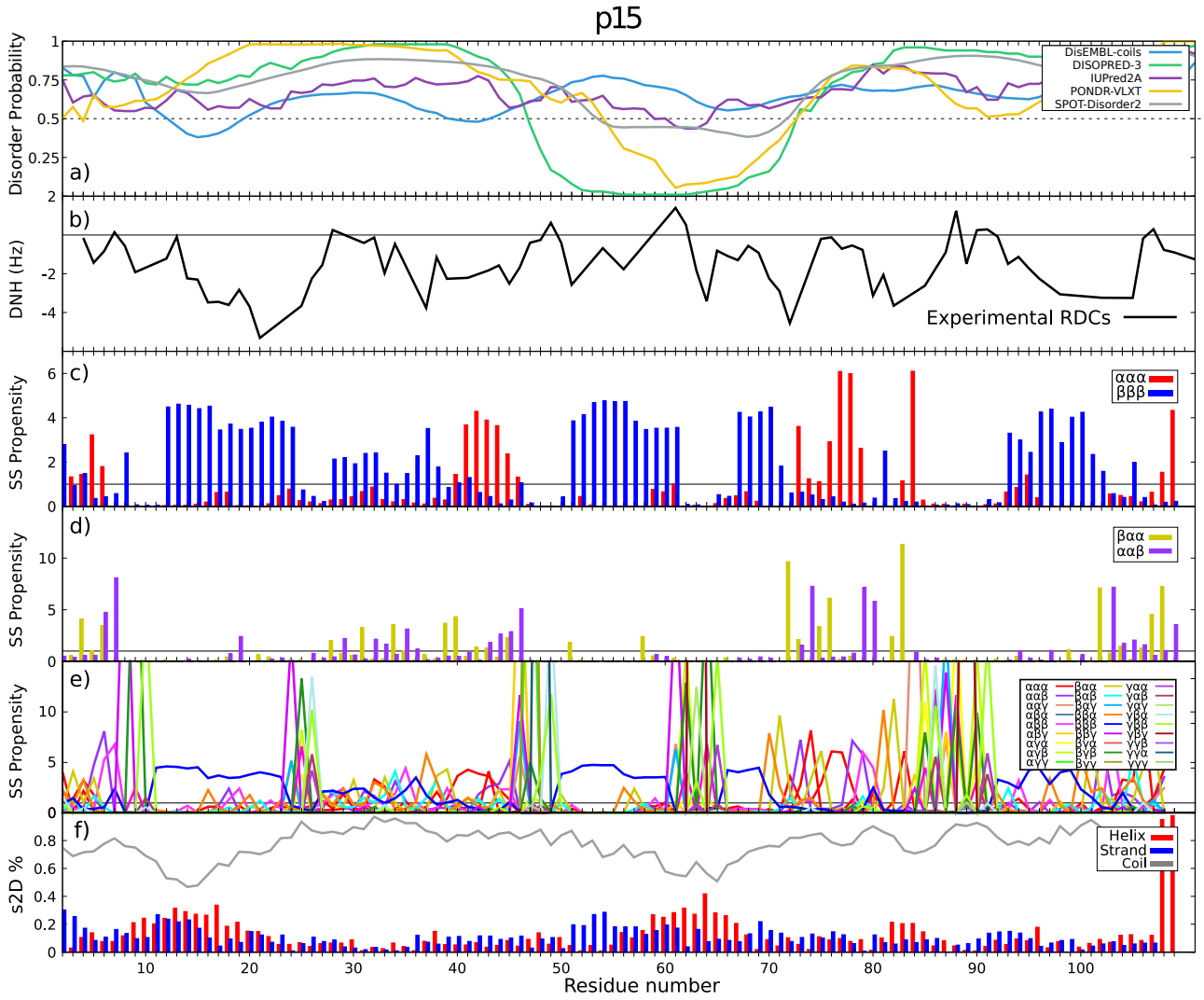


Figure S5: LS2P secondary structure prediction for p15. (a) Disorder probability predictions. (b) Experimental RDC profile. (c) Helical ($\alpha\alpha\alpha$) and extended ($\beta\beta\beta$) propensities predicted by LS2P. (d) $\beta\alpha\alpha$ and $\alpha\alpha\beta$ propensities predicted by LS2P, allowing to identify β-turns. (e) Propensities predicted by LS2P for the 27 structural classes. For clarity purposes, values are truncated at 15. (f) Secondary structure populations predicted by s2D.

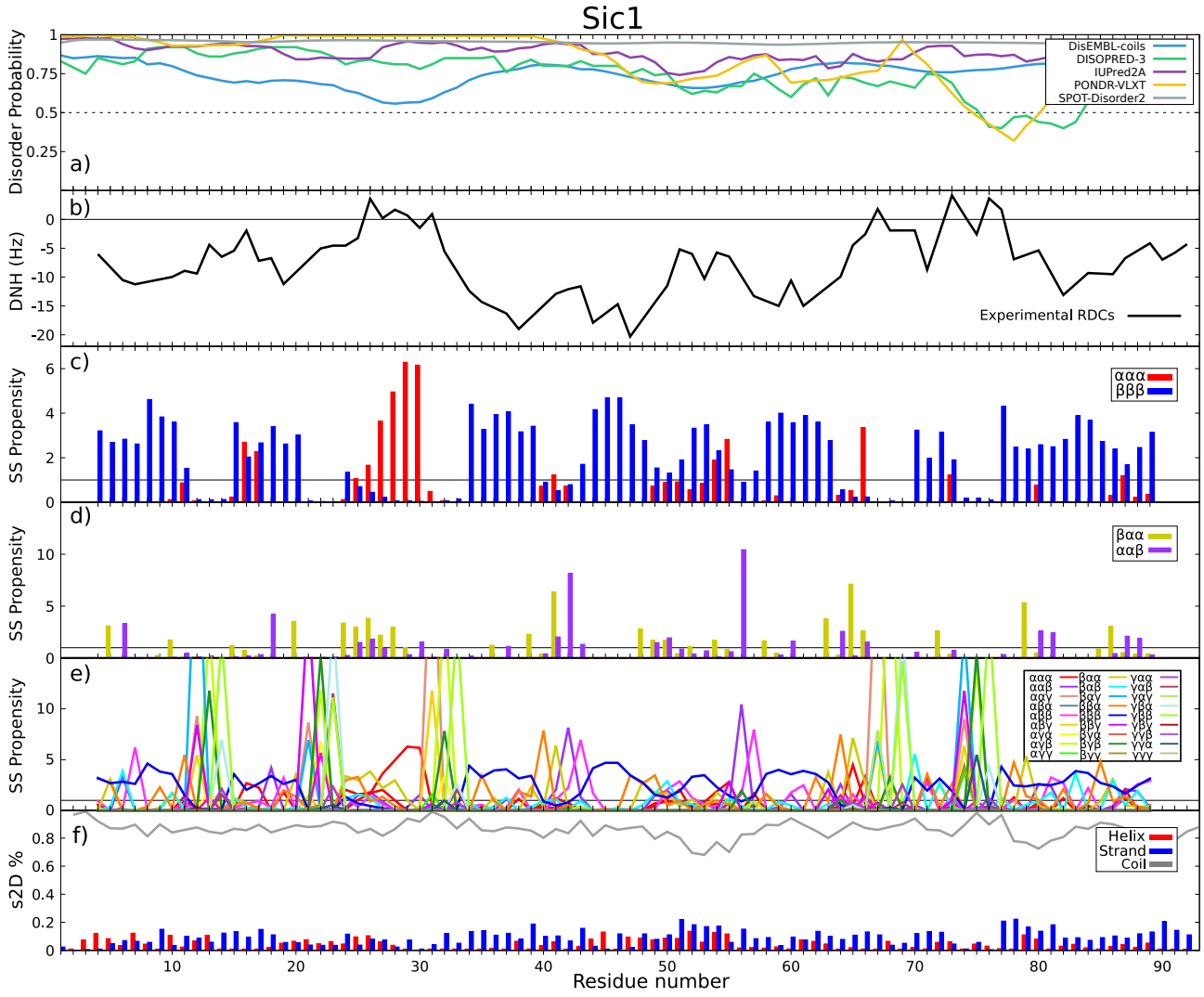


Figure S6: LS2P secondary structure prediction for Sic1. (a) Disorder probability predictions. (b) Experimental RDC profile. (c) Helical ($\alpha\alpha\alpha$) and extended ($\beta\beta\beta$) propensities predicted by LS2P. (d) $\beta\alpha\alpha$ and $\alpha\alpha\beta$ propensities predicted by LS2P, allowing to identify β-turns. (e) Propensities predicted by LS2P for the 27 structural classes. For clarity purposes, values are truncated at 15. (f) Secondary structure populations predicted by s2D.

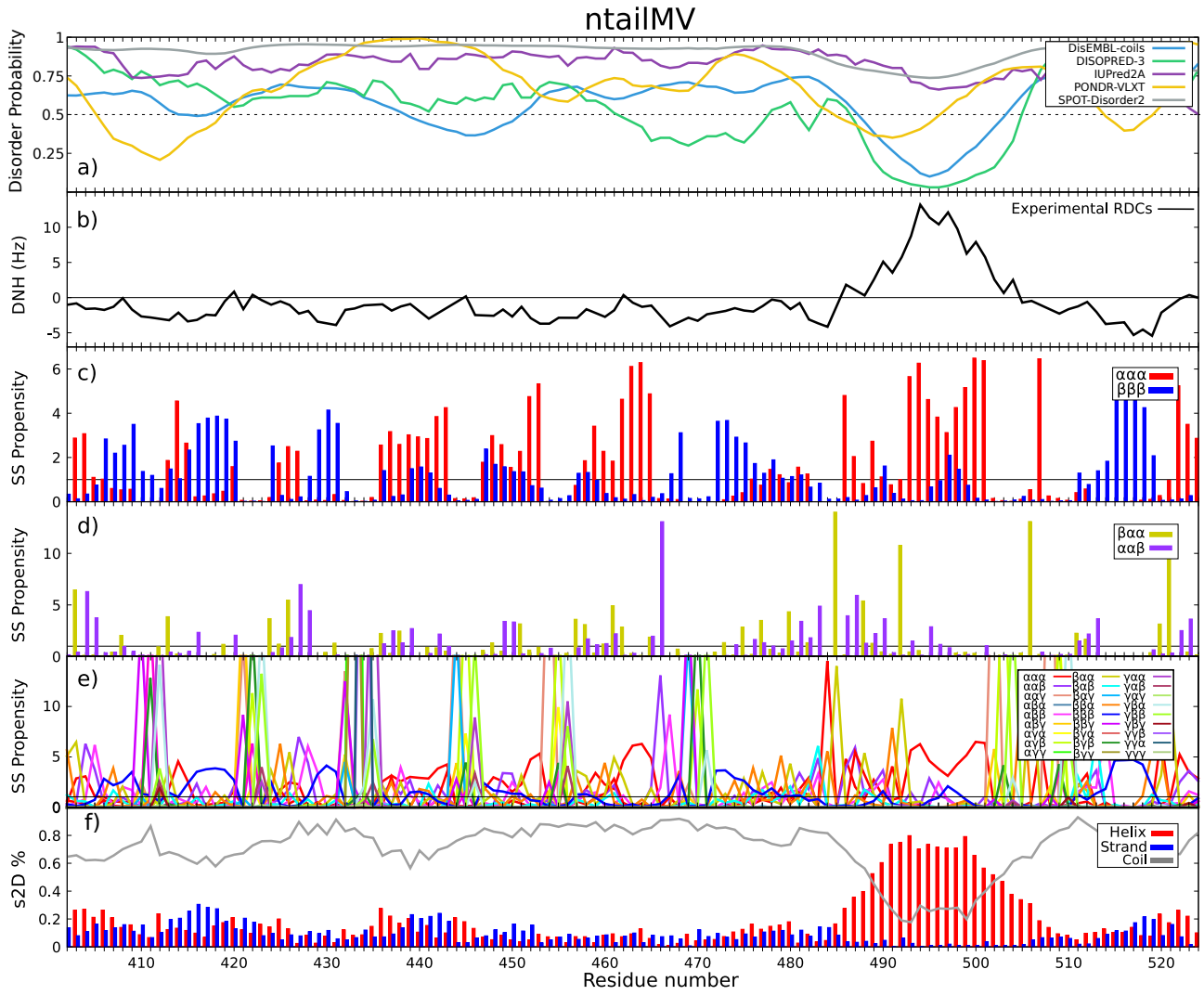


Figure S 7: LS2P secondary structure prediction for ntailMV. (a) Disorder probability predictions. (b) Experimental RDC profile. (c) Helical ($\alpha\alpha\alpha$) and extended ($\beta\beta\beta$) propensities predicted by LS2P. (d) $\beta\alpha\alpha$ and $\alpha\alpha\beta$ propensities predicted by LS2P, allowing to identify β-turns. (e) Propensities predicted by LS2P for the 27 structural classes. For clarity purposes, values are truncated at 15. (f) Secondary structure populations predicted by s2D.

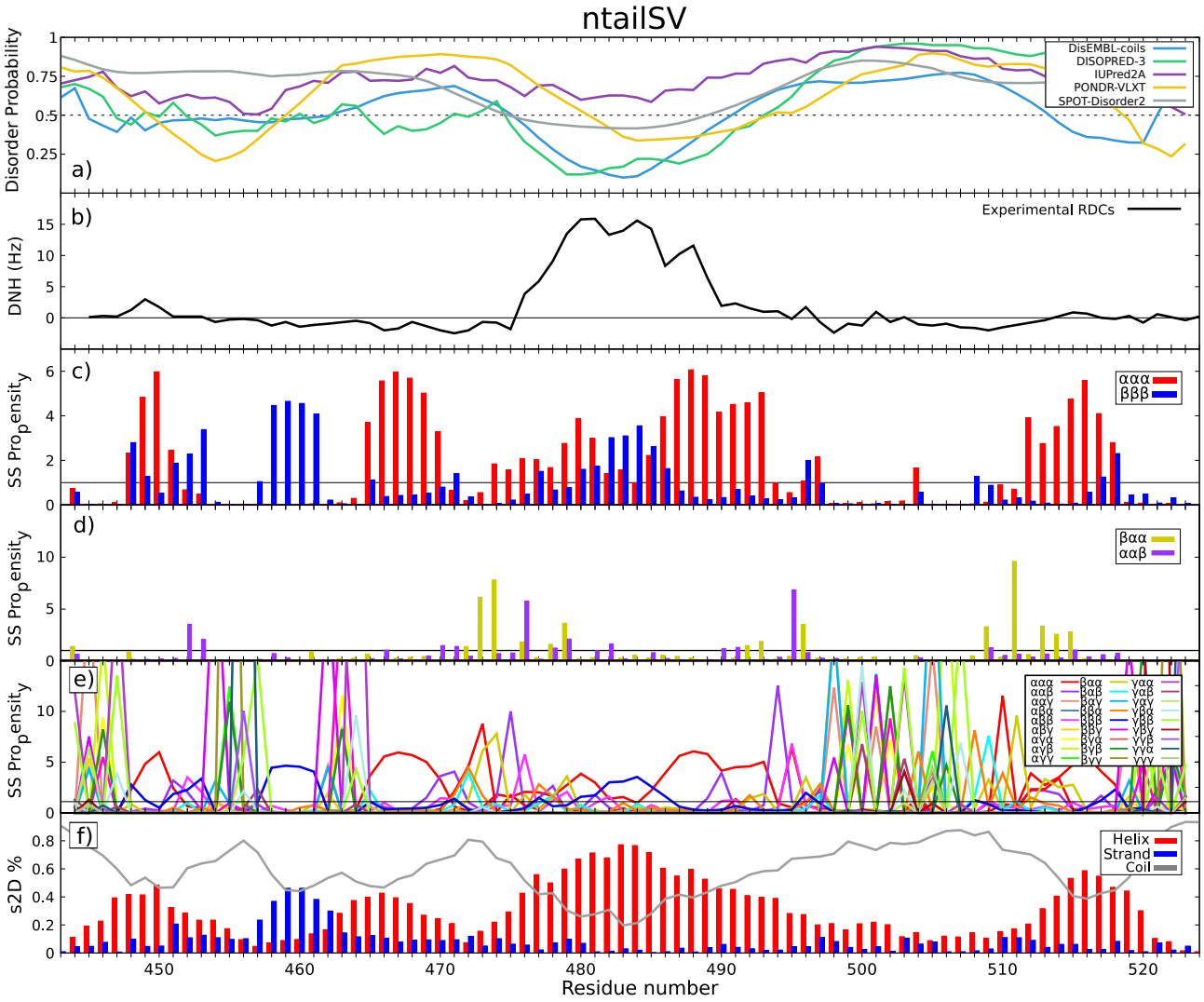


Figure S 8: LS2P secondary structure prediction for ntailSV. (a) Disorder probability predictions. (b) Experimental RDC profile. (c) Helical ($\alpha\alpha\alpha$) and extended ($\beta\beta\beta$) propensities predicted by LS2P. (d) $\beta\alpha\alpha$ and $\alpha\alpha\beta$ propensities predicted by LS2P, allowing to identify β-turns. (e) Propensities predicted by LS2P for the 27 structural classes. For clarity purposes, values are truncated at 15. (f) Secondary structure populations predicted by s2D.

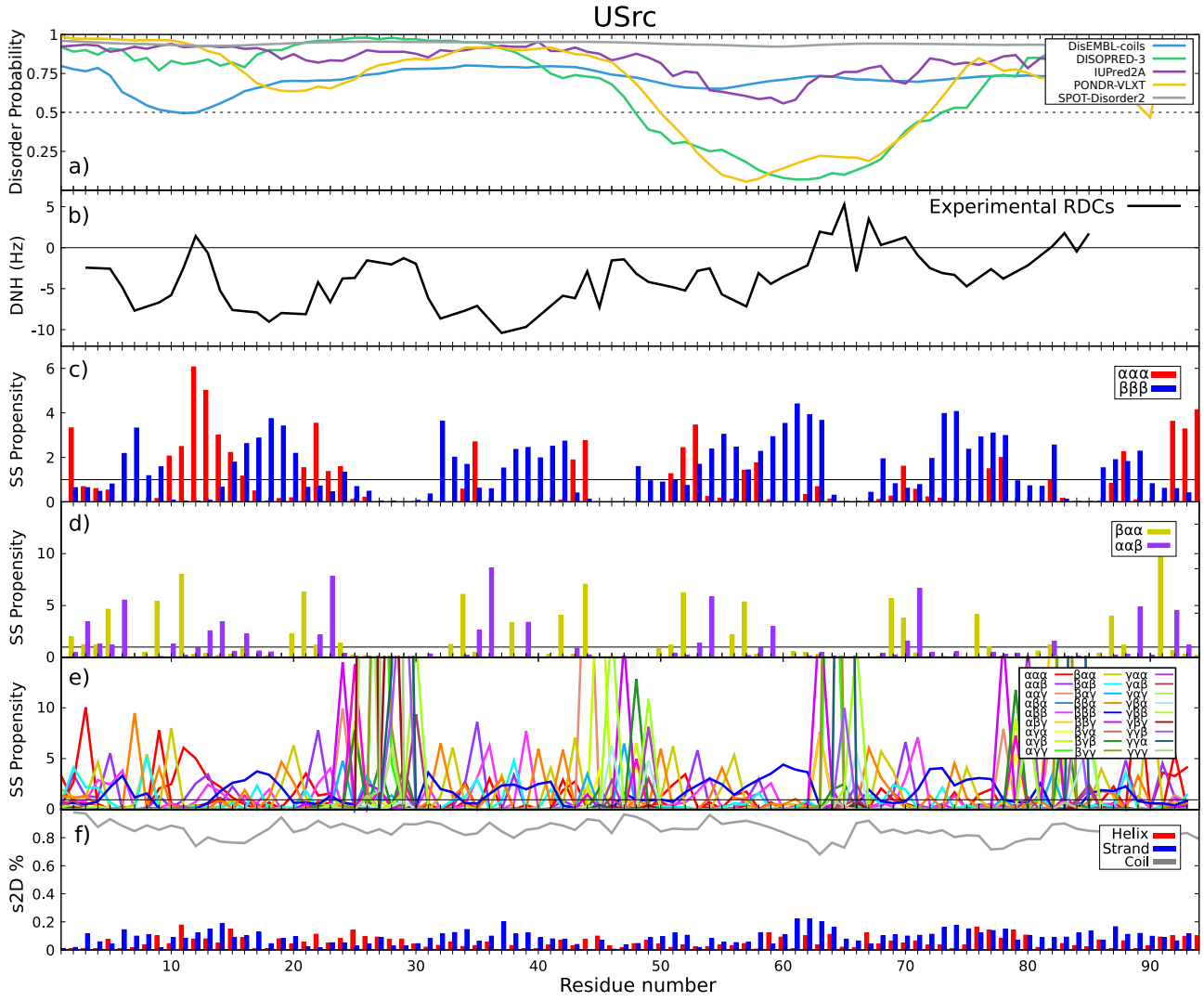


Figure S9: LS2P secondary structure prediction for USrc. (a) Disorder probability predictions. (b) Experimental RDC profile. (c) Helical ($\alpha\alpha\alpha$) and extended ($\beta\beta\beta$) propensities predicted by LS2P. (d) $\beta\alpha\alpha$ and $\alpha\alpha\beta$ propensities predicted by LS2P, allowing to identify β-turns. (e) Propensities predicted by LS2P for the 27 structural classes. For clarity purposes, values are truncated at 15. (f) Secondary structure populations predicted by s2D.

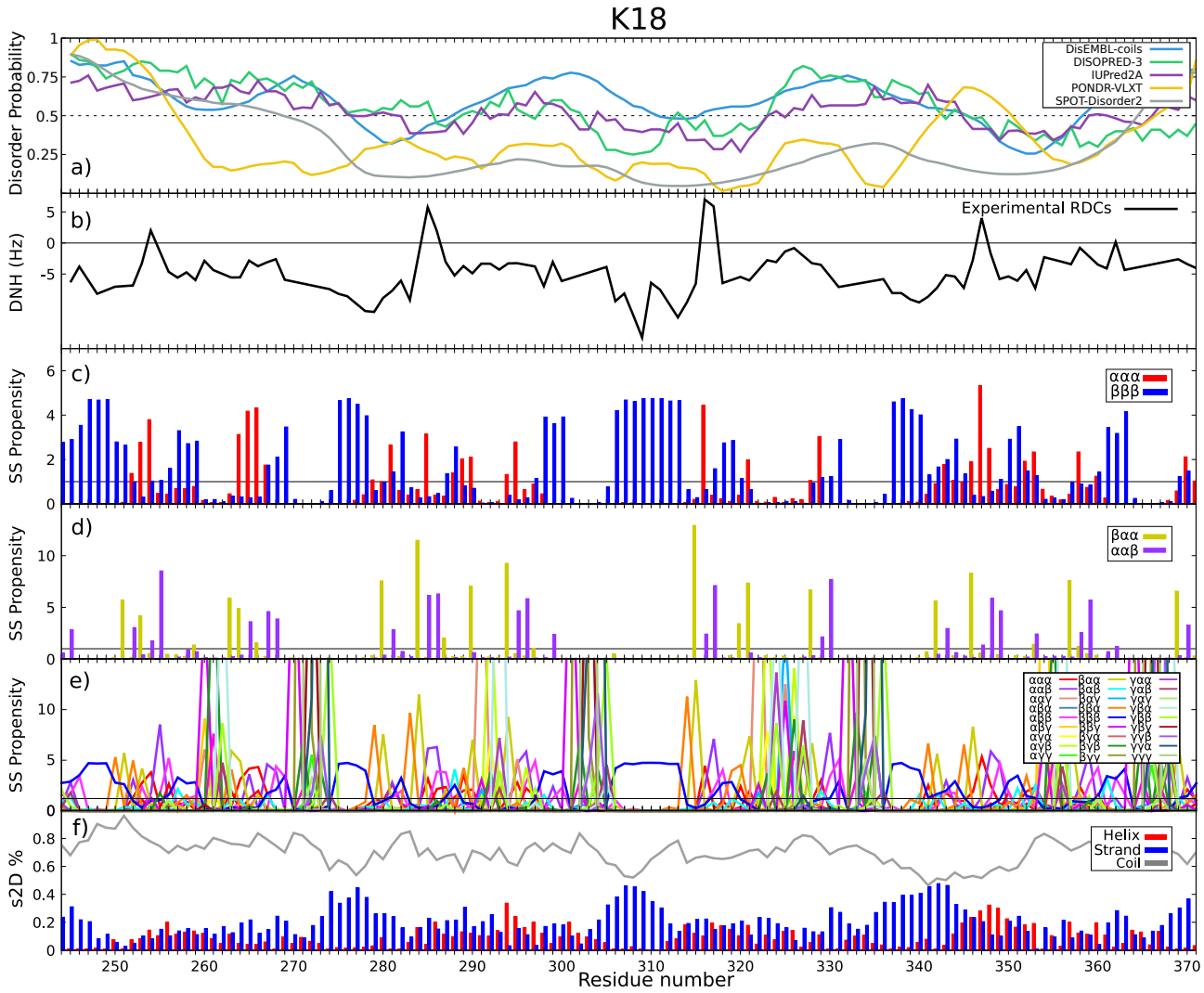


Figure S 10: LS2P secondary structure prediction for K18. (a) Disorder probability predictions. (b) Experimental RDC profile. (c) Helical ($\alpha\alpha\alpha$) and extended ($\beta\beta\beta$) propensities predicted by LS2P. (d) $\beta\alpha\alpha$ and $\alpha\alpha\beta$ propensities predicted by LS2P, allowing to identify β-turns. (e) Propensities predicted by LS2P for the 27 structural classes. For clarity purposes, values are truncated at 15. (f) Secondary structure populations predicted by s2D.

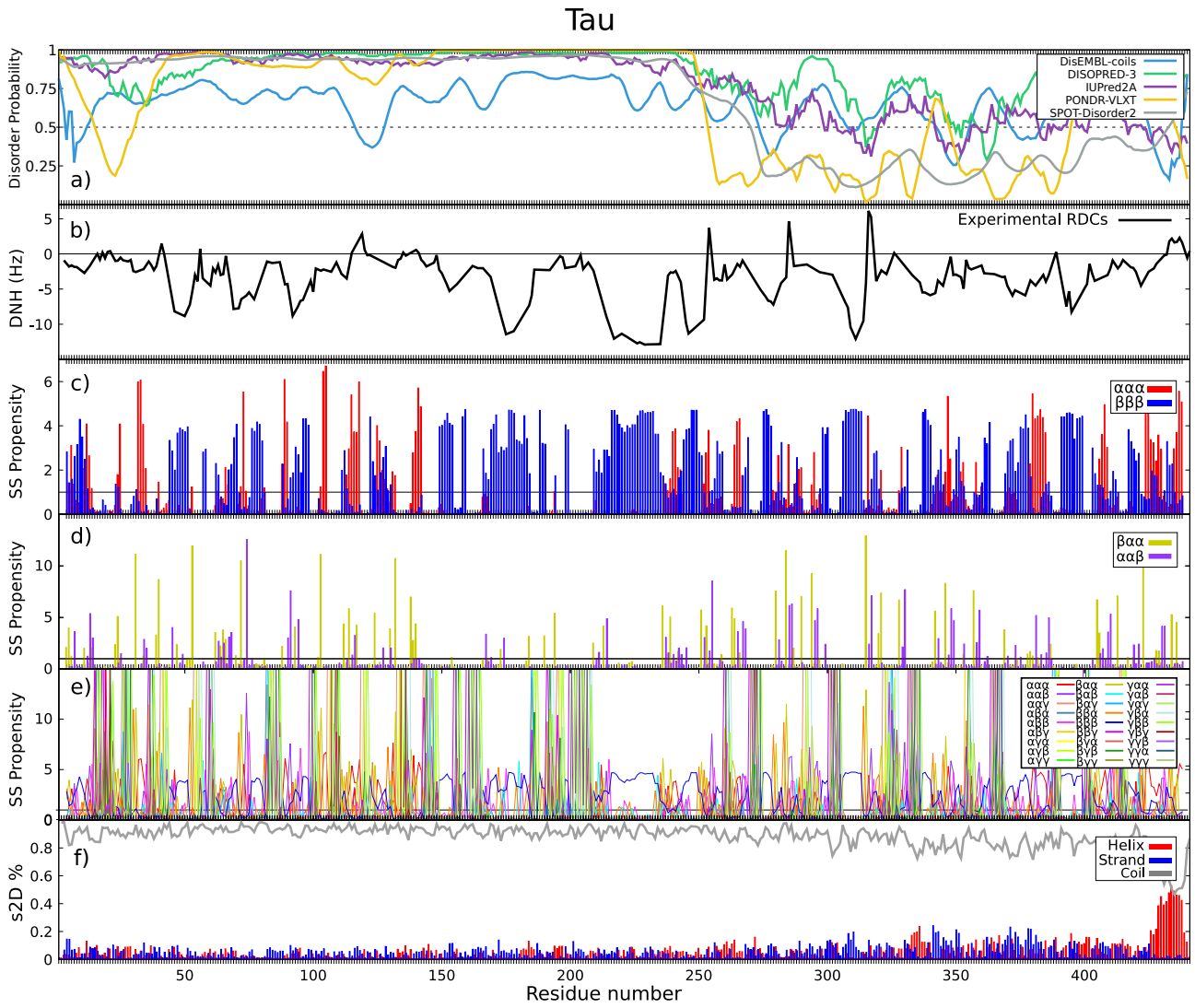


Figure S11: LS2P secondary structure prediction for Tau. (a) Disorder probability predictions. (b) Experimental RDC profile. (c) Helical ($\alpha\alpha\alpha$) and extended ($\beta\beta\beta$) propensities predicted by LS2P. (d) $\beta\alpha\alpha$ and $\alpha\alpha\beta$ propensities predicted by LS2P, allowing to identify β-turns. (e) Propensities predicted by LS2P for the 27 structural classes. For clarity purposes, values are truncated at 15. (f) Secondary structure populations predicted by s2D.