

Supplementary Information

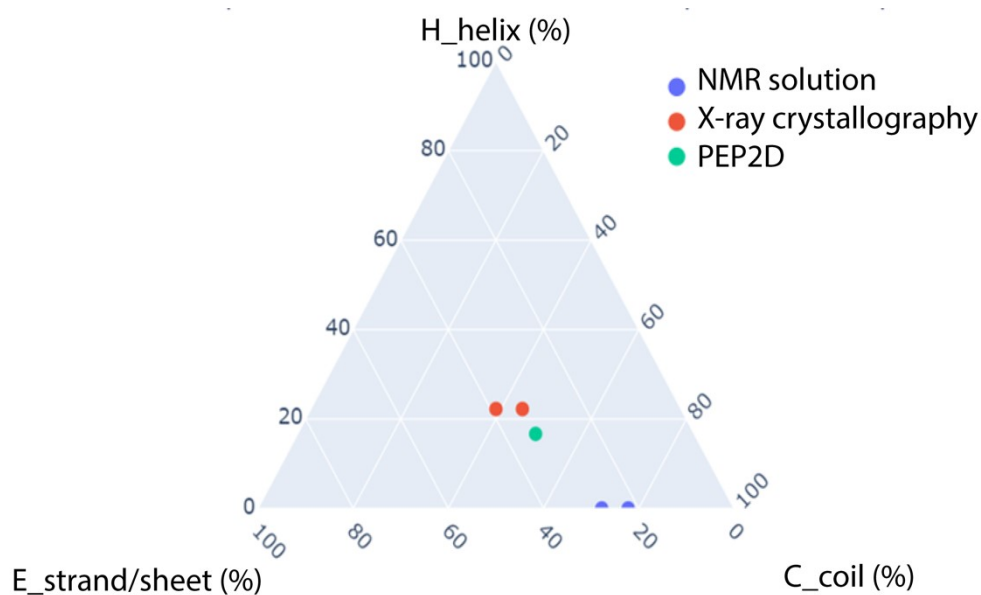


Figure S1. Differences in secondary structures annotations for PDBe structures corresponding to the same sequence under different experimental conditions. For example, five tridimensional structures of Human Beta-Defensin 1 – Nuclear Magnetic Resonance (NMR) solution structures 1KJ5 and 1E4S, X-ray crystal structures 1IJU and 1IJV, and secondary structure prediction using PEP2D.

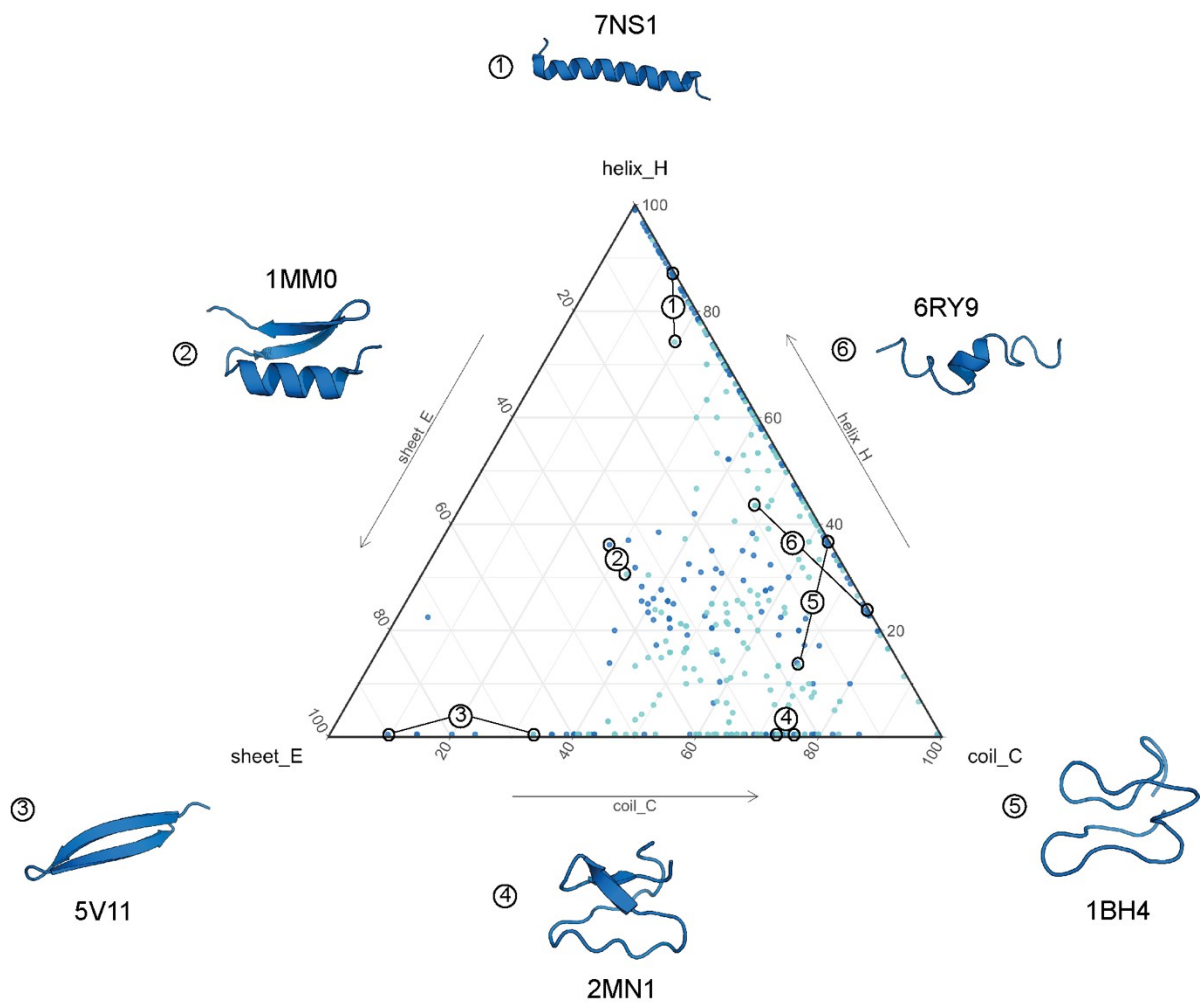


Figure S2. Ternary plot illustrating the secondary structure compositions for the 261 GRAMPA-related PDB structures as ground truth (blue) and their corresponding PEP2D predictions (cyan). The following six examples serve as structural markers; (1) pepG1 (PDB ID: 7NS1), (2) termicin (1MM0), (3) synthetic arenicin-3 analogue (5V11), (4) kalata B1[W23WW] (2MN1), (5) circulin A (1BH4) and (6) pleurocidin-like peptide 1a-1 (6RY9). The lines indicate the error between the three-state secondary structures (H, E, C) measured between PEP2D predictions and corresponding experimental references.

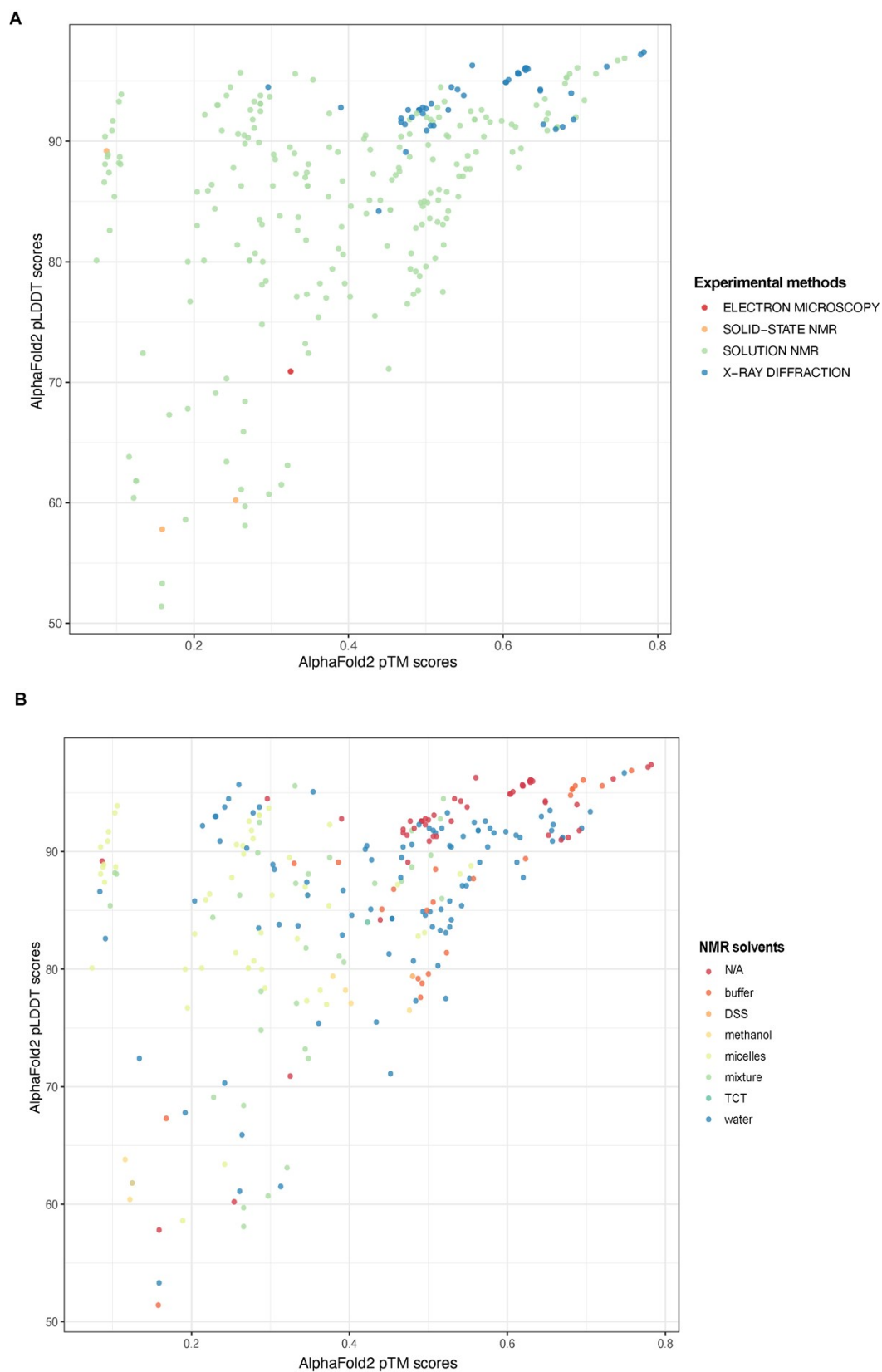


Figure S3. AlphaFold2 predictions for the 261 GRAMPA-related sequences. Scatterplot showing the corresponding predicted local distance difference test (pLDDT) and predicted template modelling (pTM) scores for all best models (rank₁) grouped by the used experimental methods (A) and different NMR solvents (B).

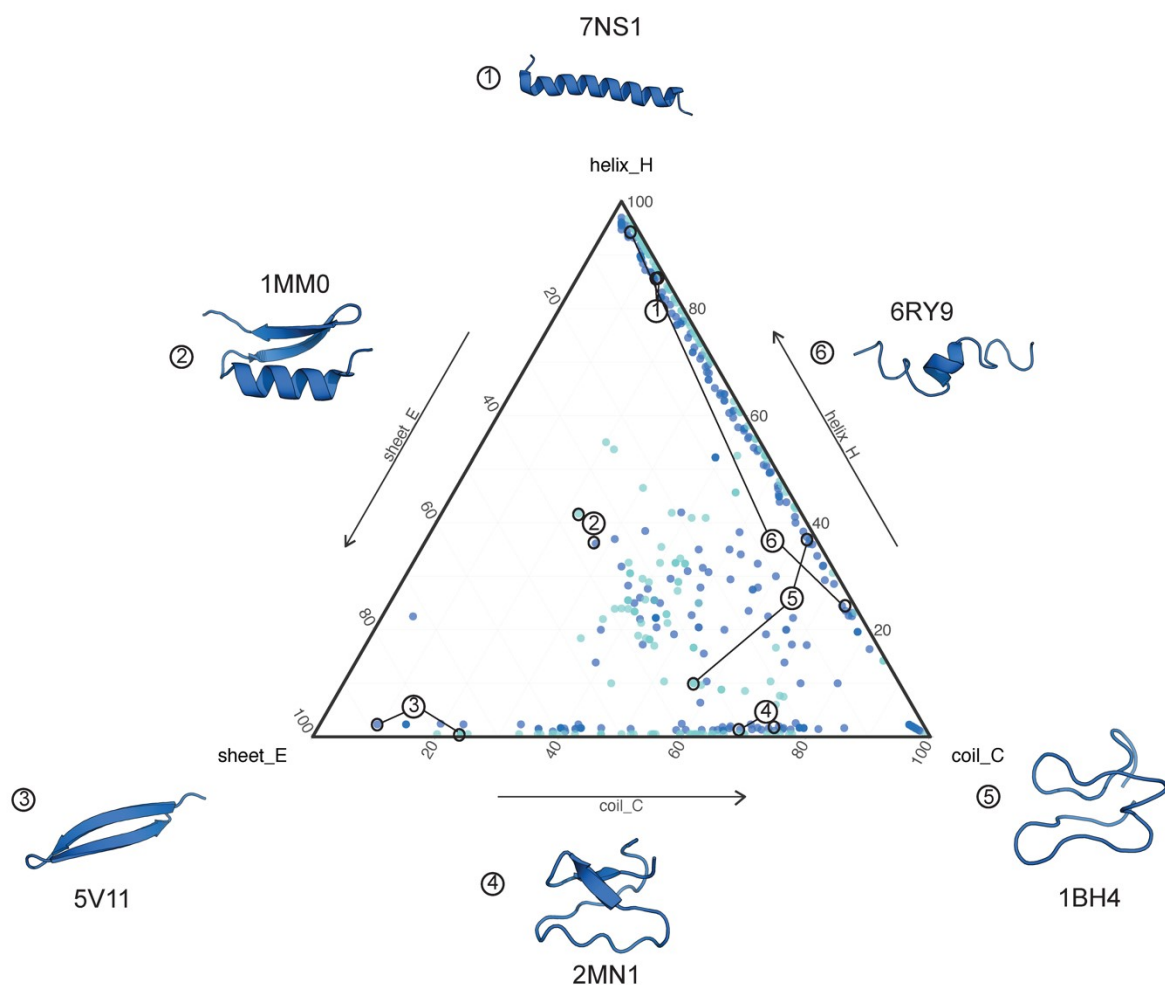


Figure S4. Ternary plot illustrating the secondary structure compositions for the 261 GRAMPA-related PDB structures as ground truth (blue) and their corresponding AlphaFold2 predictions (cyan). The following six examples serve as structural markers; (1) pepG1 (PDB ID: 7NS1), (2) termicin (1MM0), (3) synthetic arenicin-3 analogue (5V11), (4) kalata B1[W23WW] (2MN1), (5) circulin A (1BH4) and (6) pleurocidin-like peptide 1a-1 (6RY9). The lines indicate the error between the three-state secondary structures (H, E, C) measured between PEP2D predictions and corresponding experimental references.

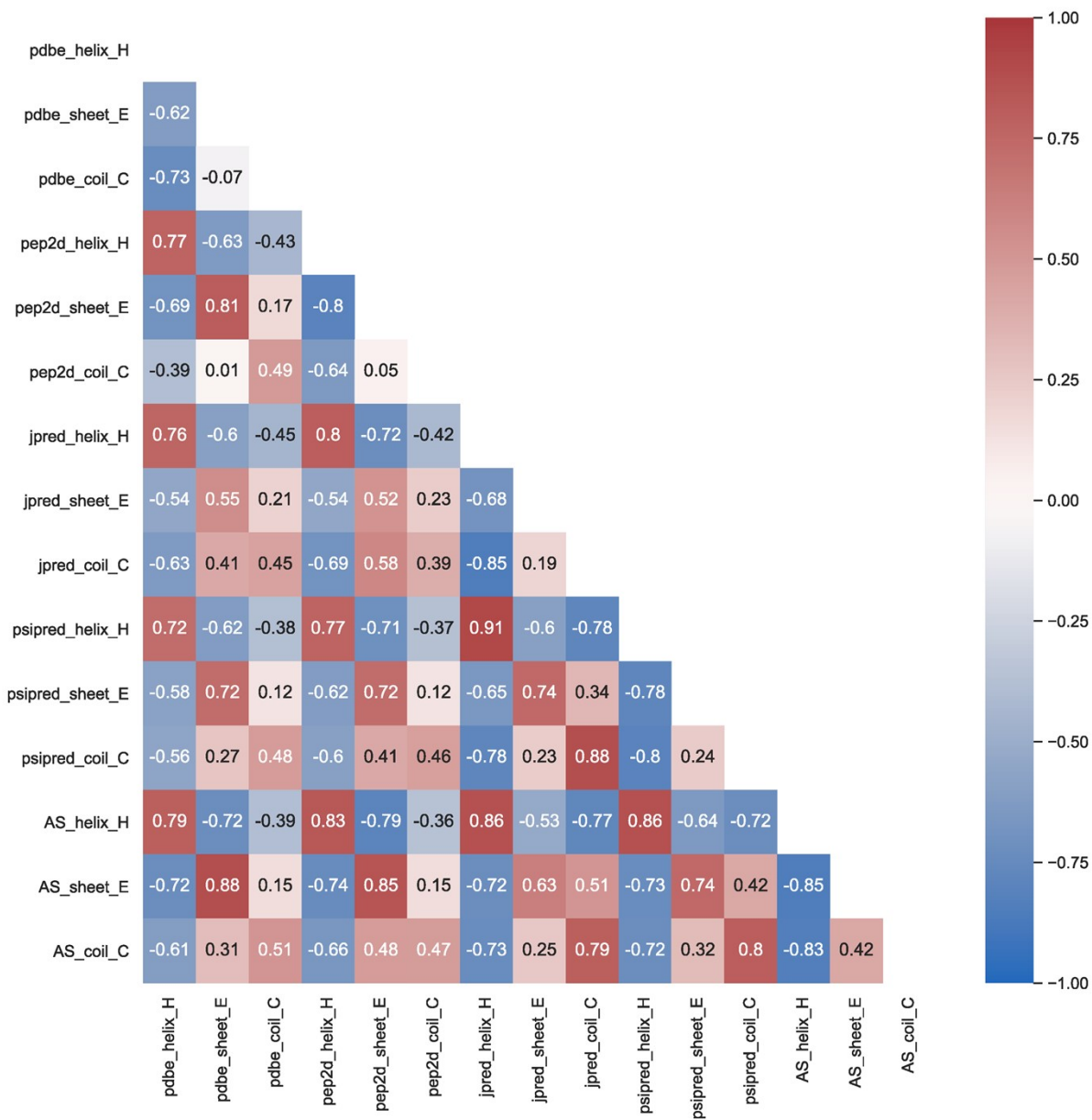


Figure S5. Pearson correlation coefficients between the secondary structure compositions from the PDBE structures (as ground truth) and the protein structure predictions (PEP2D, for the 261 GRAMPA sequences/structures).