Supplemental information

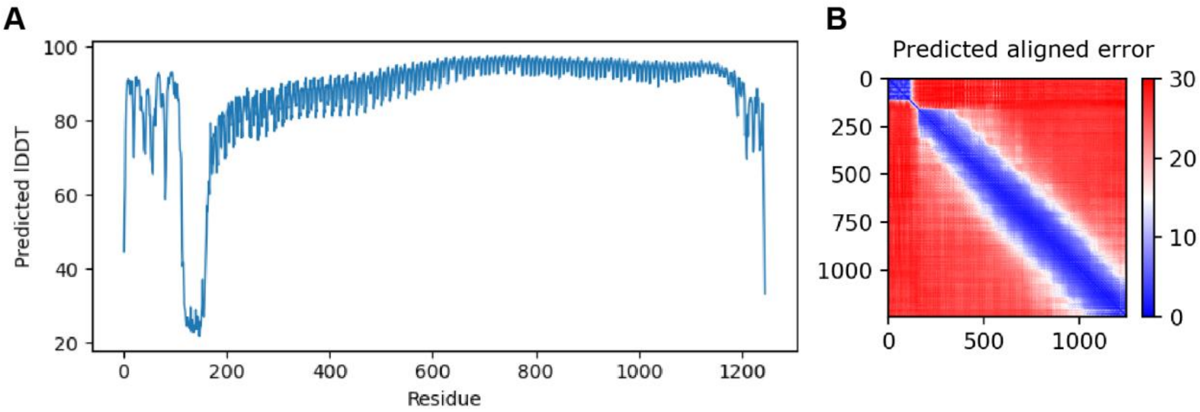

Figure S1. AlphaFold shows high confidence in overall fold of the model. A) Predicted location dependent difference test (pLDDT) shows the residue-by-residue confidence of the model generated by AlphaFold. Low values may indicate low-confidence or intrinsic disorder. **B)** Predicted aligned error (pAE) plots indicate the confidence in the relative orientation of the models. The x- and y-axes indicate residue position of the model, with low (blue) values indicating high confidence and high (red) indicating low. Rigid domains often appear as squares along the diagonal axis.

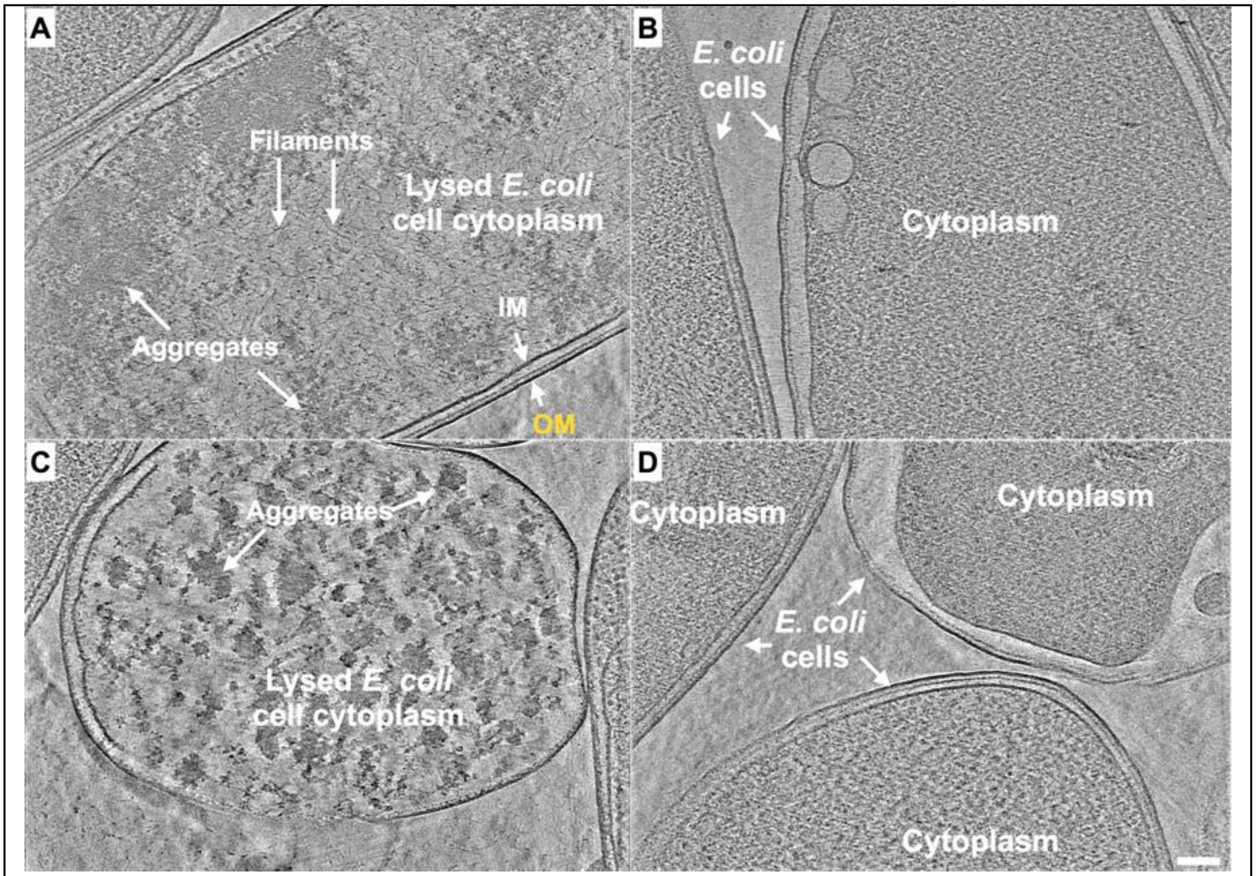

Figure S4. *E. coli* expressing INP mutant lacking R-coils show no fibre clusters as observed in those cells overexpressing wild-type INP. A-D) Representative snapshots from 3-D cryo-tomograms showing cytoplasmic and extracellular features of various *E. coli* cells overexpressing an INP mutant in which all but the C-terminal R-coil have been replaced by WO-coils. All four images are in the same scale and the scale bar represents 100 nm.

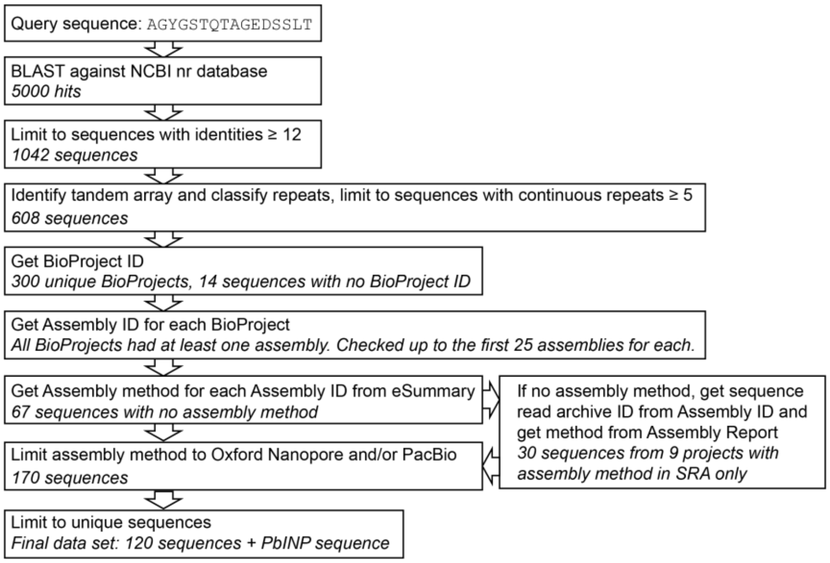

Figure S6. Flowchart and quality control steps in sequence selection for bioinformatic analysis. Ten known INPs from literature were used to generate a consensus sequence for WO-coils which was then used as a query in a BLAST against NCBI's non-redundant protein database to identify INPs. NCBI E-Utils were used to generate a data set using only genes from long-read DNA sequencing data.