Supplementary Materials for:

BioGraphX-Sol: A Structure-Free, Fully Interpretable Protein Solubility Predictor Exceeds AlphaFold3-Based Methods and Reveals Nine Universal Cross-Species Determinants

Abubakar Saeed[1,*]

[1] Independent Researcher, Chiniot, Pakistan
[1]Corresponding author: abubakarsaeed915@gmail.com

Table S1: Complete BioGraphX-Sol feature list with group assignments, statistic type, and definitions (200 features)

Feature	Group	Statistic	Definition / Interpretation
Graph-Derived Features			
Degree_mean	Graph	mean	Average number of edges per residue (connectivity).
Degree_std	Graph	std	Variability of residue connectivity.
Degree_max	Graph	max	Maximum connectivity (hub residues).
Degree_Log_mean	Graph	mean	Log-transformed mean degree.
Degree_Log_max	Graph	max	Log-transformed maximum degree.
Weighted_Degree_mean	Graph	mean	Average sum of edge weights (interaction strength).
Weighted_Degree_std	Graph	std	Variability of total interaction strength.
Weighted_Degree_max	Graph	max	Maximum interaction strength (aggregation-prone hubs).
Weighted_Degree_min	Graph	min	Minimum interaction strength.
Weighted_Degree_Log_mean	Graph	mean	Log-transformed mean weighted degree.
Weighted_Degree_Log_std	Graph	std	Log-transformed weighted degree variability.
Weighted_Degree_Log_max	Graph	max	Log-transformed maximum weighted degree.
Weighted_Degree_Log_min	Graph	min	Log-transformed minimum weighted degree.
Betweenness_mean	Graph	mean	Average shortest-path centrality (buried core).
Betweenness_std	Graph	std	Heterogeneity of centrality (structural irregularity).
Betweenness_max	Graph	max	Peak centrality (structural hubs).
Betweenness_Log_mean	Graph	mean	Log-transformed mean betweenness.
Betweenness_Log_std	Graph	std	Log-transformed betweenness variability.
Betweenness_Log_max	Graph	max	Log-transformed maximum betweenness.

Feature	Group	Statistic	Definition / Interpretation
Closeness_mean	Graph	mean	Average closeness centrality (compactness).
Closeness_std	Graph	std	Variability of closeness (structural heterogeneity).
Closeness_max	Graph	max	Maximum closeness (central core).
Closeness_min	Graph	min	Minimum closeness (peripheral residues).
Closeness_Log_mean	Graph	mean	Log-transformed mean closeness.
Closeness_Log_std	Graph	std	Log-transformed closeness variability.
Closeness_Log_max	Graph	max	Log-transformed maximum closeness.
Closeness_Log_min	Graph	min	Log-transformed minimum closeness.
Clustering_mean	Graph	mean	Average local clustering coefficient (cohesiveness).
Clustering_std	Graph	std	Variability of clustering (ordered/disordered patches).
Hybrid_Ratio_mean	Graph	mean	Fraction of edges with two interaction types.
Hybrid_Ratio_std	Graph	std	Variability of hybrid interaction density.
Local_Edge_Density_mean	Graph	mean	Average packing density in 21-residue windows.
Local_Edge_Density_std	Graph	std	Variability of local packing density.
Local_Edge_Density_max	Graph	max	Maximum local packing density (compact regions).
Local_Edge_Density_min	Graph	min	Minimum local packing density (disordered loops).
Surface_Proxy_mean	Graph	mean	Average inferred solvent exposure from graph centrality.
Surface_Proxy_std	Graph	std	Variability of inferred surface exposure.
Surface_Proxy_max	Graph	max	Maximum inferred surface exposure.
Frustration_mean	Graph	mean	Average variance of incident edge energies (conflicting constraints).
Frustration_std	Graph	std	Heterogeneity of frustration across residues.
Frustration_Entropy_Win7_std	Graph	std	Local variability of frustration entropy (7-residue windows).
Contact_Order_mean	Graph	mean	Average sequence separation of non-backbone edges (folding order).
Contact_Order_std	Graph	std	Variability of contact order.
Global Physicochemical Features			
pI	Global	global	Isoelectric point; proteins near pH 7 have low net charge $\rightarrow$ aggregation.
Charge_Density	Global	global	Net charge per residue.
Net_Charge_pH7	Global	global	Total charge at neutral pH (proxy).
Net_Charge_Global_mean	Global	mean	Whole-protein net charge (averaged over seeds).
Net_Charge_Global_min	Global	min	Minimum net charge across seeds.
Net_Charge_Global_max	Global	max	Maximum net charge across seeds.
Net_Charge_Global_std	Global	std	Variability of net charge.
Aromaticity	Global	global	Fraction of aromatic residues (F, Y, W).
Instability_Index	Global	global	Guruprasad–Pandit instability index.
Aliphatic_Index	Global	global	Ikai aliphatic index (A, V, I, L content).

Feature	Group	Statistic	Definition / Interpretation
Cysteine_Count	Global	global	Number of cysteines (disulfide potential).
GRAVY_Global_min	Global	min	Grand average hydropathy.
Max_Hydrophobic_Stretch	Global	global	Longest contiguous run of hydrophobic residues.
Mean_Hydrophobic_Stretch_Length	Global	global	Average length of hydrophobic stretches ≥ 5.
Mean_WALTZ	Global	global	Mean amyloidogenic propensity (WALTZ scale).
Sequence Physicochemical Features			
Local_Hydrophobicity_mean	SeqPhys	mean	Average 7-residue hydrophobicity.
Local_Hydrophobicity_std	SeqPhys	std	Variability of local hydrophobicity.
Local_Hydrophobicity_max	SeqPhys	max	Peak local hydrophobicity (aggregation patches).
Local_Hydrophobicity_min	SeqPhys	min	Minimum local hydrophobicity.
Local_Charge_mean	SeqPhys	mean	Average 7-residue charge.
Local_Charge_std	SeqPhys	std	Variability of local charge.
Local_Charge_max	SeqPhys	max	Maximum local positive/negative charge.
Local_Charge_min	SeqPhys	min	Minimum local charge.
Local_Charge_LargeWin_mean	SeqPhys	mean	Average 15-residue charge (larger window).
Local_Charge_LargeWin_max	SeqPhys	max	Maximum 15-residue charge.
Local_Helix_Propensity_mean	SeqPhys	mean	Average alpha-helix propensity in 7-residue windows.
Local_Helix_Propensity_std	SeqPhys	std	Variability of helix propensity.
Local_Helix_Propensity_max	SeqPhys	max	Maximum helix propensity.
Local_Helix_Propensity_min	SeqPhys	min	Minimum helix propensity.
Local_Sheet_Propensity_mean	SeqPhys	mean	Average beta-sheet propensity.
Local_Sheet_Propensity_std	SeqPhys	std	Variability of beta-sheet propensity (aggregation risk).
Local_Sheet_Propensity_max	SeqPhys	max	Maximum beta-sheet propensity.
Local_Sheet_Propensity_min	SeqPhys	min	Minimum beta-sheet propensity.
Predicted_Helix_mean	SeqPhys	mean	Fraction of residues predicted helical.
Predicted_Helix_std	SeqPhys	std	Variability of helical content.
Predicted_Sheet_mean	SeqPhys	mean	Fraction of residues predicted sheet.
Predicted_Coil_mean	SeqPhys	mean	Fraction of residues predicted coil.
Predicted_Coil_std	SeqPhys	std	Variability of coil content.
Predicted_Coil_max	SeqPhys	max	Maximum coil fraction.
Disorder_Res_Fraction_min	SeqPhys	min	Minimum fraction of disorder-promoting residues in a window.
Order_Res_Fraction_mean	SeqPhys	mean	Average fraction of order-promoting residues.
Order_Res_Fraction_std	SeqPhys	std	Variability of order fraction.
Order_Res_Fraction_max	SeqPhys	max	Maximum order fraction.
Disorder_AA_Scale_mean	SeqPhys	mean	Average disorder propensity (amino acid scale).
Disorder_AA_Scale_std	SeqPhys	std	Variability of disorder propensity.
Disorder_AA_Scale_Local_mean	SeqPhys	mean	Average local (7-residue) disorder propensity.

Feature	Group	Statistic	Definition / Interpretation
Disorder_AA_Scale_Local_std	SeqPhys	std	Variability of local disorder propensity.
Disorder_AA_Scale_Local_min	SeqPhys	min	Minimum local disorder propensity.
Disorder_AA_Scale_Local_max	SeqPhys	max	Maximum local disorder propensity.
Disorder_AA_Scale_Cterm	SeqPhys	Cterm	Mean disorder propensity at C-terminal region.
Global_Disorder_Composition_Bias_mean	SeqPhys	mean	Ratio of total disorder-promoting to order-promoting residues.
Global_Disorder_Composition_Bias_std	SeqPhys	std	Variability of composition bias.
Global_Disorder_Composition_Bias_min	SeqPhys	min	Minimum composition bias.
Global_Disorder_Composition_Bias_max	SeqPhys	max	Maximum composition bias.
NTerm_Distance_mean	SeqPhys	mean	Average relative position from N-terminus.
NTerm_Distance_std	SeqPhys	std	Variability of N-terminal distance.
CTerm_Distance_mean	SeqPhys	mean	Average relative position from C-terminus.
CTerm_Distance_std	SeqPhys	std	Variability of C-terminal distance.
Length_Normalized_std	SeqPhys	std	Normalized length (L/500) variability.
Hydro_Autocorr_Lag1_mean	SeqPhys	mean	Average lag-1 autocorrelation of hydrophobicity.
Hydro_Autocorr_Lag1_std	SeqPhys	std	Variability of hydrophobicity autocorrelation.
Hydro_Autocorr_Lag1_max	SeqPhys	max	Maximum hydrophobicity autocorrelation.
Charge_Autocorr_Lag1_mean	SeqPhys	mean	Average lag-1 autocorrelation of charge.
Charge_Autocorr_Lag1_std	SeqPhys	std	Variability of charge autocorrelation.
Charge_Autocorr_Lag1_max	SeqPhys	max	Maximum charge autocorrelation.
OrderDisorder_Pattern_Entropy_mean	SeqPhys	mean	Average entropy of order/disorder alternation (11-residue window).
OrderDisorder_Pattern_Entropy_std	SeqPhys	std	Variability of pattern entropy.
OrderDisorder_Pattern_Entropy_min	SeqPhys	min	Minimum pattern entropy.
Sequence_Complexity_std	SeqPhys	std	Variability of Shannon entropy in 21-residue windows.
Sequence_Complexity_max	SeqPhys	max	Maximum sequence complexity.
Complexity_Std_Win7_min	SeqPhys	min	Minimum complexity standard deviation in 7-residue windows.
Complexity_Std_Win7_max	SeqPhys	max	Maximum complexity standard deviation.
Complexity_W5_mean	SeqPhys	mean	Average Shannon entropy in 5-residue windows (short repeats).
Complexity_W5_std	SeqPhys	std	Variability of 5-residue complexity.
Complexity_W5_min	SeqPhys	min	Minimum 5-residue complexity.
Complexity_W25_std	SeqPhys	std	Variability of 25-residue complexity.
Complexity_W25_min	SeqPhys	min	Minimum 25-residue complexity.
Complexity_W50_mean	SeqPhys	mean	Average 50-residue complexity (long low-complexity tracts).
Complexity_W50_std	SeqPhys	std	Variability of 50-residue complexity.
Complexity_W50_min	SeqPhys	min	Minimum 50-residue complexity.
Complexity_W50_max	SeqPhys	max	Maximum 50-residue complexity.
Complexity_Scale_Variance_min	SeqPhys	min	Minimum variance across the four complexity scales.

Feature	Group	Statistic	Definition / Interpretation
Complexity_Scale_Variance_std	SeqPhys	std	Variability of multi-scale complexity consistency.
In_Motif_std	SeqPhys	std	Variability of known short linear motif flags.
Local_Motif_Density_mean	SeqPhys	mean	Average density of motifs in 7-residue windows.
Local_Motif_Density_std	SeqPhys	std	Variability of motif density.
Local_Bulkiness_mean	SeqPhys	mean	Average side-chain bulkiness.
Local_Bulkiness_min	SeqPhys	min	Minimum bulkiness.
Local_Bulkiness_max	SeqPhys	max	Maximum bulkiness.
IDP / Polymer Physics Features			
FCR_mean	IDP	mean	Average fraction of charged residues (K,R,D,E) in 21-residue windows.
FCR_min	IDP	min	Minimum charged residue fraction.
NCPR_mean	IDP	mean	Average net charge per residue in 21-residue windows.
NCPR_std	IDP	std	Variability of net charge per residue.
Polyampholyte_Score_mean	IDP	mean	Average FCR $\times$ (1−—NCPR—): charge-balanced polyampholyte character.
Polyampholyte_Score_std	IDP	std	Variability of polyampholyte score.
Polyampholyte_Score_min	IDP	min	Minimum polyampholyte score.
Sticker_Density_mean	IDP	mean	Average fraction of sticker residues (Y,F,W,R).
Sticker_Density_std	IDP	std	Variability of sticker density.
Sticker_Density_max	IDP	max	Maximum sticker density.
Sticker_Density_min	IDP	min	Minimum sticker density.
Spacer_Density_mean	IDP	mean	Average fraction of spacer residues (G,S,P).
Spacer_Density_std	IDP	std	Variability of spacer density.
Spacer_Density_max	IDP	max	Maximum spacer density.
Sticker_Dispersion_mean	IDP	mean	Average regularity of sticker spacing.
Sticker_Dispersion_std	IDP	std	Variability of sticker dispersion.
Sticker_Dispersion_min	IDP	min	Minimum sticker dispersion.
Local_Kappa_mean	IDP	mean	Average local deviation of absolute charge density from global net charge magnitude (15-residue windows).
Local_Kappa_min	IDP	min	Minimum local deviation (high = uniform charge distribution).
Local_Kappa_max	IDP	max	Maximum local deviation (low = charge-depleted patches).
Uversky_Distance_mean	IDP	mean	Average per-residue distance from Uversky disorder boundary.
Uversky_Distance_std	IDP	std	Variability of Uversky distance.
Uversky_Distance_max	IDP	max	Maximum Uversky distance.
Uversky_Distance_min	IDP	min	Minimum Uversky distance.
MoRF_Core_Score_std	IDP	std	Variability of molecular recognition feature core score.
MoRF_Core_Score_Nterm	IDP	Nterm	Mean MoRF core score at N-terminus.
MoRF_Core_Score_Cterm	IDP	Cterm	Mean MoRF core score at C-terminus.
MoRF_Flank_Score_mean	IDP	mean	Average flanking disorder composition around potential MoRFs.

Feature	Group	Statistic	Definition / Interpretation
MoRF_Flank_Score_std	IDP	std	Variability of MoRF flank score.
MoRF_Flank_Score_min	IDP	min	Minimum MoRF flank score.
MoRF_Flank_Score_max	IDP	max	Maximum MoRF flank score.
MoRF_Dip_Score_mean	IDP	mean	Average hydrophobic dip (local vs context).
MoRF_Dip_Score_max	IDP	max	Maximum hydrophobic dip.
MoRF_Dip_Score_min	IDP	min	Minimum hydrophobic dip.
Ramachandran_Flexibility_mean	IDP	mean	Average backbone flexibility (Ramachandran-derived).
Ramachandran_Flexibility_std	IDP	std	Variability of backbone flexibility.
ProGly_Density_mean	IDP	mean	Average fraction of Pro and Gly in 9-residue windows.
ProGly_Density_std	IDP	std	Variability of Pro/Gly density.
ProGly_Density_max	IDP	max	Maximum Pro/Gly density.
Aggregation-Specific Features			
Agg_Mean	AGG	global	Average aggregation propensity (hydrophobicity + sheet $-$ 0.5|charge|).
Agg_Max	AGG	global	Maximum aggregation propensity.
Agg_Top5_Mean	AGG	global	Mean of top-5 aggregation scores.
Agg_Hotspot_Count	AGG	global	Number of residues with aggregation score > 0.5.
Agg_Hotspot_Variance	AGG	global	Variance of hotspot positions (clustered vs dispersed).
SCD	AGG	global	Sequence charge decoration: length-normalized sum of $q_i q_j / \sqrt{\lvert i-j \rvert}$.
Charge_Asymmetry	AGG	global	Difference in net charge between N- and C-terminal halves.
Surface & Terminal Conditioning Features			
Local_Hydrophobicity_surf	SurfTerm	surf	Mean hydrophobicity over surface-exposed residues only.
Local_Hydrophobicity_Nterm	SurfTerm	Nterm	Mean hydrophobicity over first 30 residues.
Local_Hydrophobicity_Cterm	SurfTerm	Cterm	Mean hydrophobicity over last 30 residues.
Local_Charge_surf	SurfTerm	surf	Mean charge over surface residues.
Local_Charge_Nterm	SurfTerm	Nterm	Mean charge over first 30 residues.
Local_Charge_Cterm	SurfTerm	Cterm	Mean charge over last 30 residues.
Disorder_AA_Scale_surf	SurfTerm	surf	Mean disorder propensity over surface residues.
Disorder_AA_Scale_Nterm	SurfTerm	Nterm	Mean disorder propensity at N-terminus.
Local_Sheet_Propensity_surf	SurfTerm	surf	Mean beta-sheet propensity over surface residues.
Local_Sheet_Propensity_Nterm	SurfTerm	Nterm	Mean beta-sheet propensity at N-terminus.
Local_Sheet_Propensity_Cterm	SurfTerm	Cterm	Mean beta-sheet propensity at C-terminus.
Sticker_Density_surf	SurfTerm	surf	Mean sticker density over surface residues.
Sticker_Density_Nterm	SurfTerm	Nterm	Mean sticker density at N-terminus.

Feature	Group	Statistic	Definition / Interpretation
Sticker_Density_Cterm	SurfTerm	Cterm	Mean sticker density at C-terminus.
Spacer_Density_surf	SurfTerm	surf	Mean spacer density over surface residues.
Spacer_Density_Nterm	SurfTerm	Nterm	Mean spacer density at N-terminus.
Contact_Order_surf	SurfTerm	surf	Mean contact order over surface residues.
Contact_Order_Nterm	SurfTerm	Nterm	Mean contact order at N-terminus.
Contact_Order_Cterm	SurfTerm	Cterm	Mean contact order at C-terminus.
Frustration_surf	SurfTerm	surf	Mean frustration over surface residues.
Frustration_Nterm	SurfTerm	Nterm	Mean frustration at N-terminus.
Frustration_Cterm	SurfTerm	Cterm	Mean frustration at C-terminus.
Hybrid_Ratio_Nterm	SurfTerm	Nterm	Mean hybrid edge ratio at N-terminus.
Nterm_Charge	SurfTerm	Nterm	Total charge in first 10 residues.
Cterm_Charge	SurfTerm	Cterm	Total charge in last 10 residues.
Nterm_HydroFraction	SurfTerm	Nterm	Fraction of hydrophobic residues in first 20 residues.
Additional Interface & Disorder Features			
Interface_Propensity_mean	Interface	mean	Average predicted interface propensity (from graph + sequence).
Interface_Propensity_std	Interface	std	Variability of interface propensity.
Interface_Propensity_max	Interface	max	Maximum interface propensity.
Interface_Propensity_min	Interface	min	Minimum interface propensity.
Disorder_from_Ifprop_mean	Interface	mean	Average disorder inferred from interface propensity ($1 -$ propensity).
Disorder_from_Ifprop_std	Interface	std	Variability of inferred disorder.
Disorder_from_Ifprop_max	Interface	max	Maximum inferred disorder.

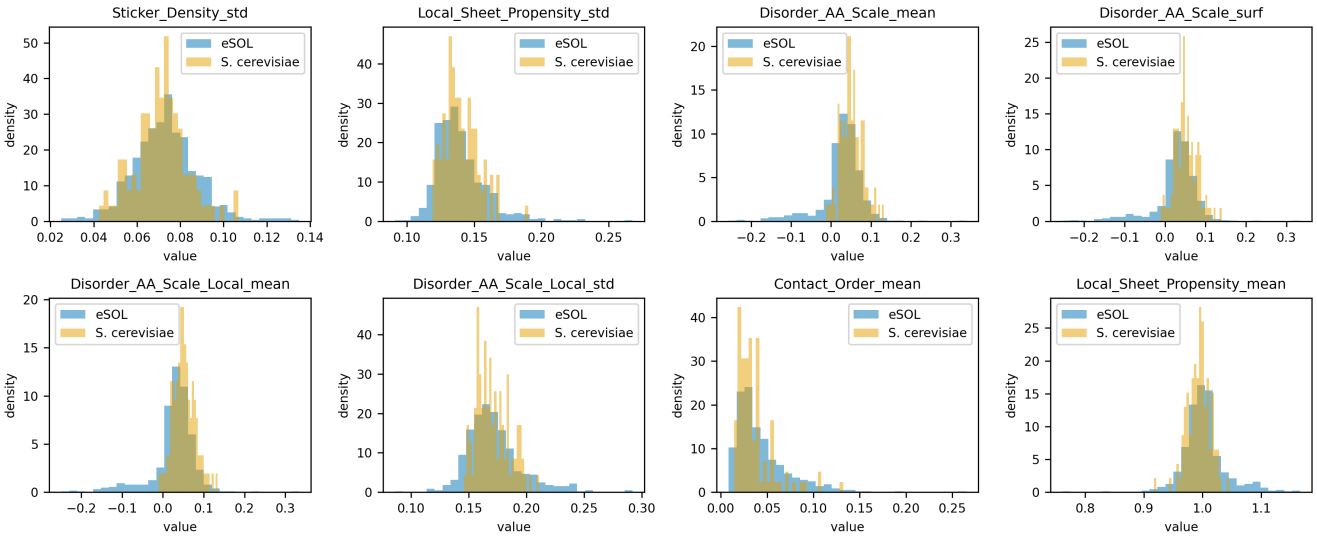


Figure S1: Distribution comparison of selected importance-divergent features on the eSOL 660-sequence test set (n=660) and the S. cerevisiae 108-sequence external set (n=108). The substantial overlap in distributions, particularly for Sticker_Density_std, indicates that importance divergence across species does not arise from simple compositional shifts.

Table S2: Alpha sweep for the XGBoost + ESM ensemble (per-seed, Method B). Alpha is the weight on the physics (XGBoost) stream; $(1 - \alpha)$ on the ESM stream. For each of the five seeds, the optimal alpha was selected by grid search maximising validation R^2 on the 268-sequence validation subset, and the corresponding seed-specific ensemble was evaluated on the full 660-sequence test set. Values are mean $\pm$ standard deviation across the five per-seed ensembles. The optimum is a broad plateau ($\Delta R^2 < 0.002$ across $\alpha \in [0.45, 0.70]$).

Alpha	Validation R^2	Test R^2	Test AUC
0.00	0.4642 ± 0.0197	0.4806 ± 0.0071	0.8721 ± 0.0042
0.05	0.4758 ± 0.0185	0.4921 ± 0.0066	0.8755 ± 0.0040
0.10	0.4865 ± 0.0173	0.5025 ± 0.0062	0.8789 ± 0.0037
0.15	0.4962 ± 0.0161	0.5118 ± 0.0057	0.8818 ± 0.0035
0.20	0.5048 ± 0.0149	0.5200 ± 0.0053	0.8847 ± 0.0032
0.25	0.5125 ± 0.0137	0.5270 ± 0.0050	0.8871 ± 0.0028
0.30	0.5191 ± 0.0125	0.5329 ± 0.0046	0.8894 ± 0.0026
0.35	0.5247 ± 0.0113	0.5378 ± 0.0043	0.8912 ± 0.0022
0.40	0.5293 ± 0.0102	0.5415 ± 0.0040	0.8928 ± 0.0019
0.45	0.5329 ± 0.0090	0.5441 ± 0.0038	0.8937 ± 0.0016
0.50	0.5355 ± 0.0079	0.5455 ± 0.0036	0.8943 ± 0.0013
0.55	$\mathbf{0.5371 \pm 0.0068}$	$\mathbf{0.5459 \pm 0.0034}$	$\mathbf{0.8946 \pm 0.0012}$
0.60	$\mathbf{0.5377 \pm 0.0057}$	$\mathbf{0.5451 \pm 0.0032}$	$\mathbf{0.8942 \pm 0.0009}$
0.65	0.5372 ± 0.0046	0.5432 ± 0.0030	0.8933 ± 0.0009
0.70	0.5358 ± 0.0036	0.5402 ± 0.0029	0.8918 ± 0.0008
0.75	0.5333 ± 0.0027	0.5361 ± 0.0029	0.8900 ± 0.0009
0.80	0.5299 ± 0.0021	0.5309 ± 0.0029	0.8878 ± 0.0009
0.85	0.5254 ± 0.0021	0.5246 ± 0.0029	0.8849 ± 0.0009
0.90	0.5199 ± 0.0027	0.5171 ± 0.0030	0.8815 ± 0.0008
0.95	0.5134 ± 0.0035	0.5085 ± 0.0031	0.8778 ± 0.0010
1.00	0.5059 ± 0.0045	0.4988 ± 0.0032	0.8734 ± 0.0012

Table S3: Alpha sweep for the Graph-NN + ESM ensemble (per-seed, Method B). Alpha is the weight on the physics (graph-only NN) stream; $(1 - \alpha)$ on the ESM stream. Same protocol as Table S2: per-seed alpha selection on the 268-sequence validation subset, evaluated on the full 660-sequence test set, then aggregated across five seeds. The larger per-seed SDs relative to the XGB+ESM ensemble reflect the additional training variability of the neural physics stream.

Alpha	Validation R^2	Test R^2	Test AUC
0.00	0.4642 ± 0.0197	0.4806 ± 0.0071	0.8721 ± 0.0042
0.05	0.4755 ± 0.0191	0.4923 ± 0.0068	0.8756 ± 0.0040
0.10	0.4857 ± 0.0186	0.5028 ± 0.0066	0.8791 ± 0.0036
0.15	0.4947 ± 0.0183	0.5120 ± 0.0066	0.8818 ± 0.0034
0.20	0.5025 ± 0.0180	0.5199 ± 0.0067	0.8845 ± 0.0032
0.25	0.5092 ± 0.0179	0.5266 ± 0.0069	0.8868 ± 0.0029
0.30	0.5148 ± 0.0179	0.5320 ± 0.0071	0.8889 ± 0.0027
0.35	0.5192 ± 0.0179	0.5361 ± 0.0074	0.8905 ± 0.0024
0.40	0.5224 ± 0.0181	0.5389 ± 0.0078	0.8917 ± 0.0020
0.45	0.5245 ± 0.0183	0.5405 ± 0.0082	0.8923 ± 0.0018
0.50	$\mathbf{0.5254 \pm 0.0186}$	$\mathbf{0.5408 \pm 0.0086}$	$\mathbf{0.8925 \pm 0.0018}$
0.55	0.5251 ± 0.0189	0.5398 ± 0.0090	0.8923 ± 0.0020
0.60	0.5237 ± 0.0193	0.5376 ± 0.0095	0.8913 ± 0.0020
0.65	0.5212 ± 0.0198	0.5340 ± 0.0099	0.8898 ± 0.0021
0.70	0.5174 ± 0.0202	0.5292 ± 0.0104	0.8876 ± 0.0024
0.75	0.5126 ± 0.0207	0.5232 ± 0.0110	0.8851 ± 0.0027
0.80	0.5065 ± 0.0212	0.5158 ± 0.0115	0.8820 ± 0.0031
0.85	0.4993 ± 0.0217	0.5072 ± 0.0121	0.8785 ± 0.0035
0.90	0.4910 ± 0.0223	0.4973 ± 0.0128	0.8746 ± 0.0039
0.95	0.4814 ± 0.0228	0.4861 ± 0.0134	0.8702 ± 0.0045
1.00	0.4708 ± 0.0234	0.4737 ± 0.0142	0.8651 ± 0.0049

Table S4: Forward feature group ablation on the eSOL 660-sequence test set. Features are added cumulatively in the order shown. Values are mean ± standard deviation over five seeds.

Configuration	R^2	**AUC**	**MCC**	ΔR^2	**ΔAUC**
All Graph-Derived Features	0.3100±0.0025	0.7863±0.0015	0.4800±0.0083	–	–
+ Global Physicochemical	0.4431±0.0023	0.8530±0.0009	0.5110±0.0095	+0.1331±0.0034	+0.0667±0.0017
+ Sequence Physicochemical	0.4781±0.0020	0.8614±0.0010	0.5297±0.0117	+0.0350±0.0030	+0.0084±0.0014
+ IDP / Polymer Physics	0.4959±0.0016	0.8714±0.0010	0.5543±0.0056	+0.0178±0.0026	+0.0100±0.0015
+ Aggregation-Specific (novel)	0.5008±0.0023	0.8749±0.0014	0.5697±0.0085	+0.0049±0.0028	+0.0035±0.0017
+ Surface & Terminal Conditioning	0.5004±0.0021	0.8745±0.0018	0.5578±0.0065	−0.0004±0.0031	−0.0004±0.0023

Table S5: Leave-one-out feature group ablation on the eSOL 660-sequence test set. Each group is removed from the complete 200-feature model. Values are mean ± standard deviation over five seeds. The final column indicates the number of seeds (out of five) showing a decrease in R^2 upon removal.

Group Removed	R^2	**AUC**	**MCC**	ΔR^2	**ΔAUC**	**Negative** ΔR^2
All Graph-Derived Features	0.4990±0.0011	0.8727±0.0007	0.5554±0.0064	-0.0027 ± 0.0023	-0.0031 ± 0.0010	5/5
Global Physicochemical	0.5058±0.0013	0.8763±0.0005	0.5687±0.0042	$+0.0040 \pm 0.0025$	$+0.0005 \pm 0.0014$	1/5
Sequence Physicochemical	0.4952±0.0007	0.8742±0.0006	0.5529±0.0096	-0.0066 ± 0.0025	-0.0016 ± 0.0020	5/5
IDP / Polymer Physics	0.4915±0.0015	0.8697±0.0010	0.5600±0.0072	-0.0103 ± 0.0027	-0.0061 ± 0.0018	5/5
Aggregation-Specific	0.4971±0.0020	0.8717±0.0012	0.5573±0.0073	-0.0047 ± 0.0032	-0.0041 ± 0.0018	5/5
Surface & Terminal	0.5006±0.0028	0.8749±0.0019	0.5617±0.0078	-0.0012 ± 0.0030	-0.0009 ± 0.0024	3/5

Table S6: Leakage-free holdout evaluation of the XGBoost + ESM ensemble on the 392-sequence subset (rows 268–659 of the test set) that was never used for alpha selection. Per-seed values are mean ± SD across five per-seed ensembles; bootstrap 95% confidence intervals (2,000 resamples) are computed on the seed-averaged ensemble prediction and capture test-set sampling uncertainty. Values are fully consistent with the Method B protocol used in Section 3.2.

Metric	**392-seq holdout**	**660-seq full test**
R^2	0.5483 ± 0.0042 [0.4951, 0.6062]	0.5451 ± 0.0041 [0.5026, 0.5919]
AUC	0.8894 ± 0.0017 [0.8603, 0.9208]	0.8939 ± 0.0007 [0.8728, 0.9181]
PR-AUC	0.8675 ± 0.0020 [0.8225, 0.9094]	0.8774 ± 0.0010 [0.8489, 0.9069]