Supplemental information

Table S1: Full summary of performance for different methods.

Method	Sequence representation	Structure representation	Learnt with ML	R
INFUSSE (S_{block} + G_{block})	One-hot encoding & LLM embeddings	Weighted Gaussian graph ($\eta = 8$)	$T_1, T_2, T_3, \mathbf{W}^{(l)}, t$	0.71
		GNM ($\epsilon = 10$Å)		0.70
		GNM ($\epsilon = 8$Å)		0.69
S_{block} + G_{block} (GCN instead of diff-GCN)	One-hot encoding & LLM embeddings	Weighted Gaussian graph ($\eta = 8$)	$T_1, T_2, T_3, \mathbf{W}^{(l)}$	0.69
		GNM ($\epsilon = 10$Å)		0.69
		GNM ($\epsilon = 8$Å)		0.68
S_{block} alone (no structure)	One-hot encoding & LLM embeddings	—	T_1, T_2, T_3	0.64
S_{block} (no LLM) + G_{block}	One-hot encoding	Weighted Gaussian graph ($\eta = 8$)	$T_1, T_3, \mathbf{W}^{(l)}, t$	0.55
LSTM[43] (SOTA for general proteins)	One-hot encoding	Raw coordinates $\mathbf{r}$, secondary structure, and chain breaks	LSTM weights	0.48
Laplacian pseudoinverse (no learning, baseline)	—	Weighted Gaussian graph ($\eta = 8$)	—	0.01
		GNM ($\epsilon = 10$Å)		0.01

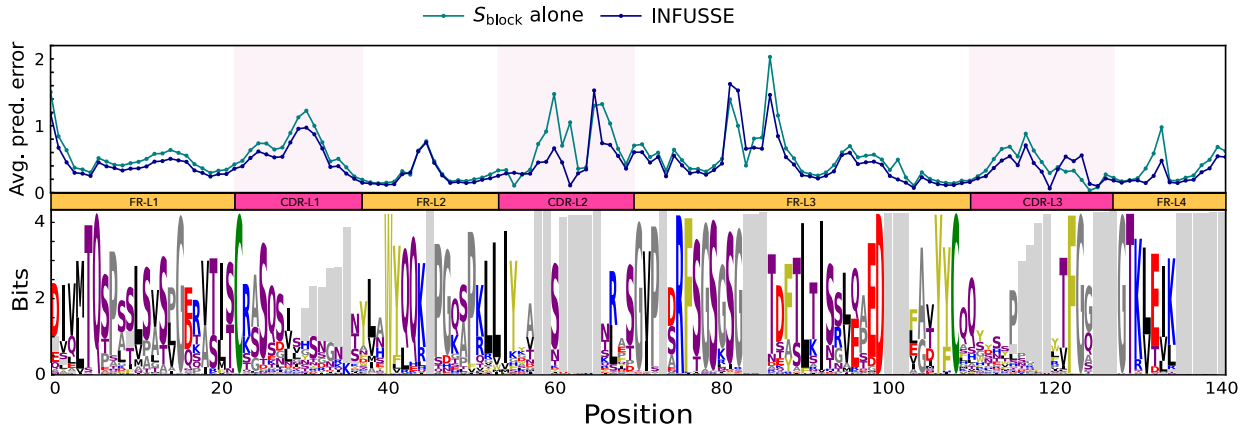

Figure S1: INFUSSE's predictions for the light chain variable region. Same plot as Figure 2A, but for the light chain variable region.

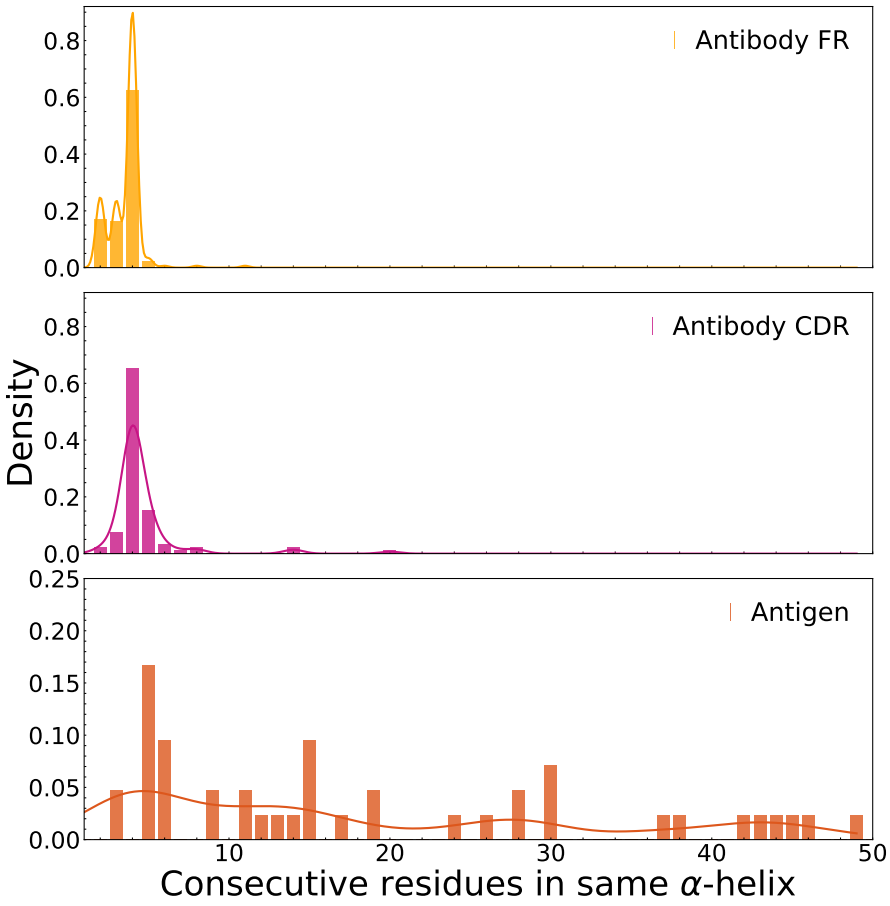

Figure S2: Density of consecutive residues classed as being part of an α-helix for the test set. Histograms depict the distribution of consecutive α-helix residues for antibody framework (FR), antibody complementarity-determining region (CDR) and antigen. Overlaid are Gaussian Kernel Density Estimator (KDE) curves, providing an estimation of the probability density function.

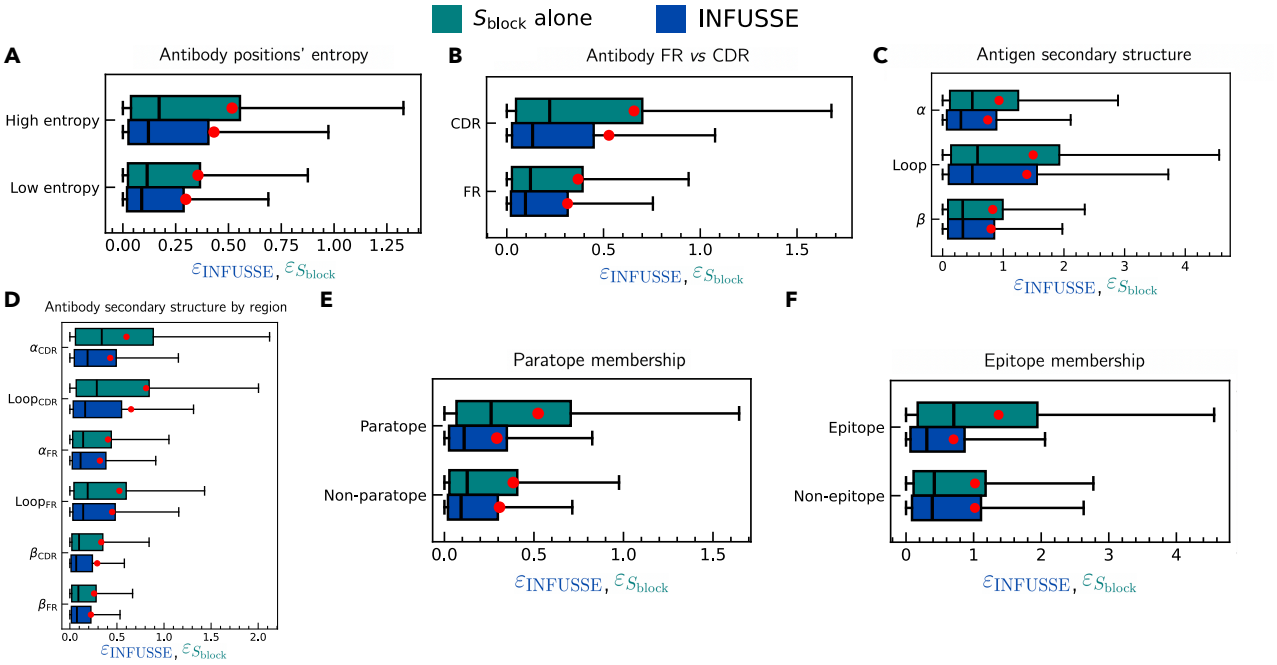

Figure S3: Prediction errors of INFUSSE, $\varepsilon^{(q)}_{\text{INFUSSE},j}$, and S_{block} alone, $\varepsilon^{(q)}_{S_{\text{block}},j}$, for each position j and test set sample q. The groups are based on: (**A**) antibody variable-region entropy (similar to Figure 2C); (**B**) antibody framework (FR) versus complementarity-determining region (CDR) (similar to Figure 3A); (**C**) antigen secondary structure (similar to Figure 3B); (**D**) antibody region and secondary structure (similar to Figure 3C); (**E**) paratope membership (similar to Figure 4B); (**F**) epitope membership (similar to Figure 4C). Red dots indicate mean values and black vertical lines denote medians.

Table S2: Summary of statistical tests on the Δ_{graph} distribution across different scenarios. Upper-triangle cells show the difference in mean $\Delta\mu$ between the two groups and the p-value, lower-triangle cells the corresponding IQR differences ΔIQR and the p-value; diagonal elements give the mean over the corresponding group (denoted by μ) and its p-value. The groups compared are based on: (**A**) antibody variable-region entropy (data shown in Figure 2C); (**B**) antibody framework (FR) versus complementarity-determining region (CDR) (data shown in Figure 3A); (**C**) antigen secondary structure (data shown in Figure 3B); (**D**) antibody region and secondary structure (data shown in Figure 3C); (**E**) paratope membership (data shown in Figure 4B); (**F**) epitope membership (data shown in Figure 4C). Statistical tests are carried out by the bootstrap as described in Methods, Section 4.10.

A

Antibody positions' entropy (Low/high threshold: 1.457 bit)

	Low	High
Low	$\mu = 0.06$ $p < 10^{-6}$	$\Delta\mu = 0.03$ $p = 3 \times 10^{-5}$
High	$\Delta\text{IQR} = 0.08$ $p < 10^{-6}$	$\mu = 0.09$ $p < 10^{-6}$

B

Antibody FR *vs* CDR

	FR	CDR
FR	$\mu = 0.05$ $p < 10^{-6}$	$\Delta\mu = 0.08$ $p < 10^{-6}$
CDR	$\Delta\text{IQR} = 0.15$ $p < 10^{-6}$	$\mu = 0.13$ $p < 10^{-6}$

C

Antigen secondary structure

	α	Loop	β
α	$\mu = 0.19$ $p < 10^{-6}$	$\Delta\mu = 0.08$ $p = 0.05$	$\Delta\mu = 0.16$ $p = 4 \times 10^{-3}$
Loop	$\Delta\text{IQR} = 0.25$ $p = 10^{-6}$	$\mu = 0.10$ $p = 7 \times 10^{-3}$	$\Delta\mu = 0.08$ $p = 0.09$(NS)
β	$\Delta\text{IQR} = 0.16$ $p = 8 \times 10^{-3}$	$\Delta\text{IQR} = 0.42$ $p < 10^{-6}$	$\mu = 0.03$ $p = 0.2$(NS)

D

Antibody secondary structure by region

	α_{CDR}	Loop_{CDR}	α_{FR}	Loop_{FR}	β_{CDR}	β_{FR}
α_{CDR}	$\mu = 0.17$ $p < 10^{-6}$	$\Delta\mu = 0.01$ $p = 0.3$(NS)	$\Delta\mu = 0.09$ $p = 7 \times 10^{-4}$	$\Delta\mu = 0.10$ $p = 4 \times 10^{-5}$	$\Delta\mu = 0.13$ $p = 10^{-6}$	$\Delta\mu = 0.14$ $p < 10^{-6}$
Loop_{CDR}	$\Delta\text{IQR} = 0.10$ $p = 5 \times 10^{-6}$	$\mu = 0.16$ $p < 10^{-6}$	$\Delta\mu = 0.07$ $p = 7 \times 10^{-6}$	$\Delta\mu = 0.08$ $p < 10^{-6}$	$\Delta\mu = 0.12$ $p < 10^{-6}$	$\Delta\mu = 0.13$ $p < 10^{-6}$
α_{FR}	$\Delta\text{IQR} = 0.27$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.17$ $p < 10^{-6}$	$\mu = 0.09$ $p < 10^{-6}$	$\Delta\mu = 0.01$ $p = 0.3$(NS)	$\Delta\mu = 0.04$ $p = 0.02$	$\Delta\mu = 0.05$ $p = 6 \times 10^{-4}$
Loop_{FR}	$\Delta\text{IQR} = 0.22$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.12$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.05$ $p = 3 \times 10^{-5}$	$\mu = 0.08$ $p < 10^{-6}$	$\Delta\mu = 0.03$ $p = 0.02$	$\Delta\mu = 0.04$ $p = 10^{-6}$
β_{CDR}	$\Delta\text{IQR} = 0.31$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.21$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.04$ $p = 0.02$	$\Delta\text{IQR} = 0.09$ $p < 10^{-6}$	$\mu = 0.04$ $p = 2 \times 10^{-4}$	$\Delta\mu = 0.01$ $p = 0.3$(NS)
β_{FR}	$\Delta\text{IQR} = 0.36$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.25$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.08$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.13$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.05$ $p = 6 \times 10^{-4}$	$\mu = 0.03$ $p < 10^{-6}$

E

Paratope membership

	Non-paratope	Paratope
Non-paratope	$\mu = 0.08$ $p < 10^{-6}$	$\Delta\mu = 0.15$ $p < 10^{-6}$
Paratope	$\Delta\text{IQR} = 0.20$ $p < 10^{-6}$	$\mu = 0.23$ $p < 10^{-6}$

F

Epitope membership

	Non-epitope	Epitope
Non-epitope	$\mu = 1.8 \times 10^{-3}$ $p = 0.5$(NS)	$\Delta\mu = 0.66$ $p < 10^{-6}$
Epitope	$\Delta\text{IQR} = 0.47$ $p < 10^{-6}$	$\mu = 0.67$ $p < 10^{-6}$

Table S3: Summary of statistical tests on the $\varepsilon_{S_{\text{block}}}$ distribution across different scenarios. Panels and cells are the same as those in Table S2, while data is presented in Figure S3. Statistical tests are carried out by the bootstrap as described in Methods, Section 4.10.

A Antibody positions' entropy

	Low	High
Low	$\mu = 0.36$	$\Delta\mu = 0.16$ $p < 10^{-6}$
High	$\Delta\text{IQR} = 0.17$ $p < 10^{-6}$	$\mu = 0.52$

B Antibody FR *vs* CDR

	FR	CDR
FR	$\mu = 0.37$	$\Delta\mu = 0.29$ $p < 10^{-6}$
CDR	$\Delta\text{IQR} = 0.29$ $p < 10^{-6}$	$\mu = 0.66$

C Antigen secondary structure

	α	Loop	β
α	$\mu = 0.93$	$\Delta\mu = 0.56$ $p < 10^{-6}$	$\Delta\mu = 0.10$ $p = 0.14$(NS)
Loop	$\Delta\text{IQR} = 0.66$ $p < 10^{-6}$	$\mu = 1.49$	$\Delta\mu = 0.66$ $p < 10^{-6}$
β	$\Delta\text{IQR} = 0.22$ $p = 0.05$	$\Delta\text{IQR} = 0.88$ $p < 10^{-6}$	$\mu = 0.83$

D Antibody secondary structure by region

	α_{CDR}	Loop_{CDR}	α_{FR}	Loop_{FR}	β_{CDR}	β_{FR}
α_{CDR}	$\mu = 0.60$	$\Delta\mu = 0.21$ $p = 10^{-6}$	$\Delta\mu = 0.20$ $p = 2 \times 10^{-3}$	$\Delta\mu = 0.08$ $p = 0.08$(NS)	$\Delta\mu = 0.27$ $p = 7 \times 10^{-5}$	$\Delta\mu = 0.34$ $p = 2 \times 10^{-6}$
Loop_{CDR}	$\Delta\text{IQR} = 0.05$ $p = 0.14$(NS)	$\mu = 0.81$	$\Delta\mu = 0.40$ $p < 10^{-6}$	$\Delta\mu = 0.28$ $p < 10^{-6}$	$\Delta\mu = 0.48$ $p < 10^{-6}$	$\Delta\mu = 0.55$ $p < 10^{-6}$
α_{FR}	$\Delta\text{IQR} = 0.42$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.36$ $p < 10^{-6}$	$\mu = 0.41$	$\Delta\mu = 0.12$ $p = 5 \times 10^{-5}$	$\Delta\mu = 0.07$ $p = 0.05$	$\Delta\mu = 0.15$ $p = 2 \times 10^{-4}$
Loop_{FR}	$\Delta\text{IQR} = 0.27$ $p = 2 \times 10^{-6}$	$\Delta\text{IQR} = 0.22$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.15$ $p = 10^{-6}$	$\mu = 0.53$	$\Delta\mu = 0.19$ $p < 10^{-6}$	$\Delta\mu = 0.27$ $p < 10^{-6}$
β_{CDR}	$\Delta\text{IQR} = 0.50$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.45$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.08$ $p = 0.03$	$\Delta\text{IQR} = 0.23$ $p < 10^{-6}$	$\mu = 0.33$	$\Delta\mu = 0.08$ $p = 0.02$
β_{FR}	$\Delta\text{IQR} = 0.57$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.52$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.15$ $p = 10^{-5}$	$\Delta\text{IQR} = 0.30$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.07$ $p = 0.01$	$\mu = 0.26$

E Paratope membership

	Non-paratope	Paratope
Non-paratope	$\mu = 0.38$	$\Delta\mu = 0.14$ $p = 5 \times 10^{-4}$
Paratope	$\Delta\text{IQR} = 0.26$ $p < 10^{-6}$	$\mu = 0.52$

F Epitope membership

	Non-epitope	Epitope
Non-epitope	$\mu = 1.02$	$\Delta\mu = 0.35$ $p = 2 \times 10^{-4}$
Epitope	$\Delta\text{IQR} = 0.71$ $p < 10^{-6}$	$\mu = 1.37$

Table S4: Summary of statistical tests on the $\varepsilon_{\text{INFUSSE}}$ distribution across different scenarios. Panels and cells are the same as those in Table S2, while data is presented in Figure S3. Statistical tests are carried out by the bootstrap as described in Methods, Section 4.10.

A Antibody positions' entropy

	Low	High
Low	$\mu = 0.30$	$\Delta\mu = 0.13$ $p < 10^{-6}$
High	$\Delta\text{IQR} = 0.11$ $p < 10^{-6}$	$\mu = 0.43$

B Antibody FR *vs* CDR

	FR	CDR
FR	$\mu = 0.31$	$\Delta\mu = 0.21$ $p < 10^{-6}$
CDR	$\Delta\text{IQR} = 0.13$ $p < 10^{-6}$	$\mu = 0.53$

C Antigen secondary structure

	α	Loop	β
α	$\mu = 0.74$	$\Delta\mu = 0.64$ $p < 10^{-6}$	$\Delta\mu = 0.05$ $p = 0.30$(NS)
Loop	$\Delta\text{IQR} = 0.64$ $p < 10^{-6}$	$\mu = 1.39$	$\Delta\mu = 0.59$ $p < 10^{-6}$
β	$\Delta\text{IQR} = 0.05$ $p = 0.30$(NS)	$\Delta\text{IQR} = 0.69$ $p < 10^{-6}$	$\mu = 0.80$

D Antibody secondary structure by region

	α_{CDR}	Loop_{CDR}	α_{FR}	Loop_{FR}	β_{CDR}	β_{FR}
α_{CDR}	$\mu = 0.43$	$\Delta\mu = 0.22$ $p < 10^{-6}$	$\Delta\mu = 0.11$ $p = 0.03$	$\Delta\mu = 0.02$ $p = 0.36$(NS)	$\Delta\mu = 0.14$ $p = 0.01$	$\Delta\mu = 0.21$ $p = 10^{-3}$
Loop_{CDR}	$\Delta\text{IQR} = 0.07$ $p = 0.03$	$\mu = 0.65$	$\Delta\mu = 0.33$ $p < 10^{-6}$	$\Delta\mu = 0.20$ $p < 10^{-6}$	$\Delta\mu = 0.36$ $p < 10^{-6}$	$\Delta\mu = 0.43$ $p < 10^{-6}$
α_{FR}	$\Delta\text{IQR} = 0.09$ $p = 0.02$	$\Delta\text{IQR} = 0.16$ $p < 10^{-6}$	$\mu = 0.32$	$\Delta\mu = 0.13$ $p = 3 \times 10^{-6}$	$\Delta\mu = 0.03$ $p = 0.25$(NS)	$\Delta\mu = 0.10$ $p = 5 \times 10^{-3}$
Loop_{FR}	$\Delta\text{IQR} = 5 \times 10^{-3}$ $p = 0.46$(NS)	$\Delta\text{IQR} = 0.06$ $p = 5 \times 10^{-4}$	$\Delta\text{IQR} = 0.09$ $p = 5 \times 10^{-5}$	$\mu = 0.45$	$\Delta\mu = 0.16$ $p < 10^{-6}$	$\Delta\mu = 0.23$ $p < 10^{-6}$
β_{CDR}	$\Delta\text{IQR} = 0.22$ $p = 10^{-6}$	$\Delta\text{IQR} = 0.29$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.13$ $p = 8 \times 10^{-5}$	$\Delta\text{IQR} = 0.22$ $p < 10^{-6}$	$\mu = 0.29$	$\Delta\mu = 0.07$ $p = 0.02$
β_{FR}	$\Delta\text{IQR} = 0.24$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.31$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.15$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.24$ $p < 10^{-6}$	$\Delta\text{IQR} = 0.02$ $p = 0.20$(NS)	$\mu = 0.22$

E Paratope membership

	Non-paratope	Paratope
Non-paratope	$\mu = 0.31$	$\Delta\mu = 0.01$ $p = 0.34$(NS)
Paratope	$\Delta\text{IQR} = 0.04$ $p = 0.05$	$\mu = 0.29$

F Epitope membership

	Non-epitope	Epitope
Non-epitope	$\mu = 1.02$	$\Delta\mu = 0.32$ $p = 7 \times 10^{-5}$
Epitope	$\Delta\text{IQR} = 0.22$ $p = 5 \times 10^{-3}$	$\mu = 0.70$