



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 20, 2026 – 09:15 AM UTC

PDB ID : 8GPU / pdb_00008gpu
Title : YFV_E_YD6Fab_prefusion
Authors : Li, Y.; Wu, L.; Qi, J.; Yan, J.; Gao, G.F.
Deposited on : 2022-08-27
Resolution : 2.79 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

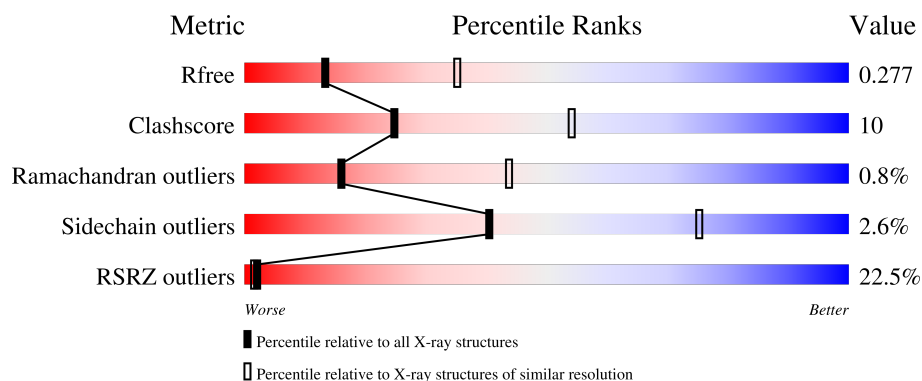
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.79 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	3866 (2.80-2.80)
Clashscore	190562	4276 (2.80-2.80)
Ramachandran outliers	187476	4196 (2.80-2.80)
Sidechain outliers	187428	4198 (2.80-2.80)
RSRZ outliers	180081	3869 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	398	<div> <div>11%</div> <div>78%</div> <div>20%</div> <div>..</div> </div>
1	B	398	<div> <div>11%</div> <div>71%</div> <div>24%</div> <div>..</div> </div>
1	E	398	<div> <div>17%</div> <div>72%</div> <div>24%</div> <div>..</div> </div>
1	I	398	<div> <div>14%</div> <div>79%</div> <div>16%</div> <div>...</div> </div>
1	M	398	<div> <div>28%</div> <div>75%</div> <div>21%</div> <div>...</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	P	398	
2	C	217	
2	F	217	
2	H	217	
2	J	217	
2	N	217	
2	Q	217	
3	D	217	
3	G	217	
3	K	217	
3	L	217	
3	O	217	
3	R	217	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 37132 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Envelope protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	392	Total	C	N	O	S	0	1	0
			2997	1877	518	582	20			
1	E	392	Total	C	N	O	S	0	0	0
			2985	1868	515	582	20			
1	B	388	Total	C	N	O	S	0	1	0
			2965	1861	512	572	20			
1	I	388	Total	C	N	O	S	0	0	0
			2952	1853	508	571	20			
1	M	388	Total	C	N	O	S	0	0	0
			2950	1851	508	571	20			
1	P	388	Total	C	N	O	S	0	0	0
			2950	1851	508	571	20			

- Molecule 2 is a protein called YD6Fab_H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	216	Total	C	N	O	S	0	0	0
			1623	1028	270	318	7			
2	F	216	Total	C	N	O	S	0	0	0
			1623	1028	270	318	7			
2	C	216	Total	C	N	O	S	0	0	0
			1623	1028	270	318	7			
2	J	211	Total	C	N	O	S	0	0	0
			1591	1010	264	310	7			
2	N	216	Total	C	N	O	S	0	0	0
			1623	1028	270	318	7			
2	Q	216	Total	C	N	O	S	0	0	0
			1623	1028	270	318	7			

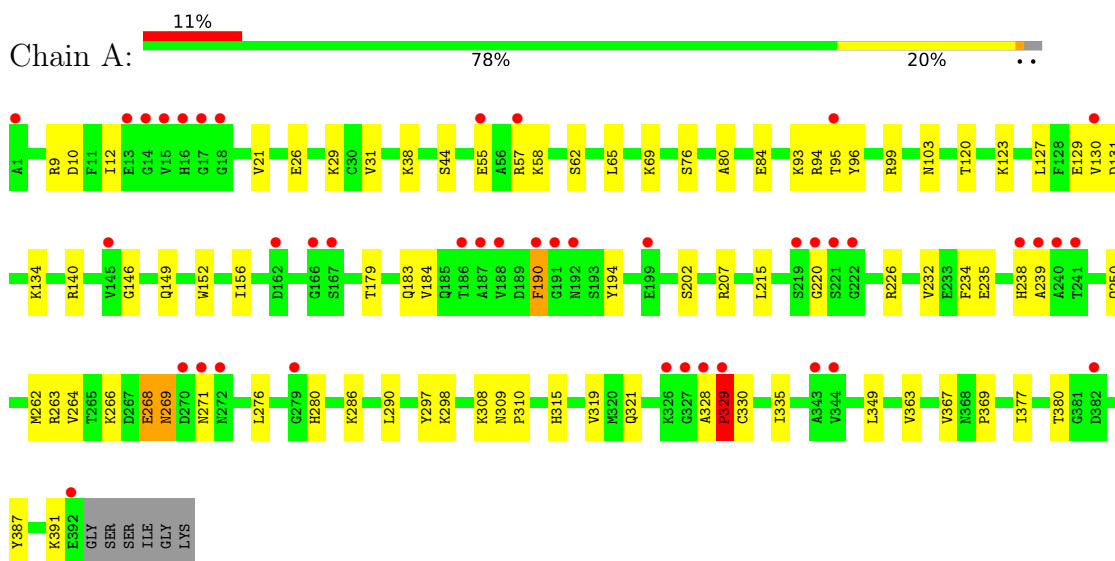
- Molecule 3 is a protein called YD6Fab_L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	214	Total	C	N	O	S	0	0	0
			1601	1004	267	325	5			
3	G	216	Total	C	N	O	S	0	0	0
			1614	1010	268	330	6			
3	D	213	Total	C	N	O	S	0	0	0
			1594	1000	266	323	5			
3	K	214	Total	C	N	O	S	0	0	0
			1601	1004	267	325	5			
3	O	217	Total	C	N	O	S	0	0	0
			1623	1015	270	332	6			
3	R	213	Total	C	N	O	S	0	0	0
			1594	1000	266	323	5			

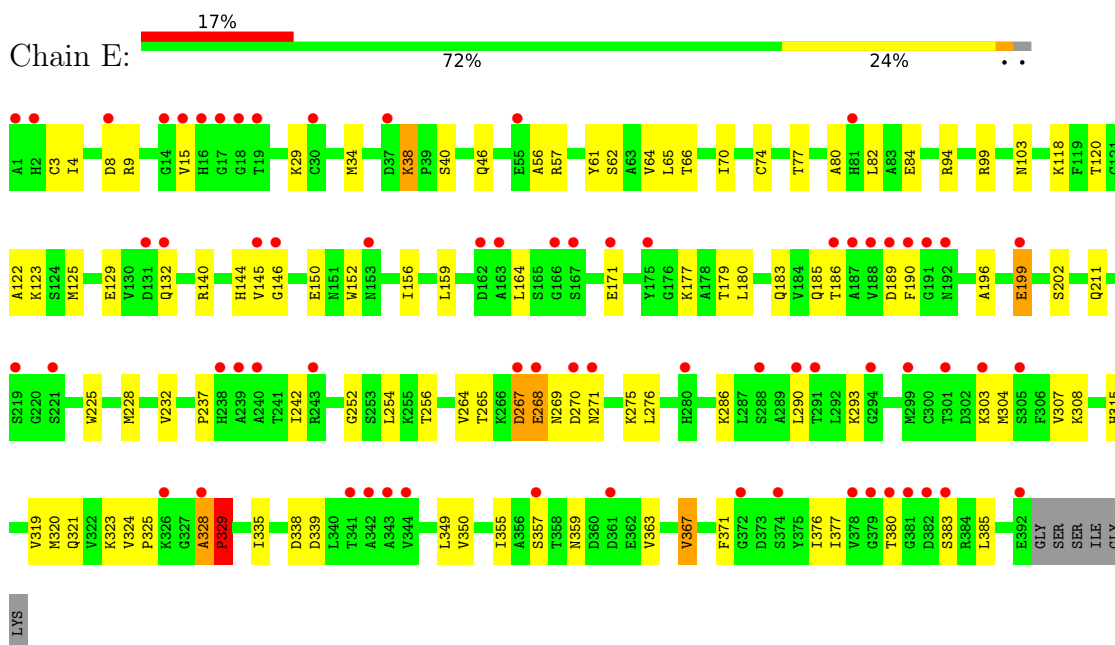
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

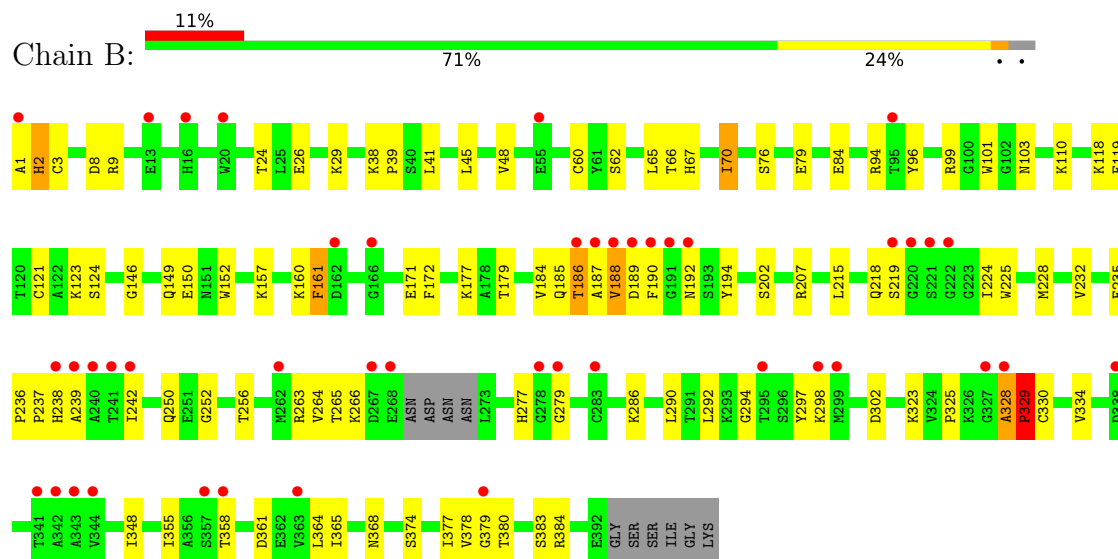
• Molecule 1: Envelope protein



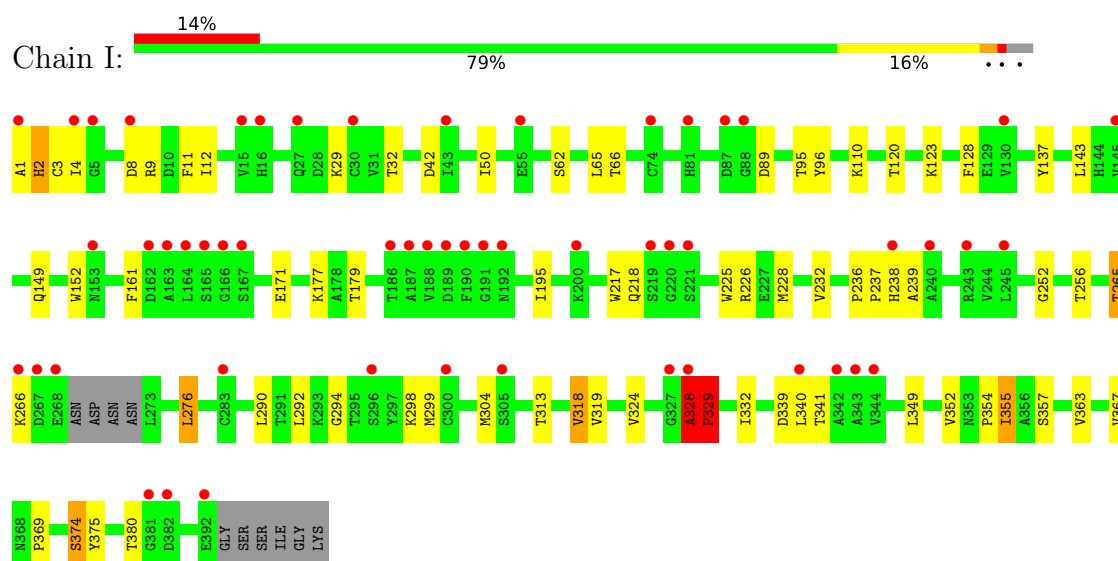
• Molecule 1: Envelope protein



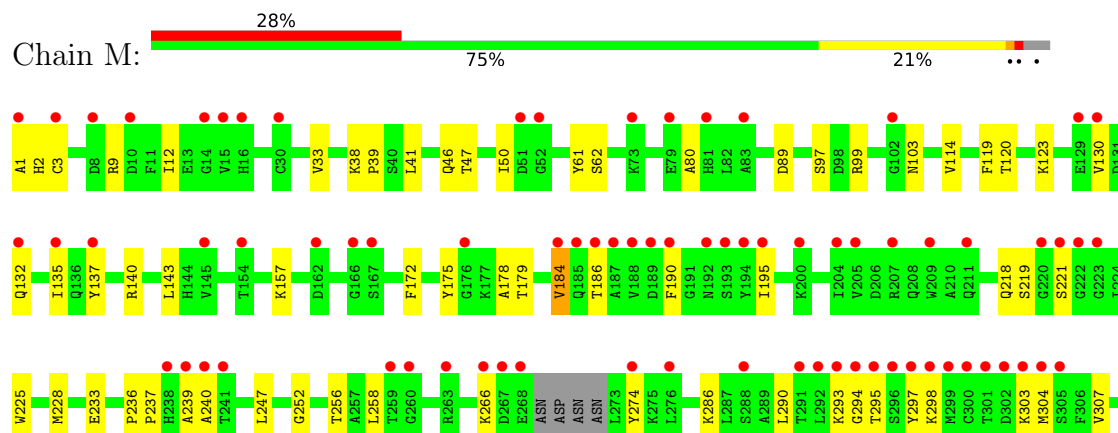
- Molecule 1: Envelope protein

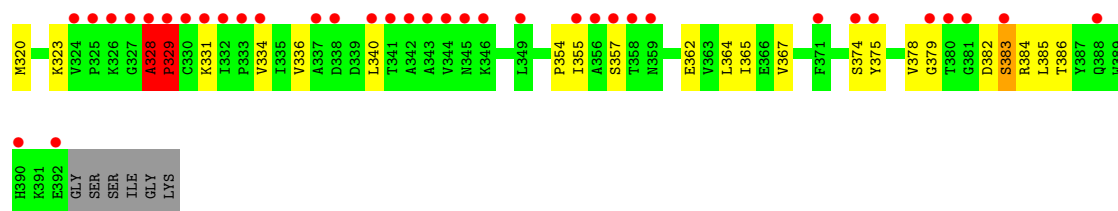


- Molecule 1: Envelope protein

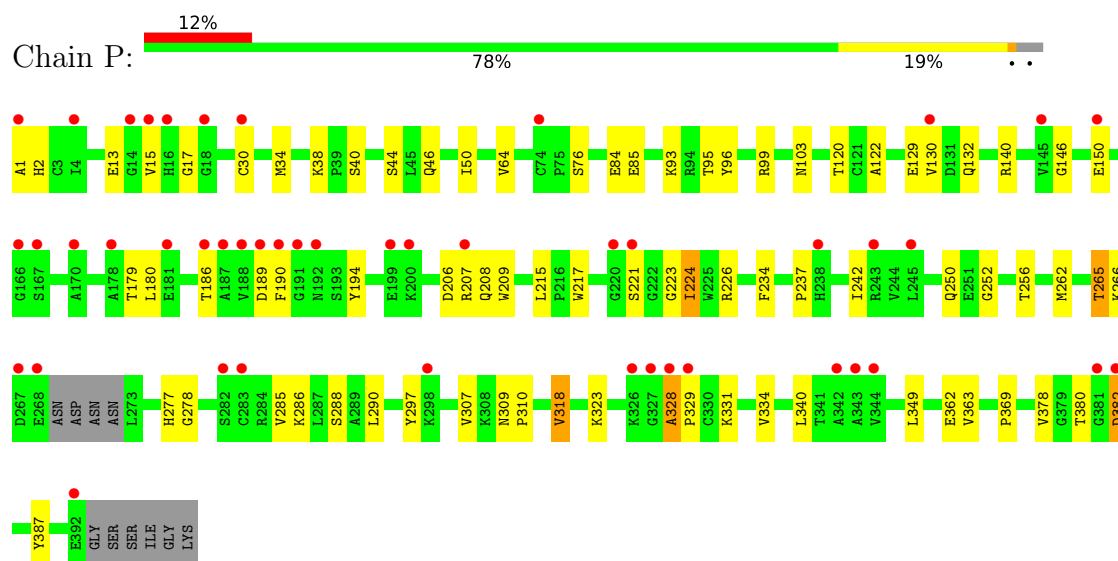


- Molecule 1: Envelope protein

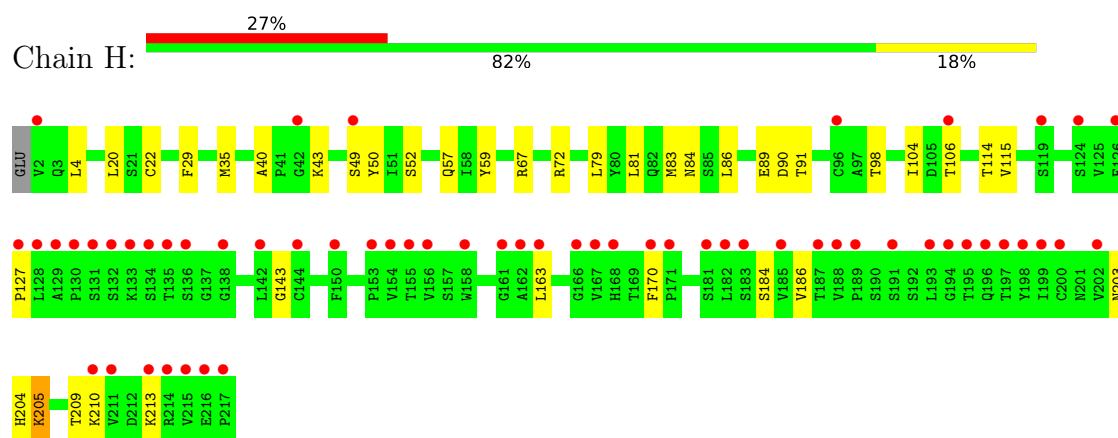




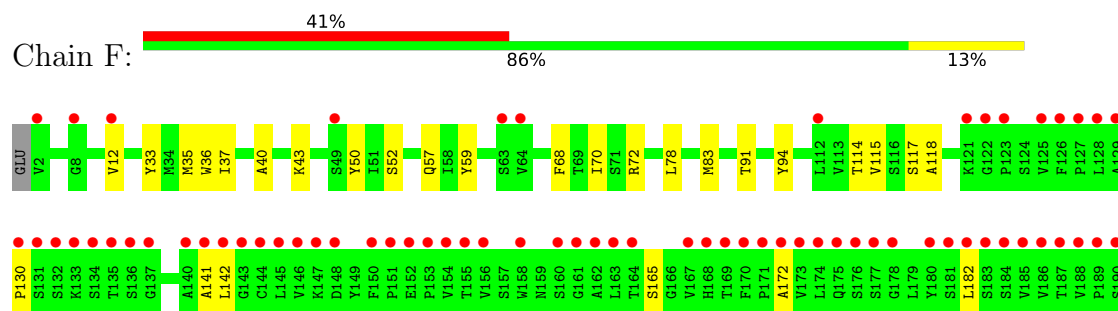
• Molecule 1: Envelope protein

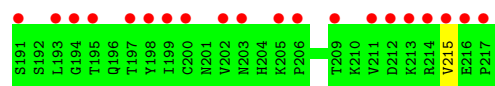


• Molecule 2: YD6Fab_H

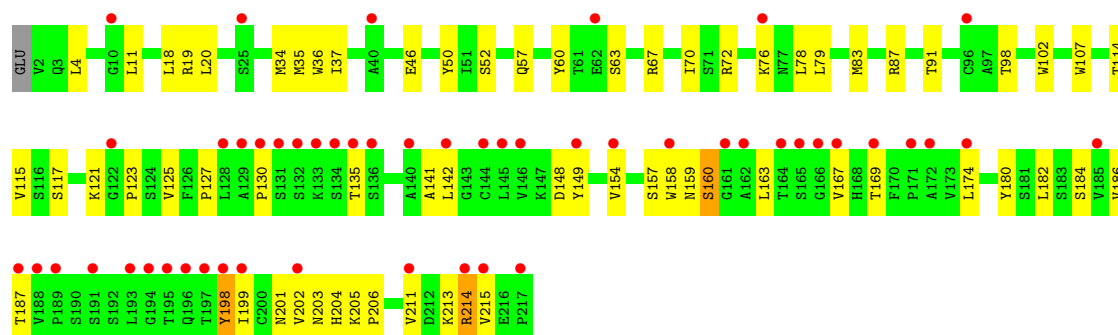


• Molecule 2: YD6Fab_H

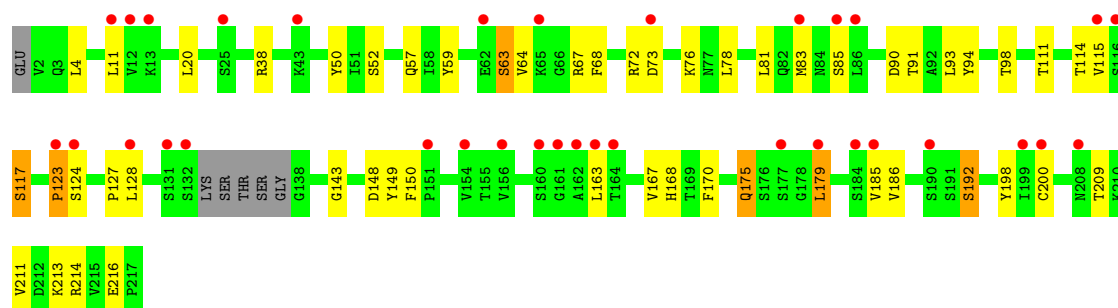




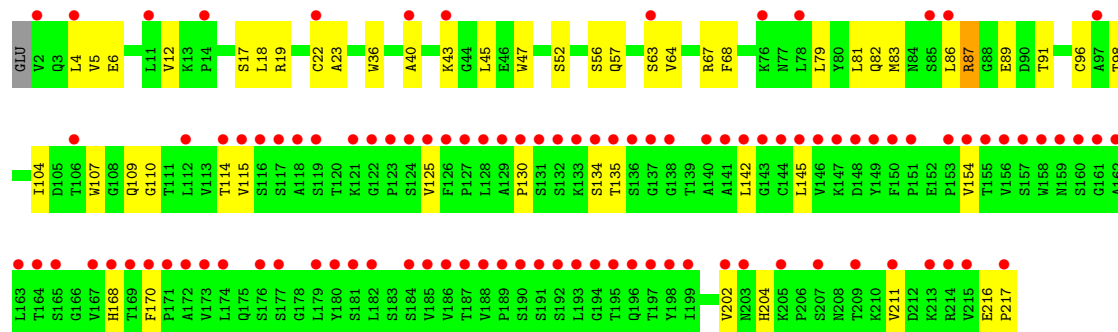
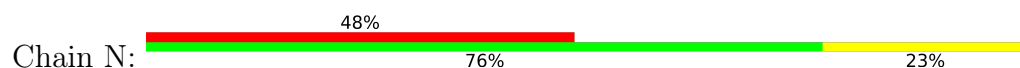
• Molecule 2: YD6Fab_H



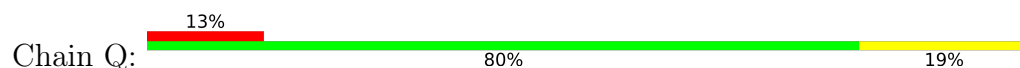
• Molecule 2: YD6Fab_H

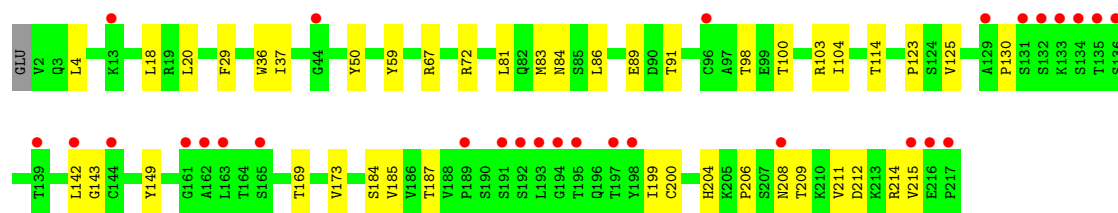


• Molecule 2: YD6Fab_H

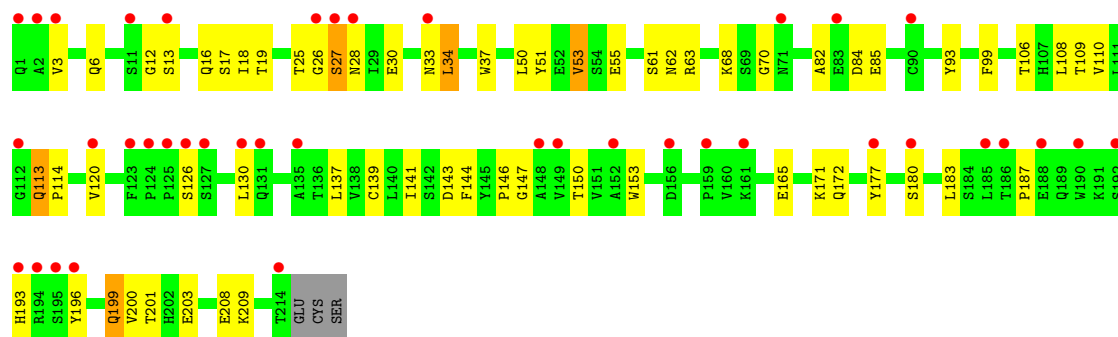


• Molecule 2: YD6Fab_H

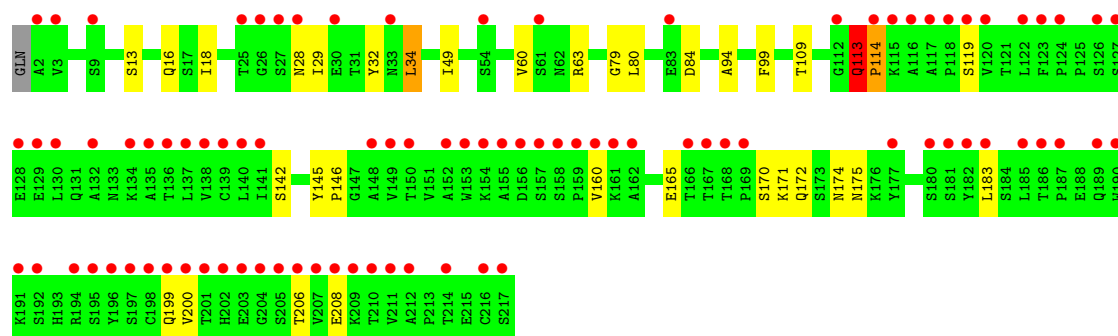
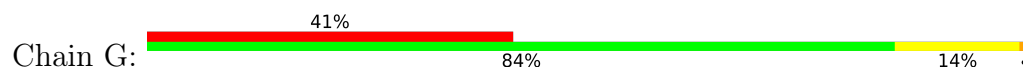




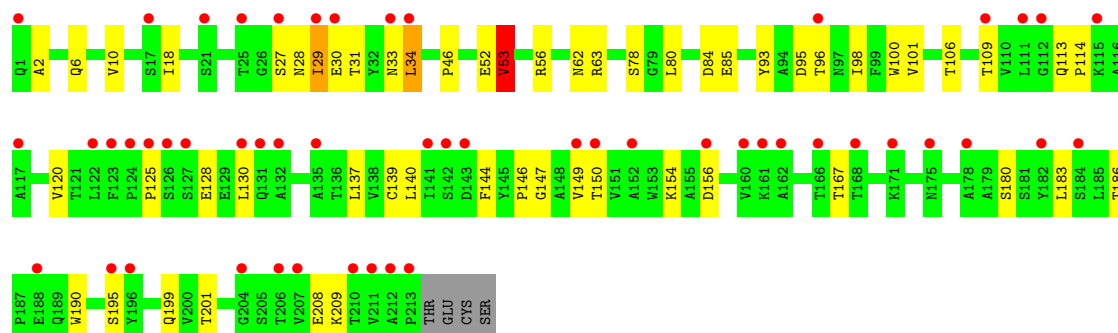
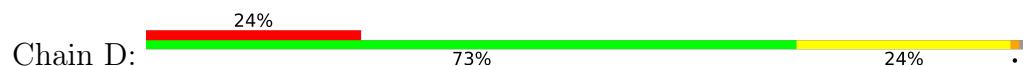
● Molecule 3: YD6Fab_L



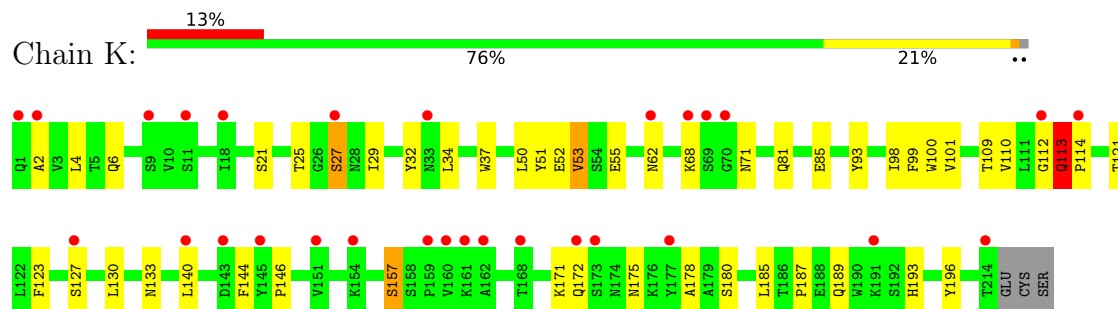
● Molecule 3: YD6Fab_L



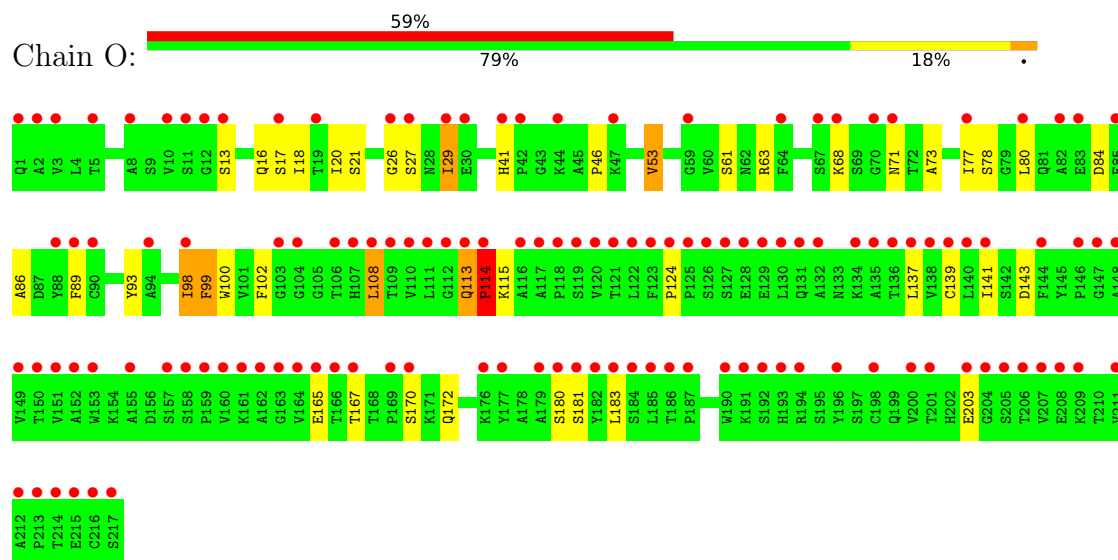
● Molecule 3: YD6Fab_L



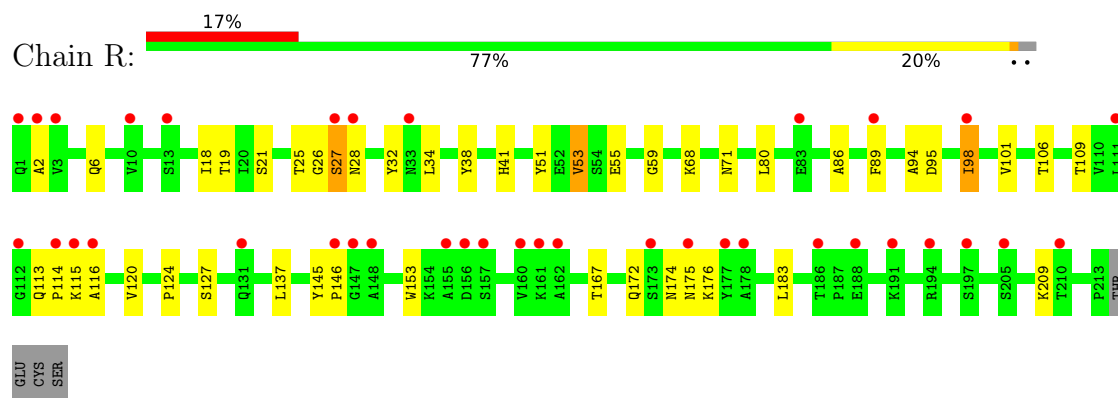
- Molecule 3: YD6Fab_L



- Molecule 3: YD6Fab_L



- Molecule 3: YD6Fab_L



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	157.26Å 278.00Å 354.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.57 – 2.79 42.57 – 2.79	Depositor EDS
% Data completeness (in resolution range)	85.4 (42.57-2.79) 85.3 (42.57-2.79)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.03 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.244 , 0.279 0.243 , 0.277	Depositor DCC
R_{free} test set	8337 reflections (4.36%)	wwPDB-VP
Wilson B-factor (Å ²)	49.9	Xtriage
Anisotropy	0.138	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 43.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.006 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.006 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	37132	wwPDB-VP
Average B, all atoms (Å ²)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.78% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.17	0/3061	0.38	0/4153
1	B	0.18	0/3025	0.40	0/4102
1	E	0.20	0/3045	0.46	4/4132 (0.1%)
1	I	0.16	0/3011	0.37	3/4084 (0.1%)
1	M	0.21	0/3009	0.48	5/4081 (0.1%)
1	P	0.18	0/3009	0.43	1/4081 (0.0%)
2	C	0.24	0/1662	0.50	2/2262 (0.1%)
2	F	0.20	0/1662	0.44	2/2262 (0.1%)
2	H	0.18	0/1662	0.40	0/2262
2	J	0.20	0/1629	0.46	1/2217 (0.0%)
2	N	0.19	0/1662	0.40	0/2262
2	Q	0.19	0/1662	0.35	0/2262
3	D	0.28	0/1634	0.56	0/2232
3	G	0.27	0/1654	0.54	3/2258 (0.1%)
3	K	0.28	0/1641	0.55	2/2242 (0.1%)
3	L	0.32	0/1641	0.58	0/2242
3	O	0.32	0/1663	0.61	1/2270 (0.0%)
3	R	0.31	0/1634	0.52	0/2232
All	All	0.22	0/37966	0.46	24/51636 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1
1	I	0	1
3	K	0	1
All	All	0	3

There are no bond length outliers.

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	P	382	ASP	N-CA-C	-12.33	98.24	113.28
1	M	383	SER	N-CA-C	-8.24	102.78	112.92
3	O	114	PRO	N-CA-CB	-8.14	94.70	103.25
2	F	118	ALA	CA-C-N	6.82	134.23	122.64
2	F	118	ALA	C-N-CA	6.82	134.23	122.64

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	328	ALA	Mainchain
1	I	328	ALA	Mainchain
3	K	112	GLY	Mainchain

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2997	0	2934	56	0
1	B	2965	0	2908	81	0
1	E	2985	0	2911	58	0
1	I	2952	0	2893	40	0
1	M	2950	0	2886	53	0
1	P	2950	0	2886	57	0
2	C	1623	0	1595	48	0
2	F	1623	0	1595	17	0
2	H	1623	0	1595	26	0
2	J	1591	0	1561	38	0
2	N	1623	0	1595	39	0
2	Q	1623	0	1595	26	0
3	D	1594	0	1541	38	0
3	G	1614	0	1553	20	0
3	K	1601	0	1548	41	0
3	L	1601	0	1548	39	0
3	O	1623	0	1564	45	0
3	R	1594	0	1541	32	0
All	All	37132	0	36249	716	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 10.

The worst 5 of 716 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:P:328:ALA:CB	1:P:329:PRO:HD2	1.72	1.19
2:C:76:LYS:O	2:C:78:LEU:HD23	1.43	1.18
3:O:113:GLN:CG	3:O:114:PRO:HD3	1.76	1.14
1:P:328:ALA:HB1	1:P:329:PRO:CD	1.79	1.13
3:O:113:GLN:CB	3:O:114:PRO:HD3	1.78	1.12

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	391/398 (98%)	369 (94%)	19 (5%)	3 (1%)	16	44
1	B	385/398 (97%)	363 (94%)	19 (5%)	3 (1%)	16	44
1	E	390/398 (98%)	367 (94%)	19 (5%)	4 (1%)	12	38
1	I	384/398 (96%)	363 (94%)	19 (5%)	2 (0%)	24	55
1	M	384/398 (96%)	365 (95%)	17 (4%)	2 (0%)	24	55
1	P	384/398 (96%)	364 (95%)	19 (5%)	1 (0%)	36	66
2	C	214/217 (99%)	203 (95%)	10 (5%)	1 (0%)	24	55
2	F	214/217 (99%)	208 (97%)	6 (3%)	0	100	100
2	H	214/217 (99%)	205 (96%)	9 (4%)	0	100	100
2	J	207/217 (95%)	197 (95%)	8 (4%)	2 (1%)	12	38
2	N	214/217 (99%)	202 (94%)	12 (6%)	0	100	100
2	Q	214/217 (99%)	205 (96%)	9 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	D	211/217 (97%)	199 (94%)	9 (4%)	3 (1%)	9	30
3	G	214/217 (99%)	197 (92%)	15 (7%)	2 (1%)	14	41
3	K	212/217 (98%)	194 (92%)	16 (8%)	2 (1%)	14	41
3	L	212/217 (98%)	198 (93%)	10 (5%)	4 (2%)	6	22
3	O	215/217 (99%)	201 (94%)	8 (4%)	6 (3%)	4	14
3	R	211/217 (97%)	196 (93%)	13 (6%)	2 (1%)	14	41
All	All	4870/4992 (98%)	4596 (94%)	237 (5%)	37 (1%)	16	44

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	329	PRO
3	L	26	GLY
1	E	267	ASP
1	E	268	GLU
1	E	329	PRO

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	327/330 (99%)	321 (98%)	6 (2%)	51	82
1	B	322/330 (98%)	309 (96%)	13 (4%)	28	63
1	E	324/330 (98%)	312 (96%)	12 (4%)	30	65
1	I	320/330 (97%)	311 (97%)	9 (3%)	38	73
1	M	319/330 (97%)	315 (99%)	4 (1%)	61	86
1	P	319/330 (97%)	310 (97%)	9 (3%)	38	73
2	C	182/183 (100%)	178 (98%)	4 (2%)	45	78
2	F	182/183 (100%)	179 (98%)	3 (2%)	55	83
2	H	182/183 (100%)	178 (98%)	4 (2%)	45	78
2	J	178/183 (97%)	171 (96%)	7 (4%)	28	64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	N	182/183 (100%)	180 (99%)	2 (1%)	65	87
2	Q	182/183 (100%)	181 (100%)	1 (0%)	81	93
3	D	179/183 (98%)	175 (98%)	4 (2%)	45	78
3	G	182/183 (100%)	176 (97%)	6 (3%)	33	69
3	K	180/183 (98%)	176 (98%)	4 (2%)	45	78
3	L	180/183 (98%)	171 (95%)	9 (5%)	22	54
3	O	183/183 (100%)	176 (96%)	7 (4%)	29	64
3	R	179/183 (98%)	176 (98%)	3 (2%)	53	83
All	All	4102/4176 (98%)	3995 (97%)	107 (3%)	40	75

5 of 107 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	34	LEU
2	J	115	VAL
1	P	288	SER
3	D	149	VAL
1	I	318	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
2	N	84	ASN
3	R	41	HIS
2	N	175	GLN
1	P	230	HIS
1	B	149	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	392/398 (98%)	0.74	42 (10%) 11 8	19, 42, 93, 181	1 (0%)
1	B	388/398 (97%)	0.83	44 (11%) 10 7	21, 47, 99, 144	1 (0%)
1	E	392/398 (98%)	1.00	68 (17%) 4 3	24, 51, 118, 163	0
1	I	388/398 (97%)	0.85	54 (13%) 6 5	21, 46, 95, 186	0
1	M	388/398 (97%)	1.49	111 (28%) 1 1	32, 60, 131, 235	0
1	P	388/398 (97%)	0.85	46 (11%) 9 7	18, 47, 98, 175	0
2	C	216/217 (99%)	1.27	51 (23%) 2 1	21, 67, 159, 238	0
2	F	216/217 (99%)	1.90	88 (40%) 0 0	26, 96, 210, 253	0
2	H	216/217 (99%)	1.32	59 (27%) 1 1	20, 74, 168, 234	0
2	J	211/217 (97%)	1.06	34 (16%) 4 4	17, 66, 100, 162	0
2	N	216/217 (99%)	2.21	104 (48%) 0 0	38, 113, 227, 268	0
2	Q	216/217 (99%)	0.65	29 (13%) 7 5	9, 44, 137, 215	0
3	D	213/217 (98%)	1.31	52 (24%) 2 1	22, 73, 124, 173	0
3	G	216/217 (99%)	1.72	89 (41%) 0 0	30, 87, 177, 237	0
3	K	214/217 (98%)	1.05	29 (13%) 7 5	21, 66, 114, 153	0
3	L	214/217 (98%)	1.16	40 (18%) 3 2	22, 72, 126, 186	0
3	O	217/217 (100%)	2.68	128 (58%) 0 0	46, 140, 231, 292	0
3	R	213/217 (98%)	1.10	37 (17%) 4 3	21, 67, 115, 146	0
All	All	4914/4992 (98%)	1.22	1105 (22%) 2 1	9, 60, 164, 292	2 (0%)

The worst 5 of 1105 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	238[A]	HIS	8.9
3	O	123	PHE	8.3
2	F	127	PRO	8.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	D	123	PHE	7.6
1	P	187	ALA	7.5

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.