



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 28, 2024 – 01:24 PM EDT

PDB ID : 8ULK
Title : Prefusion RSV F bound by neutralizing antibody 1G12
Authors : Harshbarger, W.D.; Andreano, E.; Malito, E.
Deposited on : 2023-10-16
Resolution : 4.28 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	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

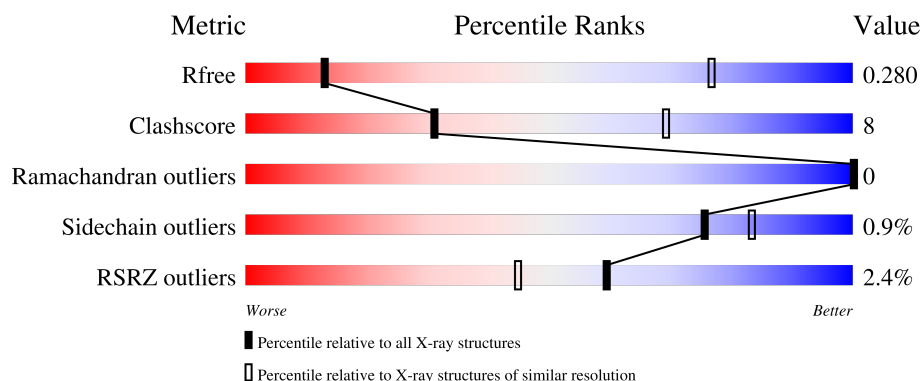
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.28 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}	164625	1024 (4.70-3.86)
Clashscore	180529	1016 (4.66-3.90)
Ramachandran outliers	177936	1010 (4.72-3.84)
Sidechain outliers	177891	1019 (4.72-3.82)
RSRZ outliers	164620	1022 (4.70-3.86)

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	72	<div> <div>3%</div> <div> <div></div> <div>78%</div> <div>22%</div> </div> </div>
1	B	72	<div> <div>%</div> <div> <div></div> <div>79%</div> <div>21%</div> </div> </div>
1	C	72	<div> <div>%</div> <div> <div></div> <div>74%</div> <div>26%</div> </div> </div>
2	D	235	<div> <div>%</div> <div> <div></div> <div>80%</div> <div>11%</div> <div>9%</div> </div> </div>
2	F	235	<div> <div>2%</div> <div> <div></div> <div>81%</div> <div>10%</div> <div>9%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	235	<div><div></div><div>%</div><div><div></div><div>80%</div><div>11%</div><div>9%</div></div></div>
3	E	214	<div><div></div><div>7%</div><div><div></div><div>78%</div><div>22%</div></div></div>
3	G	214	<div><div></div><div>3%</div><div><div></div><div>81%</div><div>19%</div></div></div>
3	L	214	<div><div></div><div>5%</div><div><div></div><div>90%</div><div>10%</div></div></div>
4	I	414	<div><div></div><div>2%</div><div><div></div><div>65%</div><div>26%</div><div>9%</div></div></div>
4	J	414	<div><div></div><div>2%</div><div><div></div><div>70%</div><div>21%</div><div>9%</div></div></div>
4	K	414	<div><div></div><div></div><div><div></div><div>66%</div><div>25%</div><div>9%</div></div></div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 20187 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 Fusion glycoprotein F2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	72	Total	C	N	O	S	0	0	0
			571	361	93	114	3			
1	B	72	Total	C	N	O	S	0	0	0
			571	361	93	114	3			
1	C	72	Total	C	N	O	S	0	0	0
			571	361	93	114	3			

- Molecule 2 is a protein called 1G12 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	214	Total	C	N	O	S	0	0	0
			1606	1014	270	315	7			
2	H	214	Total	C	N	O	S	0	0	0
			1606	1014	270	315	7			
2	F	214	Total	C	N	O	S	0	0	0
			1606	1014	270	315	7			

- Molecule 3 is a protein called 1G12 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	214	Total	C	N	O	S	0	0	0
			1641	1036	281	318	6			
3	L	214	Total	C	N	O	S	0	0	0
			1641	1036	281	318	6			
3	G	214	Total	C	N	O	S	0	0	0
			1641	1036	281	318	6			

- Molecule 4 is a protein called Fusion glycoprotein F0,Fibritin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	I	377	Total	C	N	O	S	0	0	0
			2911	1841	480	570	20			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	J	377	Total	C	N	O	S	0	0	0
			2911	1841	480	570	20			
4	K	377	Total	C	N	O	S	0	0	0
			2911	1841	480	570	20			

There are 45 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	155	CYS	SER	conflict	UNP A0A088S9A7
I	190	PHE	SER	conflict	UNP A0A088S9A7
I	207	LEU	VAL	conflict	UNP A0A088S9A7
I	290	CYS	SER	conflict	UNP A0A088S9A7
I	514	SER	-	linker	UNP A0A088S9A7
I	515	ALA	-	linker	UNP A0A088S9A7
I	516	ILE	-	linker	UNP A0A088S9A7
I	517	GLY	-	linker	UNP A0A088S9A7
I	539	LEU	PHE	conflict	UNP P10104
I	545	GLY	-	expression tag	UNP P10104
I	546	GLY	-	expression tag	UNP P10104
I	547	LEU	-	expression tag	UNP P10104
I	548	VAL	-	expression tag	UNP P10104
I	549	PRO	-	expression tag	UNP P10104
I	550	ARG	-	expression tag	UNP P10104
J	155	CYS	SER	conflict	UNP A0A088S9A7
J	190	PHE	SER	conflict	UNP A0A088S9A7
J	207	LEU	VAL	conflict	UNP A0A088S9A7
J	290	CYS	SER	conflict	UNP A0A088S9A7
J	514	SER	-	linker	UNP A0A088S9A7
J	515	ALA	-	linker	UNP A0A088S9A7
J	516	ILE	-	linker	UNP A0A088S9A7
J	517	GLY	-	linker	UNP A0A088S9A7
J	539	LEU	PHE	conflict	UNP P10104
J	545	GLY	-	expression tag	UNP P10104
J	546	GLY	-	expression tag	UNP P10104
J	547	LEU	-	expression tag	UNP P10104
J	548	VAL	-	expression tag	UNP P10104
J	549	PRO	-	expression tag	UNP P10104
J	550	ARG	-	expression tag	UNP P10104
K	155	CYS	SER	conflict	UNP A0A088S9A7
K	190	PHE	SER	conflict	UNP A0A088S9A7
K	207	LEU	VAL	conflict	UNP A0A088S9A7
K	290	CYS	SER	conflict	UNP A0A088S9A7

Continued on next page...

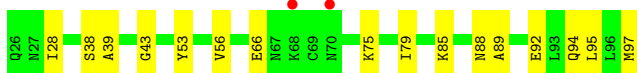
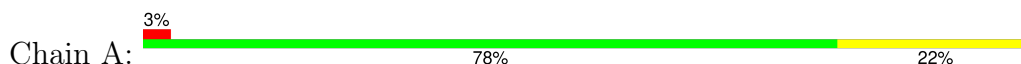
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
K	514	SER	-	linker	UNP A0A088S9A7
K	515	ALA	-	linker	UNP A0A088S9A7
K	516	ILE	-	linker	UNP A0A088S9A7
K	517	GLY	-	linker	UNP A0A088S9A7
K	539	LEU	PHE	conflict	UNP P10104
K	545	GLY	-	expression tag	UNP P10104
K	546	GLY	-	expression tag	UNP P10104
K	547	LEU	-	expression tag	UNP P10104
K	548	VAL	-	expression tag	UNP P10104
K	549	PRO	-	expression tag	UNP P10104
K	550	ARG	-	expression tag	UNP P10104

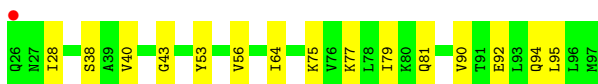
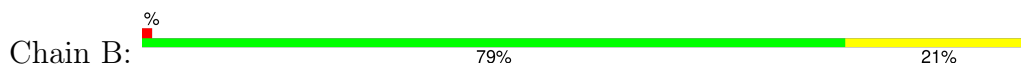
3 Residue-property plots [i](#)

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.

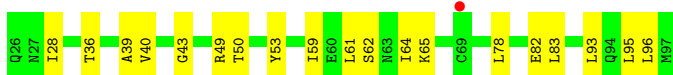
- Molecule 1: Fusion glycoprotein F2



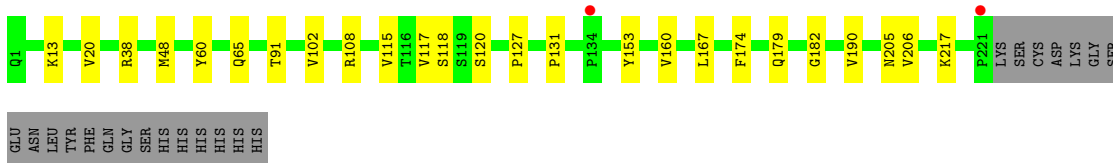
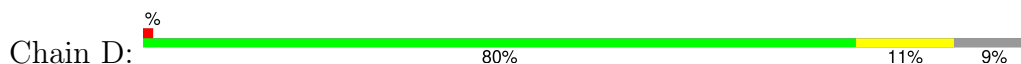
- Molecule 1: Fusion glycoprotein F2



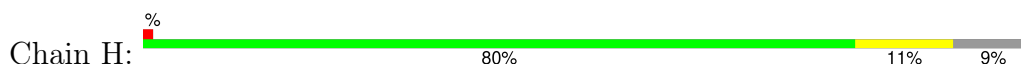
- Molecule 1: Fusion glycoprotein F2

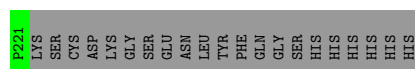


- Molecule 2: 1G12 Fab heavy chain

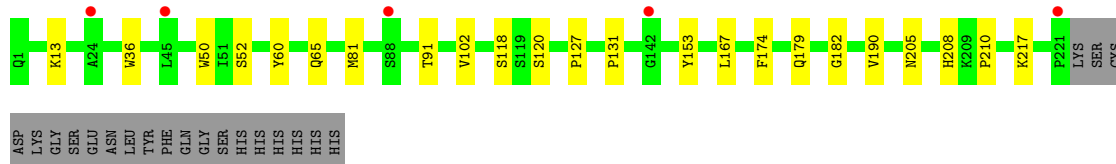
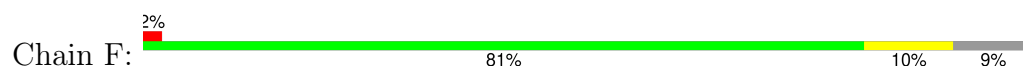


- Molecule 2: 1G12 Fab heavy chain

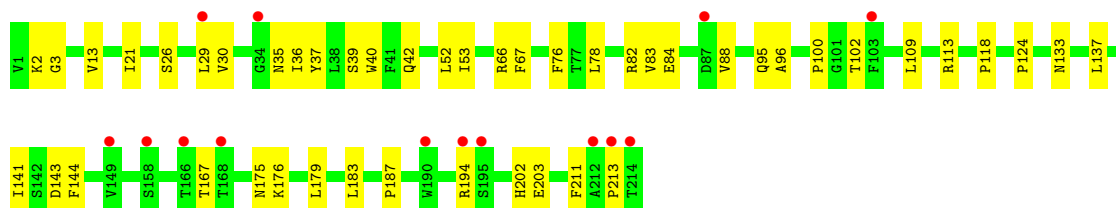
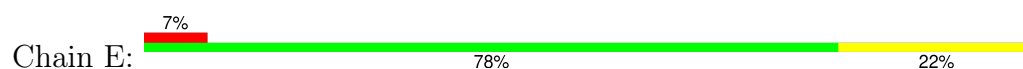




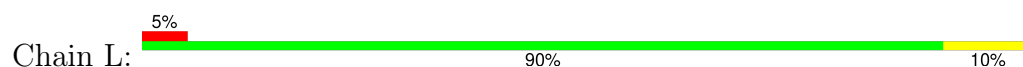
- Molecule 2: 1G12 Fab heavy chain



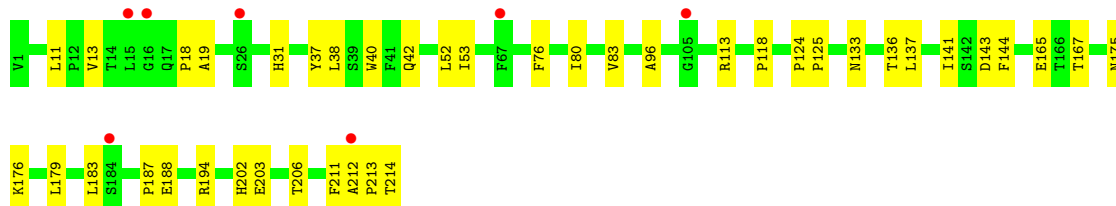
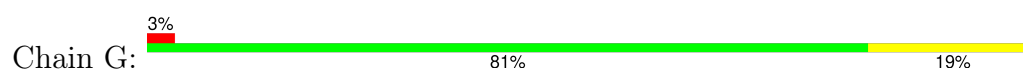
- Molecule 3: 1G12 Fab light chain



- Molecule 3: 1G12 Fab light chain

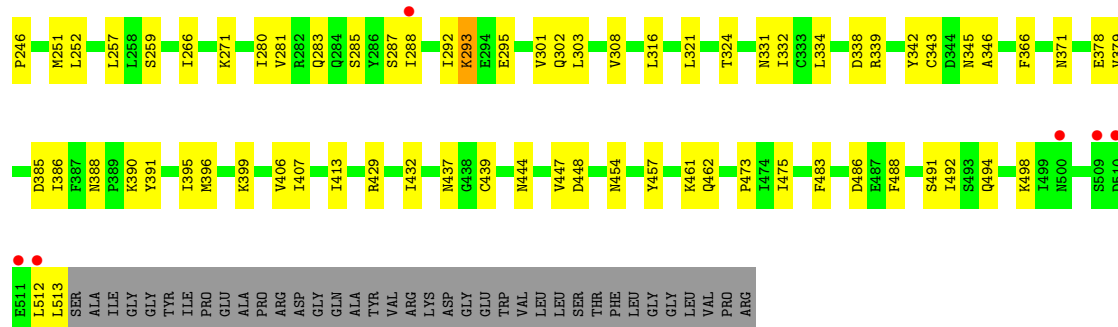


- Molecule 3: 1G12 Fab light chain

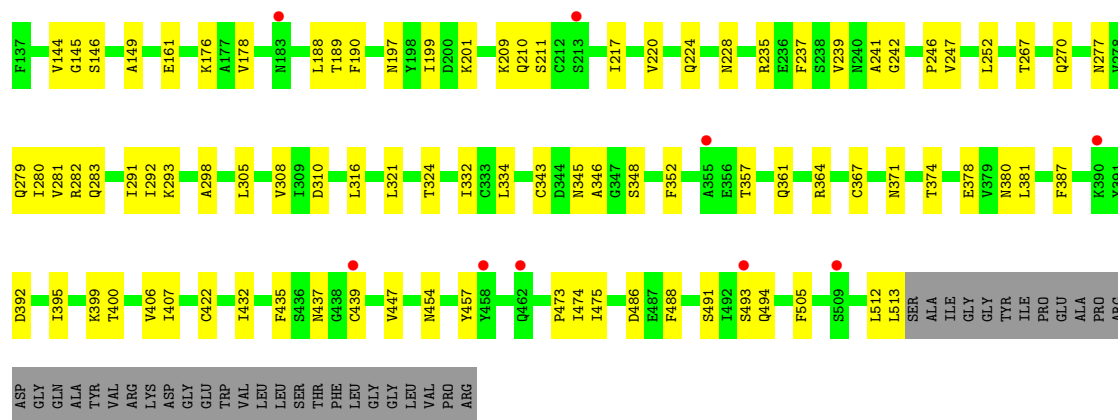


- Molecule 4: Fusion glycoprotein F0,Fibritin

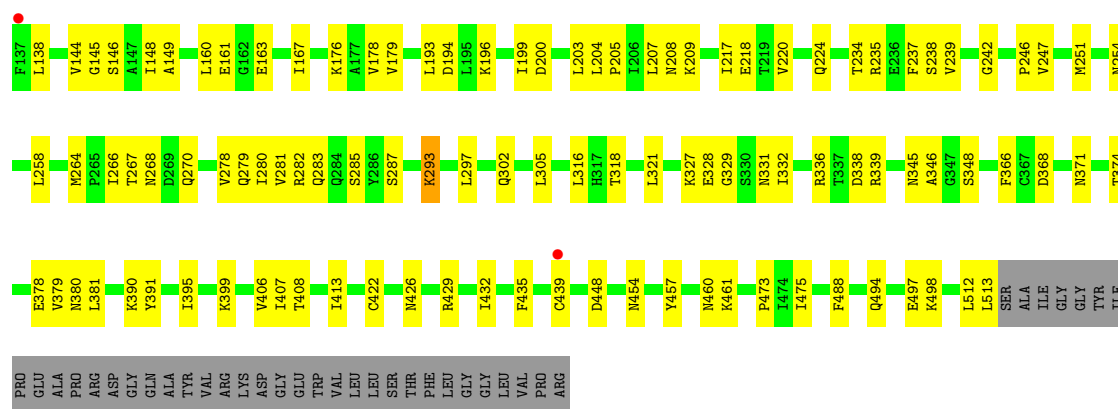




• Molecule 4: Fusion glycoprotein F0,Fibrin



• Molecule 4: Fusion glycoprotein F0,Fibrin



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	143.98Å 174.60Å 226.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.95 – 4.28 46.95 – 4.28	Depositor EDS
% Data completeness (in resolution range)	72.8 (46.95-4.28) 67.6 (46.95-4.28)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.67 (at 4.29Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???), PHENIX 1.20.1.4487	Depositor
R, R_{free}	0.244 , 0.279 0.245 , 0.280	Depositor DCC
R_{free} test set	38004 reflections (6.77%)	wwPDB-VP
Wilson B-factor (Å ²)	60.3	Xtriage
Anisotropy	0.717	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 128.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.79	EDS
Total number of atoms	20187	wwPDB-VP
Average B, all atoms (Å ²)	108.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.70% 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	0.24	0/577	0.48	0/777
1	B	0.24	0/577	0.49	0/777
1	C	0.24	0/577	0.51	0/777
2	D	0.25	0/1645	0.51	0/2245
2	F	0.25	0/1645	0.51	0/2245
2	H	0.25	0/1645	0.51	0/2245
3	E	0.26	0/1685	0.51	0/2294
3	G	0.25	0/1685	0.50	0/2294
3	L	0.25	0/1685	0.51	0/2294
4	I	0.24	0/2955	0.46	0/4007
4	J	0.25	0/2955	0.47	0/4007
4	K	0.25	0/2955	0.49	0/4007
All	All	0.25	0/20586	0.49	0/27969

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
4	I	0	1
4	J	0	1
4	K	0	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	I	293	LYS	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
4	J	293	LYS	Peptide
4	K	293	LYS	Peptide

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	571	0	581	13	0
1	B	571	0	581	15	0
1	C	571	0	581	17	0
2	D	1606	0	1564	15	0
2	F	1606	0	1564	13	0
2	H	1606	0	1564	15	0
3	E	1641	0	1603	36	0
3	G	1641	0	1603	26	0
3	L	1641	0	1603	14	0
4	I	2911	0	2945	85	0
4	J	2911	0	2945	71	0
4	K	2911	0	2945	88	0
All	All	20187	0	20079	326	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 8.

The worst 5 of 326 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)
4:K:246:PRO:HB3	4:K:283:GLN:HA	1.54	0.87
4:I:246:PRO:HB3	4:I:283:GLN:HA	1.58	0.84
1:B:95:LEU:HD11	4:K:278:VAL:HG23	1.62	0.82
4:J:374:THR:HG21	4:K:454:ASN:H	1.43	0.81
3:E:42:GLN:HB2	3:E:52:LEU:HD11	1.63	0.80

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	70/72 (97%)	65 (93%)	5 (7%)	0	100	100
1	B	70/72 (97%)	65 (93%)	5 (7%)	0	100	100
1	C	70/72 (97%)	64 (91%)	6 (9%)	0	100	100
2	D	210/235 (89%)	206 (98%)	4 (2%)	0	100	100
2	F	210/235 (89%)	206 (98%)	4 (2%)	0	100	100
2	H	210/235 (89%)	206 (98%)	4 (2%)	0	100	100
3	E	212/214 (99%)	202 (95%)	10 (5%)	0	100	100
3	G	212/214 (99%)	202 (95%)	10 (5%)	0	100	100
3	L	212/214 (99%)	203 (96%)	9 (4%)	0	100	100
4	I	375/414 (91%)	352 (94%)	23 (6%)	0	100	100
4	J	375/414 (91%)	353 (94%)	22 (6%)	0	100	100
4	K	375/414 (91%)	351 (94%)	24 (6%)	0	100	100
All	All	2601/2805 (93%)	2475 (95%)	126 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	65/65 (100%)	63 (97%)	2 (3%)	35	56
1	B	65/65 (100%)	65 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	65/65 (100%)	64 (98%)	1 (2%)	60	75
2	D	178/197 (90%)	177 (99%)	1 (1%)	84	88
2	F	178/197 (90%)	177 (99%)	1 (1%)	84	88
2	H	178/197 (90%)	178 (100%)	0	100	100
3	E	186/186 (100%)	185 (100%)	1 (0%)	86	90
3	G	186/186 (100%)	185 (100%)	1 (0%)	86	90
3	L	186/186 (100%)	185 (100%)	1 (0%)	86	90
4	I	345/373 (92%)	342 (99%)	3 (1%)	75	83
4	J	345/373 (92%)	341 (99%)	4 (1%)	67	79
4	K	345/373 (92%)	340 (99%)	5 (1%)	62	76
All	All	2322/2463 (94%)	2302 (99%)	20 (1%)	75	83

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	J	505	PHE
4	K	268	ASN
4	K	439	CYS
4	K	380	ASN
2	F	205	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 ⓘ

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	D	1
2	F	1
2	H	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	134:PRO	C	142:GLY	N	12.20
1	F	134:PRO	C	142:GLY	N	12.12
1	H	134:PRO	C	142:GLY	N	11.95

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	72/72 (100%)	0.48	2 (2%) 55 41	68, 104, 168, 206	0
1	B	72/72 (100%)	0.46	1 (1%) 73 58	60, 98, 163, 200	0
1	C	72/72 (100%)	0.49	1 (1%) 73 58	54, 86, 155, 168	0
2	D	214/235 (91%)	0.44	2 (0%) 81 66	55, 96, 138, 165	0
2	F	214/235 (91%)	0.46	5 (2%) 61 45	71, 106, 140, 158	0
2	H	214/235 (91%)	0.29	3 (1%) 73 58	47, 97, 144, 160	0
3	E	214/214 (100%)	0.86	14 (6%) 26 24	66, 120, 171, 194	0
3	G	214/214 (100%)	0.70	7 (3%) 49 38	86, 132, 192, 207	0
3	L	214/214 (100%)	0.53	10 (4%) 37 31	55, 118, 173, 191	0
4	I	377/414 (91%)	0.46	7 (1%) 66 50	58, 108, 160, 215	0
4	J	377/414 (91%)	0.44	9 (2%) 59 45	56, 108, 164, 220	0
4	K	377/414 (91%)	0.27	2 (0%) 87 76	47, 88, 150, 209	0
All	All	2631/2805 (93%)	0.47	63 (2%) 59 45	47, 107, 163, 220	0

The worst 5 of 63 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	26	GLN	5.0
3	E	166	THR	3.8
4	J	213	SER	3.6
3	E	195	SER	3.6
3	E	214	THR	3.5

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 monosaccharides 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.