wwPDB X-ray Structure Validation Summary Report (i)

Sep 24, 2023 – 11:08 AM EDT

PDB ID	:	5U68
Title	:	Structural basis for antibody cross-neutralization of respiratory syncytial virus and human metapneumovirus
Authors	:	Wen, X.; Jardetzky, T.S.
Deposited on	:	2016-12-07
Resolution	:	3.08 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 (i) 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 (1)) were used in the production of this report:

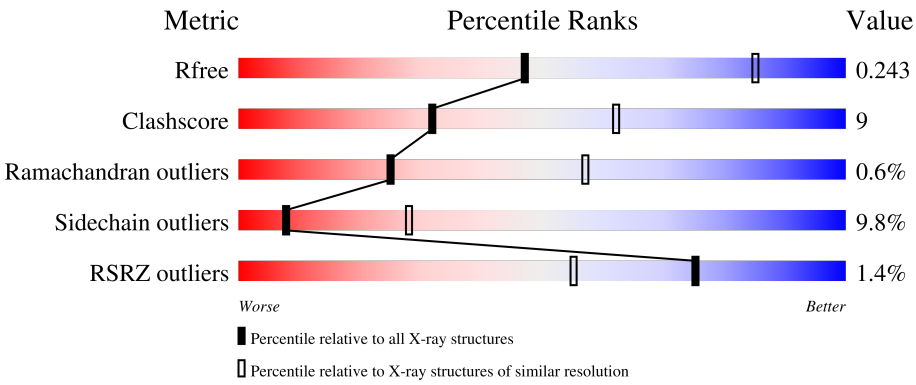
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.08 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}	130704	1447 (3.10-3.06)
Clashscore	141614	1546 (3.10-3.06)
Ramachandran outliers	138981	1487 (3.10-3.06)
Sidechain outliers	138945	1486 (3.10-3.06)
RSRZ outliers	127900	1416 (3.10-3.06)

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 <=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	562	% 53% 23% • 20%
1	B	562	% 55% 22% • 20%
1	C	562	% 59% 18% • 20%
2	E	294	62% 17% • 21%
2	F	294	59% 18% • 21%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	294	4% 65% 13% • 21%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 31169 atoms, of which 15527 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 Chimera protein of Fusion glycoprotein F0 and Envelope glycoprotein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	448	6971	2191	3503	573	681	23	0	0	0
1	B	448	6966	2191	3498	573	681	23	0	0	0
1	C	448	6972	2191	3504	573	681	23	0	0	0

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	ALA	PRO	conflict	UNP P03420
A	155	CYS	SER	conflict	UNP P03420
A	190	PHE	SER	conflict	UNP P03420
A	207	LEU	VAL	conflict	UNP P03420
A	290	CYS	SER	conflict	UNP P03420
A	379	VAL	ILE	conflict	UNP P03420
A	447	VAL	MET	conflict	UNP P03420
A	514	SER	-	linker	UNP P03420
A	515	ALA	-	linker	UNP P03420
A	516	ILE	-	linker	UNP P03420
A	517	GLY	-	linker	UNP P03420
A	546	GLY	-	expression tag	UNP M1E1E4
A	547	LEU	-	expression tag	UNP M1E1E4
A	548	VAL	-	expression tag	UNP M1E1E4
A	549	PRO	-	expression tag	UNP M1E1E4
A	550	ARG	-	expression tag	UNP M1E1E4
A	551	GLY	-	expression tag	UNP M1E1E4
A	552	SER	-	expression tag	UNP M1E1E4
A	553	SER	-	expression tag	UNP M1E1E4
A	554	ALA	-	expression tag	UNP M1E1E4
A	555	HIS	-	expression tag	UNP M1E1E4
A	556	HIS	-	expression tag	UNP M1E1E4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	557	HIS	-	expression tag	UNP M1E1E4
A	558	HIS	-	expression tag	UNP M1E1E4
A	559	HIS	-	expression tag	UNP M1E1E4
A	560	HIS	-	expression tag	UNP M1E1E4
A	561	HIS	-	expression tag	UNP M1E1E4
A	562	HIS	-	expression tag	UNP M1E1E4
B	102	ALA	PRO	conflict	UNP P03420
B	155	CYS	SER	conflict	UNP P03420
B	190	PHE	SER	conflict	UNP P03420
B	207	LEU	VAL	conflict	UNP P03420
B	290	CYS	SER	conflict	UNP P03420
B	379	VAL	ILE	conflict	UNP P03420
B	447	VAL	MET	conflict	UNP P03420
B	514	SER	-	linker	UNP P03420
B	515	ALA	-	linker	UNP P03420
B	516	ILE	-	linker	UNP P03420
B	517	GLY	-	linker	UNP P03420
B	546	GLY	-	expression tag	UNP M1E1E4
B	547	LEU	-	expression tag	UNP M1E1E4
B	548	VAL	-	expression tag	UNP M1E1E4
B	549	PRO	-	expression tag	UNP M1E1E4
B	550	ARG	-	expression tag	UNP M1E1E4
B	551	GLY	-	expression tag	UNP M1E1E4
B	552	SER	-	expression tag	UNP M1E1E4
B	553	SER	-	expression tag	UNP M1E1E4
B	554	ALA	-	expression tag	UNP M1E1E4
B	555	HIS	-	expression tag	UNP M1E1E4
B	556	HIS	-	expression tag	UNP M1E1E4
B	557	HIS	-	expression tag	UNP M1E1E4
B	558	HIS	-	expression tag	UNP M1E1E4
B	559	HIS	-	expression tag	UNP M1E1E4
B	560	HIS	-	expression tag	UNP M1E1E4
B	561	HIS	-	expression tag	UNP M1E1E4
B	562	HIS	-	expression tag	UNP M1E1E4
C	102	ALA	PRO	conflict	UNP P03420
C	155	CYS	SER	conflict	UNP P03420
C	190	PHE	SER	conflict	UNP P03420
C	207	LEU	VAL	conflict	UNP P03420
C	290	CYS	SER	conflict	UNP P03420
C	379	VAL	ILE	conflict	UNP P03420
C	447	VAL	MET	conflict	UNP P03420
C	514	SER	-	linker	UNP P03420

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	515	ALA	-	linker	UNP P03420
C	516	ILE	-	linker	UNP P03420
C	517	GLY	-	linker	UNP P03420
C	546	GLY	-	expression tag	UNP M1E1E4
C	547	LEU	-	expression tag	UNP M1E1E4
C	548	VAL	-	expression tag	UNP M1E1E4
C	549	PRO	-	expression tag	UNP M1E1E4
C	550	ARG	-	expression tag	UNP M1E1E4
C	551	GLY	-	expression tag	UNP M1E1E4
C	552	SER	-	expression tag	UNP M1E1E4
C	553	SER	-	expression tag	UNP M1E1E4
C	554	ALA	-	expression tag	UNP M1E1E4
C	555	HIS	-	expression tag	UNP M1E1E4
C	556	HIS	-	expression tag	UNP M1E1E4
C	557	HIS	-	expression tag	UNP M1E1E4
C	558	HIS	-	expression tag	UNP M1E1E4
C	559	HIS	-	expression tag	UNP M1E1E4
C	560	HIS	-	expression tag	UNP M1E1E4
C	561	HIS	-	expression tag	UNP M1E1E4
C	562	HIS	-	expression tag	UNP M1E1E4

- Molecule 2 is a protein called MPE8.

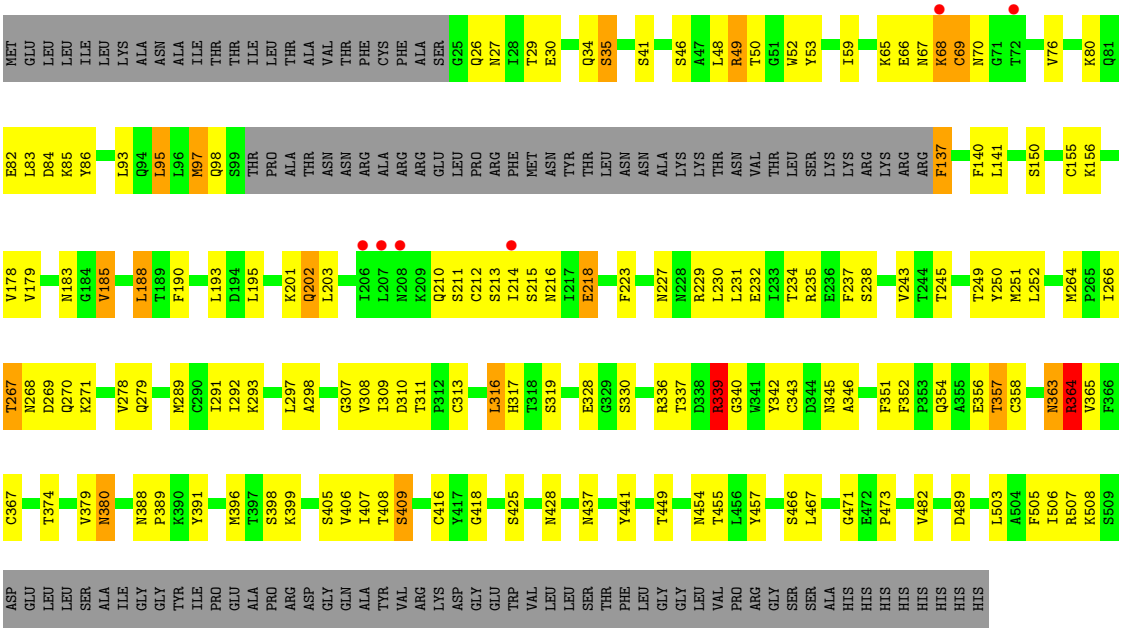
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	E	233	3420	1090	1674	298	352	6	0	0	0
2	F	233	3420	1090	1674	298	352	6	0	0	0
2	G	233	3420	1090	1674	298	352	6	0	0	0

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.

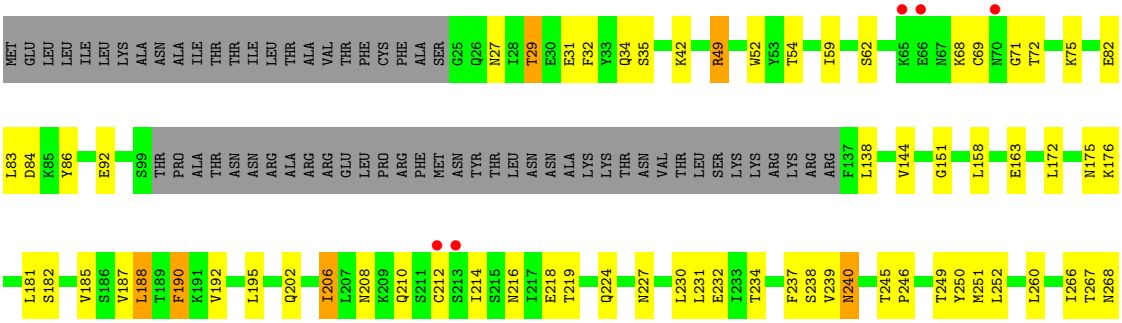
• Molecule 1: Chimera protein of Fusion glycoprotein F0 and Envelope glycoprotein

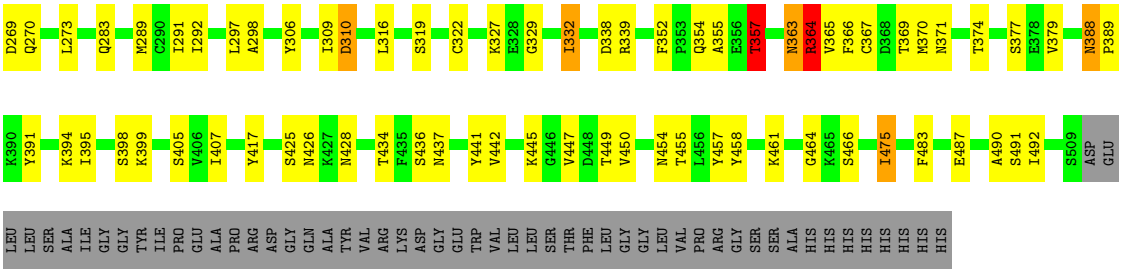
Chain A:



• Molecule 1: Chimera protein of Fusion glycoprotein F0 and Envelope glycoprotein

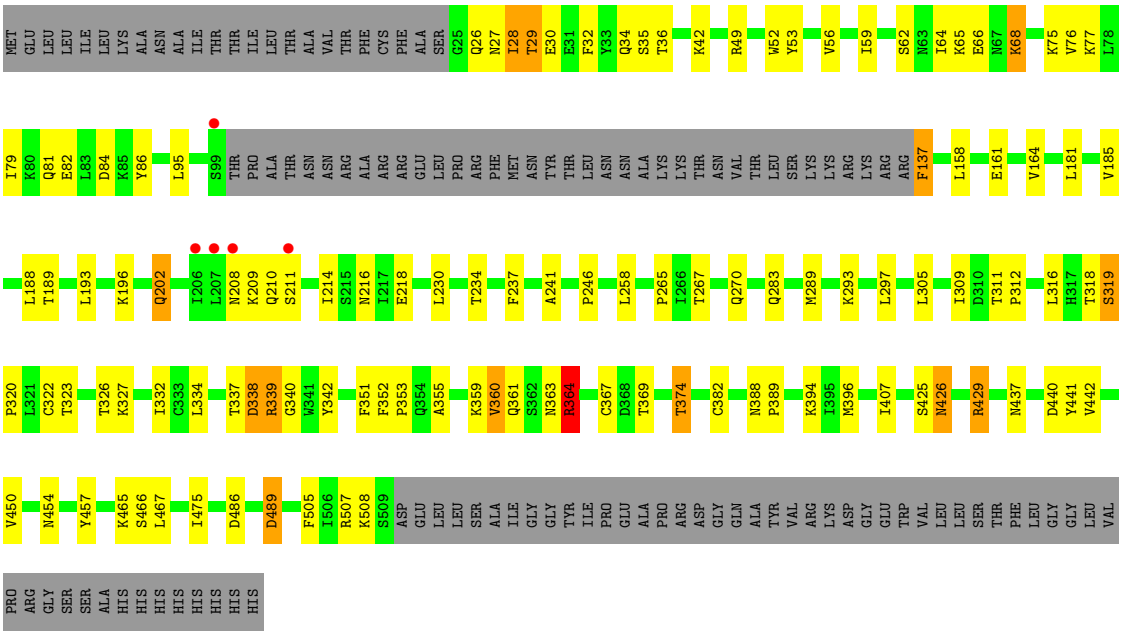
Chain B:





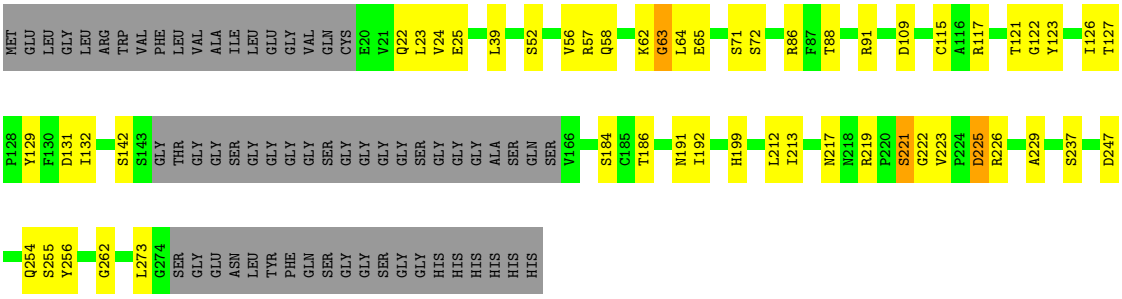
- Molecule 1: Chimera protein of Fusion glycoprotein F0 and Envelope glycoprotein

Chain C:



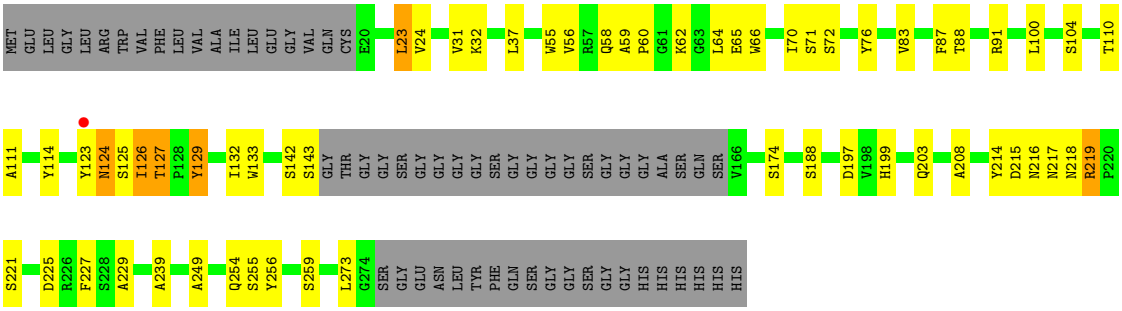
- Molecule 2: MPE8

Chain E:



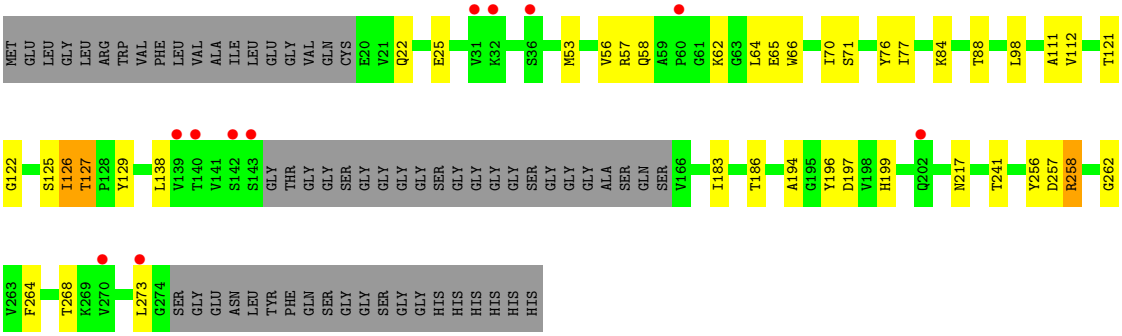
- Molecule 2: MPE8

Chain F:



- Molecule 2: MPE8

Chain G:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	155.54Å 155.54Å 275.93Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.19 – 3.08 48.19 – 3.08	Depositor EDS
% Data completeness (in resolution range)	74.6 (48.19-3.08) 74.6 (48.19-3.08)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.43 (at 3.07Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.184 , 0.239 0.189 , 0.243	Depositor DCC
R_{free} test set	2687 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	67.5	Xtriage
Anisotropy	0.031	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 24.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.032 for -h,-k,l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	31169	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.93% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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	1.01	5/3518 (0.1%)	0.96	7/4764 (0.1%)
1	B	1.13	3/3518 (0.1%)	0.99	5/4764 (0.1%)
1	C	0.88	2/3518 (0.1%)	0.93	2/4764 (0.0%)
2	E	0.83	0/1783	0.87	1/2423 (0.0%)
2	F	1.22	3/1783 (0.2%)	0.98	2/2423 (0.1%)
2	G	0.61	0/1783	0.74	2/2423 (0.1%)
All	All	0.98	13/15903 (0.1%)	0.93	19/21561 (0.1%)

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	A	0	1
1	B	0	2
1	C	0	1
All	All	0	4

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	208	ASN	CB-CG	6.51	1.66	1.51
1	B	31	GLU	CG-CD	5.99	1.60	1.51
2	F	129	TYR	CE2-CZ	-5.95	1.30	1.38
1	A	343	CYS	CB-SG	-5.84	1.72	1.81
2	F	239	ALA	CA-CB	-5.68	1.40	1.52

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	364	ARG	NE-CZ-NH1	9.19	124.90	120.30
1	A	339	ARG	NE-CZ-NH2	-7.83	116.39	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	364	ARG	NE-CZ-NH1	6.54	123.57	120.30
2	F	219	ARG	NE-CZ-NH1	-6.42	117.09	120.30
2	G	122	GLY	N-CA-C	6.40	129.10	113.10

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	67	ASN	Peptide
1	B	355	ALA	Peptide
1	B	72	THR	Peptide
1	C	26	GLN	Peptide

5.2 Too-close contacts (i)

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	3468	3503	3510	87	0
1	B	3468	3498	3510	69	0
1	C	3468	3504	3510	56	0
2	E	1746	1674	1676	30	0
2	F	1746	1674	1676	35	0
2	G	1746	1674	1676	19	0
All	All	15642	15527	15558	276	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 9.

The worst 5 of 276 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:B:374:THR:HG21	1:C:454:ASN:H	1.53	0.73
1:A:266:ILE:HA	2:E:126:ILE:HG23	1.72	0.71
2:E:226:ARG:NH1	2:E:247:ASP:OD2	2.23	0.71
2:F:56:VAL:HG12	2:F:66:TRP:HA	1.71	0.71
1:A:407:ILE:HD11	1:A:457:TYR:HB3	1.74	0.68

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	444/562 (79%)	404 (91%)	35 (8%)	5 (1%)	14	44
1	B	444/562 (79%)	396 (89%)	45 (10%)	3 (1%)	22	54
1	C	444/562 (79%)	404 (91%)	39 (9%)	1 (0%)	47	77
2	E	229/294 (78%)	211 (92%)	16 (7%)	2 (1%)	17	49
2	F	229/294 (78%)	209 (91%)	20 (9%)	0	100	100
2	G	229/294 (78%)	212 (93%)	16 (7%)	1 (0%)	34	66
All	All	2019/2568 (79%)	1836 (91%)	171 (8%)	12 (1%)	25	57

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	27	ASN
1	A	69	CYS
1	A	202	GLN
1	B	357	THR
1	C	202	GLN

5.3.2 Protein sidechains (i)

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	408/504 (81%)	368 (90%)	40 (10%)	8	28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	408/504 (81%)	358 (88%)	50 (12%)	4	18
1	C	408/504 (81%)	361 (88%)	47 (12%)	5	21
2	E	190/228 (83%)	177 (93%)	13 (7%)	16	44
2	F	190/228 (83%)	176 (93%)	14 (7%)	13	41
2	G	190/228 (83%)	179 (94%)	11 (6%)	20	50
All	All	1794/2196 (82%)	1619 (90%)	175 (10%)	8	28

5 of 175 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	326	THR
2	E	132	ILE
1	C	359	LYS
1	C	465	LYS
2	F	23	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	317	HIS
1	B	159	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides 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 (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	448/562 (79%)	-0.29	6 (1%)	77	58	31, 51, 111, 155	0
1	B	448/562 (79%)	-0.33	5 (1%)	80	62	26, 45, 102, 171	0
1	C	448/562 (79%)	-0.25	5 (1%)	80	62	29, 60, 118, 166	0
2	E	233/294 (79%)	-0.20	0	100	100	40, 65, 96, 115	0
2	F	233/294 (79%)	-0.41	1 (0%)	92	84	29, 41, 60, 95	0
2	G	233/294 (79%)	0.29	11 (4%)	31	14	58, 101, 134, 166	0
All	All	2043/2568 (79%)	-0.23	28 (1%)	75	55	26, 55, 116, 171	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	206	ILE	7.0
1	C	208	ASN	6.1
1	A	207	LEU	4.8
1	C	207	LEU	4.7
1	A	208	ASN	4.6

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.