



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 09:47 am BST

PDB ID : 4F15  
Title : Molecular basis of infectivity of 2009 pandemic H1N1 influenza A viruses  
Authors : Kim, K.H.; Cho, K.J.; Lee, J.H.; Park, Y.H.; Khan, T.G.; Lee, J.Y.; Kang, S.H.; Alam, I.  
Deposited on : 2012-05-06  
Resolution : 2.81 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

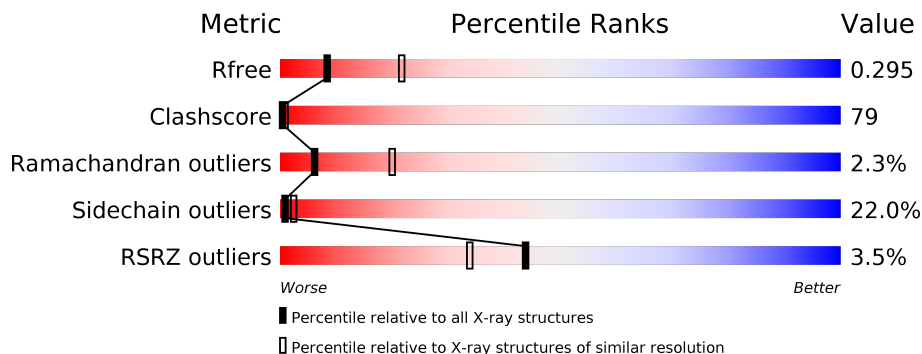
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.81 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	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)
RSRZ outliers	127900	3552 (2.84-2.80)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	518	
1	D	518	
1	G	518	
1	J	518	
2	B	219	
2	E	219	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	219	<p>2% 24% 58% 13% 5%</p>
2	K	219	<p>26% 54% 15% 5%</p>
3	C	218	<p>16% 55% 18% 7%</p>
3	F	218	<p>16% 52% 22% 7%</p>
3	I	218	<p>16% 55% 19% 7%</p>
3	L	218	<p>14% 57% 18% 7%</p>

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 19900 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 Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	234	Total 1813	C 1152	N 311	O 344	S 6	0	0	0
1	D	246	Total 1872	C 1187	N 323	O 356	S 6	0	0	0
1	G	227	Total 1778	C 1131	N 304	O 337	S 6	0	0	0
1	J	255	Total 1918	C 1215	N 332	O 365	S 6	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	ALA	-	EXPRESSION TAG	UNP C5MQE6
A	-7	ASP	-	EXPRESSION TAG	UNP C5MQE6
A	-6	PRO	-	EXPRESSION TAG	UNP C5MQE6
A	-5	GLY	-	EXPRESSION TAG	UNP C5MQE6
A	-4	TYR	-	EXPRESSION TAG	UNP C5MQE6
A	-3	LEU	-	EXPRESSION TAG	UNP C5MQE6
A	-2	LEU	-	EXPRESSION TAG	UNP C5MQE6
A	-1	GLU	-	EXPRESSION TAG	UNP C5MQE6
A	0	PHE	-	EXPRESSION TAG	UNP C5MQE6
A	507	ARG	-	EXPRESSION TAG	UNP C5MQE6
A	508	SER	-	EXPRESSION TAG	UNP C5MQE6
A	509	LEU	-	EXPRESSION TAG	UNP C5MQE6
A	510	VAL	-	EXPRESSION TAG	UNP C5MQE6
A	511	PRO	-	EXPRESSION TAG	UNP C5MQE6
A	512	ARG	-	EXPRESSION TAG	UNP C5MQE6
D	-8	ALA	-	EXPRESSION TAG	UNP C5MQE6
D	-7	ASP	-	EXPRESSION TAG	UNP C5MQE6
D	-6	PRO	-	EXPRESSION TAG	UNP C5MQE6
D	-5	GLY	-	EXPRESSION TAG	UNP C5MQE6
D	-4	TYR	-	EXPRESSION TAG	UNP C5MQE6
D	-3	LEU	-	EXPRESSION TAG	UNP C5MQE6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-2	LEU	-	EXPRESSION TAG	UNP C5MQE6
D	-1	GLU	-	EXPRESSION TAG	UNP C5MQE6
D	0	PHE	-	EXPRESSION TAG	UNP C5MQE6
D	507	ARG	-	EXPRESSION TAG	UNP C5MQE6
D	508	SER	-	EXPRESSION TAG	UNP C5MQE6
D	509	LEU	-	EXPRESSION TAG	UNP C5MQE6
D	510	VAL	-	EXPRESSION TAG	UNP C5MQE6
D	511	PRO	-	EXPRESSION TAG	UNP C5MQE6
D	512	ARG	-	EXPRESSION TAG	UNP C5MQE6
G	-8	ALA	-	EXPRESSION TAG	UNP C5MQE6
G	-7	ASP	-	EXPRESSION TAG	UNP C5MQE6
G	-6	PRO	-	EXPRESSION TAG	UNP C5MQE6
G	-5	GLY	-	EXPRESSION TAG	UNP C5MQE6
G	-4	TYR	-	EXPRESSION TAG	UNP C5MQE6
G	-3	LEU	-	EXPRESSION TAG	UNP C5MQE6
G	-2	LEU	-	EXPRESSION TAG	UNP C5MQE6
G	-1	GLU	-	EXPRESSION TAG	UNP C5MQE6
G	0	PHE	-	EXPRESSION TAG	UNP C5MQE6
G	507	ARG	-	EXPRESSION TAG	UNP C5MQE6
G	508	SER	-	EXPRESSION TAG	UNP C5MQE6
G	509	LEU	-	EXPRESSION TAG	UNP C5MQE6
G	510	VAL	-	EXPRESSION TAG	UNP C5MQE6
G	511	PRO	-	EXPRESSION TAG	UNP C5MQE6
G	512	ARG	-	EXPRESSION TAG	UNP C5MQE6
J	-8	ALA	-	EXPRESSION TAG	UNP C5MQE6
J	-7	ASP	-	EXPRESSION TAG	UNP C5MQE6
J	-6	PRO	-	EXPRESSION TAG	UNP C5MQE6
J	-5	GLY	-	EXPRESSION TAG	UNP C5MQE6
J	-4	TYR	-	EXPRESSION TAG	UNP C5MQE6
J	-3	LEU	-	EXPRESSION TAG	UNP C5MQE6
J	-2	LEU	-	EXPRESSION TAG	UNP C5MQE6
J	-1	GLU	-	EXPRESSION TAG	UNP C5MQE6
J	0	PHE	-	EXPRESSION TAG	UNP C5MQE6
J	507	ARG	-	EXPRESSION TAG	UNP C5MQE6
J	508	SER	-	EXPRESSION TAG	UNP C5MQE6
J	509	LEU	-	EXPRESSION TAG	UNP C5MQE6
J	510	VAL	-	EXPRESSION TAG	UNP C5MQE6
J	511	PRO	-	EXPRESSION TAG	UNP C5MQE6
J	512	ARG	-	EXPRESSION TAG	UNP C5MQE6

- Molecule 2 is a protein called Fab fragment, heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	208	Total	C	N	O	S	0	0	0
			1544	962	268	307	7			
2	E	208	Total	C	N	O	S	0	0	0
			1544	962	268	307	7			
2	H	208	Total	C	N	O	S	0	0	0
			1544	962	268	307	7			
2	K	208	Total	C	N	O	S	0	0	0
			1544	962	268	307	7			

- Molecule 3 is a protein called Fab fragment, light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	203	Total	C	N	O	S	0	0	0
			1557	975	263	313	6			
3	F	203	Total	C	N	O	S	0	0	0
			1557	975	263	313	6			
3	I	203	Total	C	N	O	S	0	0	0
			1557	975	263	313	6			
3	L	203	Total	C	N	O	S	0	0	0
			1557	975	263	313	6			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	11	Total	O	0	0
			11	11		
4	B	8	Total	O	0	0
			8	8		
4	C	11	Total	O	0	0
			11	11		
4	D	7	Total	O	0	0
			7	7		
4	E	11	Total	O	0	0
			11	11		
4	F	9	Total	O	0	0
			9	9		
4	G	8	Total	O	0	0
			8	8		
4	H	10	Total	O	0	0
			10	10		
4	I	12	Total	O	0	0
			12	12		
4	J	7	Total	O	0	0
			7	7		

*Continued on next page...*

*Continued from previous page...*

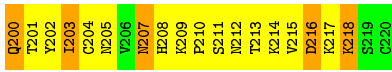
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	K	6	Total 6	O 6	0	0
4	L	15	Total 15	O 15	0	0



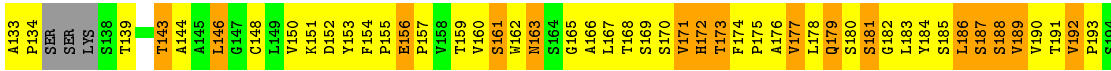
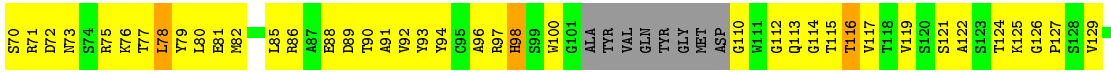




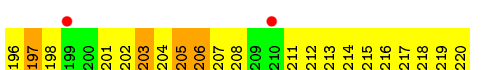
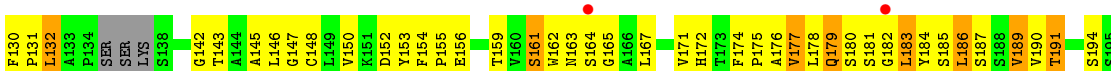
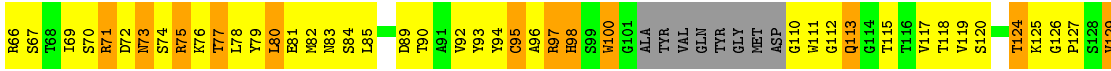
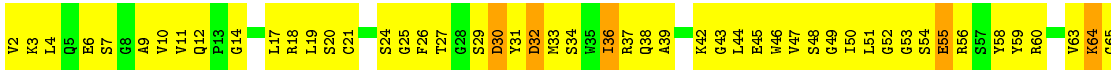




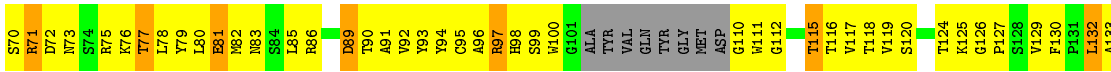
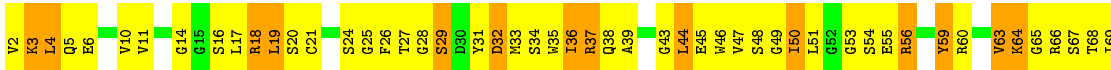
• Molecule 2: Fab fragment, heavy chain



• Molecule 2: Fab fragment, heavy chain



• Molecule 2: Fab fragment, heavy chain



T201  
Y202  
I203  
C204  
N205  
V206  
M207  
H208  
K209  
P210  
S211  
N212  
T213  
K214  
V215  
D216  
K217  
K218  
S219  
C220

- Molecule 3: Fab fragment, light chain

Chain C: 16% 55% 18% 7%

ASP I1 Q2 M3 T4 S6 F7 A8 S9 L10 A11 V12 S13 S14 S15 S16 S17 S18 S19 S20 S21 S22 S23 S24 S25 S26 S27 S28 S29 ASN TVR GLY ILE N34 F35 F36 F37 F38 F39 F40 F41 F42 F43 F44 F45 F46 F47 F48 F49 F50 F51 F52 F53 F54 F55 F56 F57 F58 F59 F60 F61 F62 F63 F64 F65 F66 F67 F68 F69 F70 F71 F72 F73 F74 F75 F76 F77 F78 F79 F80 F81 F82 F83 F84 F85 F86 F87 F88 F89 F90 F91 F92 F93 F94 F95 F96 F97 F98 F99 T100 T101 T102 T103 T104 T105 T106 T107 T108 T109 T110 T111 T112 T113 T114 T115 T116 T117 T118 T119 T120 T121 T122 T123 T124

F65 S66 G67 S68 G69 S70 G71 F72 D73 F74 F75 F76 F77 F78 M79 N79 F80 F81 F82 F83 F84 F85 F86 F87 F88 F89 F90 F91 F92 F93 F94 F95 F96 F97 F98 F99 T100 T101 T102 T103 T104 T105 T106 T107 T108 T109 T110 T111 T112 T113 T114 T115 T116 T117 T118 T119 T120 T121 T122 T123 T124

S125 E126 Q127 L128 M129 S130 G131 G132 A133 S134 S135 V136 V137 F138 L139 M140 M141 F142 F143 F144 K145 K146 I147 I148 V149 K150 W151 K152 K153 I154 D154 G155 S156 S157 R158 F159 M160 M161 M162 L163 S165 S166 S167 T167 D168 Q169 D170 D171 K172 K173 K174 S174 S175 Y176 S177 M178 S179 S180 M181 T182 T183 T184

THR LYS ASP GLU E189 R191 H192 M193 S194 Y195 T196 C197 E198 T203 S204 T205 S206 P207 I208 Y209 K210 S211 F212 N213 ARG ASN GLU CYS

- Molecule 3: Fab fragment, light chain

Chain F: 16% 52% 22% 7%

ASP I1 Q2 M3 T4 S6 F7 A8 S9 L10 A11 V12 S13 S14 S15 S16 S17 S18 S19 S20 S21 S22 S23 S24 S25 S26 S27 S28 S29 ASN TVR GLY ILE N34 F35 F36 F37 F38 F39 F40 F41 F42 F43 F44 F45 F46 F47 F48 F49 F50 F51 F52 F53 F54 F55 F56 F57 F58 F59 F60 F61 F62 F63 F64 F65 F66 F67 F68 F69 F70 F71 F72 F73 F74 F75 F76 F77 F78 F79 F80 F81 F82 F83 F84 F85 F86 F87 F88 F89 F90 F91 F92 F93 F94 F95 F96 F97 F98 F99 T100 T101 T102 T103 T104 T105 T106 T107 T108 T109 T110 T111 T112 T113 T114 T115 T116 T117 T118 T119 T120 T121 T122 T123 T124

R64 F65 S66 G67 S70 G71 F72 L76 F77 M78 F79 F80 F81 F82 F83 F84 F85 F86 F87 F88 F89 F90 F91 F92 F93 F94 F95 F96 F97 F98 F99 T100 T101 T102 T103 T104 T105 T106 T107 T108 T109 T110 T111 T112 T113 T114 T115 T116 T117 T118 T119 T120 T121 T122 T123 T124

T129 S130 G131 G132 A133 V135 V136 C137 F138 F139 M140 M141 F142 F143 F144 K145 ASP I147 H148 V149 W151 K152 L153 D154 G155 S156 S157 R158 F159 M160 M161 M162 L163 S165 S166 S167 T167 D168 Q169 D170 D171 K172 K173 K174 S174 S175 Y176 S177 M178 S179 S180 T181 T182 T183 T184 THR LYS ASP GLU

Y189 E190 H192 M193 S194 Y195 T196 C197 E198 T203 S204 T205 T208 V209 K210 N213 ARG ASN GLU CYS

- Molecule 3: Fab fragment, light chain

Chain I: 16% 55% 19% 7%

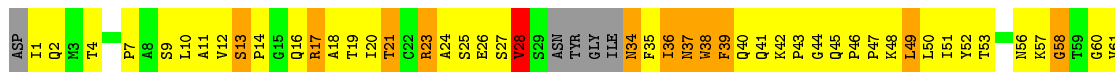
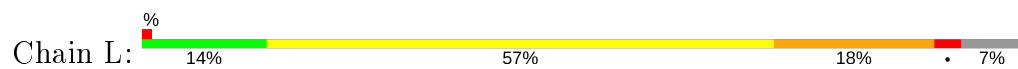
ASP I1 Q2 M3 T4 S6 F7 A8 S9 L10 A11 V12 S13 Q16 R17 A18 T19 I20 T21 A24 S25 E26 S27 V28 S29 ASN TVR GLY ILE N34 F35 F36 F37 F38 F39 F40 F41 F42 F43 F44 F45 F46 F47 F48 F49 F50 F51 F52 F53 F54 F55 F56 F57 F58 F59 F60 F61 F62 F63 F64 F65 F66 F67 F68 F69 F70 F71 F72 F73 F74 F75 F76 F77 F78 F79 F80 F81 F82 F83 F84 F85 F86 F87 F88 F89 F90 F91 F92 F93 F94 F95 F96 F97 F98 F99 T100 T101 T102 T103 T104 T105 T106 T107 T108 T109 T110 T111 T112 T113 T114 T115 T116 T117 T118 T119 T120 T121 T122 T123 T124

F65 S66 G67 S70 G71 F72 T75 F76 F77 F78 M79 F80 F81 F82 F83 F84 F85 F86 F87 F88 F89 F90 F91 F92 F93 F94 F95 F96 F97 F98 F99 T100 T101 T102 T103 T104 T105 T106 T107 T108 T109 T110 T111 T112 T113 T114 T115 T116 T117 T118 T119 T120 T121 T122 T123 T124 E126 F127

L128 T129 S130 G131 G132 A133 S134 V135 V136 C137 F138 F139 M140 M141 F142 F143 F144 K145 ASP I147 M148 V149 W151 K152 L153 D154 G155 S156 S157 R158 F159 M160 M161 M162 L163 S165 S166 S167 T167 D168 Q169 D170 D171 K172 K173 K174 S174 S175 Y176 S177 M178 S179 S180 T181 T182 T183 T184 THR LYS ASP



• Molecule 3: Fab fragment, light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.70Å 90.13Å 238.18Å 90.00° 90.05° 90.00°	Depositor
Resolution (Å)	49.68 – 2.81 49.68 – 2.81	Depositor EDS
% Data completeness (in resolution range)	87.5 (49.68-2.81) 84.4 (49.68-2.81)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.46 (at 2.81Å)	Xtrriage
Refinement program	PHENIX 1.7.1_743	Depositor
R, $R_{free}$	0.233 , 0.289 0.236 , 0.295	Depositor DCC
$R_{free}$ test set	3381 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.2	Xtrriage
Anisotropy	0.531	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 21.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.457 for h,-k,-l	Xtrriage
Reported twinning fraction	0.492 for h,-k,-l	Depositor
Outliers	2 of 66880 reflections (0.003%)	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	19900	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 50.84 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.1107e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.39	0/1861	0.63	0/2523
1	D	0.41	0/1919	0.67	0/2602
1	G	0.42	0/1826	0.69	0/2474
1	J	0.41	0/1965	0.67	1/2667 (0.0%)
2	B	0.45	0/1577	0.76	3/2141 (0.1%)
2	E	0.42	0/1577	0.74	3/2141 (0.1%)
2	H	0.43	0/1577	0.73	0/2141
2	K	0.43	0/1577	0.72	0/2141
3	C	0.50	0/1590	0.86	4/2157 (0.2%)
3	F	0.51	0/1590	0.79	1/2157 (0.0%)
3	I	0.78	1/1591 (0.1%)	0.85	5/2160 (0.2%)
3	L	0.84	1/1591 (0.1%)	0.90	7/2160 (0.3%)
All	All	0.52	2/20241 (0.0%)	0.75	24/27464 (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	2
1	D	0	3
1	G	0	3
1	J	0	2
2	B	0	1
2	E	0	1
2	H	0	1
2	K	0	1
3	C	0	8
3	F	0	13
3	I	0	6
3	L	0	2
All	All	0	43

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	L	145	LYS	C-N	27.80	1.98	1.34
3	I	145	LYS	C-N	24.11	1.89	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	L	145	LYS	O-C-N	-12.89	102.07	122.70
3	L	145	LYS	C-N-CA	9.08	144.39	121.70
2	E	43	GLY	N-CA-C	-6.80	96.09	113.10
2	B	43	GLY	N-CA-C	-6.76	96.19	113.10
3	C	94	THR	N-CA-C	-6.63	93.11	111.00

There are no chirality outliers.

5 of 43 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	129	ASN	Peptide
1	A	74	SER	Peptide
2	B	44	LEU	Peptide
3	C	28	VAL	Peptide
3	C	8	ALA	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	1813	0	1716	191	0
1	D	1872	0	1743	214	0
1	G	1778	0	1702	239	0
1	J	1918	0	1760	219	0
2	B	1544	0	1505	262	0
2	E	1544	0	1505	254	0
2	H	1544	0	1505	274	0
2	K	1544	0	1505	247	0
3	C	1557	0	1503	343	0
3	F	1557	0	1503	355	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	I	1557	0	1503	317	0
3	L	1557	0	1503	332	0
4	A	11	0	0	2	0
4	B	8	0	0	3	0
4	C	11	0	0	3	0
4	D	7	0	0	1	0
4	E	11	0	0	3	0
4	F	9	0	0	5	0
4	G	8	0	0	2	0
4	H	10	0	0	3	0
4	I	12	0	0	2	0
4	J	7	0	0	3	0
4	K	6	0	0	4	0
4	L	15	0	0	8	0
All	All	19900	0	18953	3059	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 79.

The worst 5 of 3059 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)
3:I:145:LYS:C	3:I:147:ILE:N	1.89	1.25
3:C:38:TRP:CD2	3:C:39:PHE:HA	1.74	1.21
2:B:171:VAL:HG21	3:C:176:TYR:CE1	1.76	1.19
3:L:145:LYS:C	3:L:147:ILE:N	1.98	1.16
2:E:32:ASP:HB3	2:E:51:LEU:HA	1.27	1.15

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	230/518 (44%)	198 (86%)	29 (13%)	3 (1%)	12	34
1	D	240/518 (46%)	199 (83%)	35 (15%)	6 (2%)	5	18
1	G	223/518 (43%)	193 (86%)	27 (12%)	3 (1%)	12	34
1	J	249/518 (48%)	211 (85%)	35 (14%)	3 (1%)	13	37
2	B	202/219 (92%)	175 (87%)	25 (12%)	2 (1%)	15	42
2	E	202/219 (92%)	172 (85%)	29 (14%)	1 (0%)	29	59
2	H	202/219 (92%)	174 (86%)	26 (13%)	2 (1%)	15	42
2	K	202/219 (92%)	177 (88%)	21 (10%)	4 (2%)	7	23
3	C	193/218 (88%)	150 (78%)	37 (19%)	6 (3%)	4	13
3	F	193/218 (88%)	142 (74%)	38 (20%)	13 (7%)	1	3
3	I	195/218 (89%)	141 (72%)	45 (23%)	9 (5%)	2	7
3	L	195/218 (89%)	145 (74%)	44 (23%)	6 (3%)	4	13
All	All	2526/3820 (66%)	2077 (82%)	391 (16%)	58 (2%)	6	20

5 of 58 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	8	ALA
3	F	173	ASP
3	L	144	PRO
3	L	147	ILE
1	A	120	THR

### 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	190/451 (42%)	161 (85%)	29 (15%)	2	8
1	D	190/451 (42%)	154 (81%)	36 (19%)	1	4
1	G	190/451 (42%)	157 (83%)	33 (17%)	2	5
1	J	190/451 (42%)	156 (82%)	34 (18%)	2	5

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	173/182 (95%)	130 (75%)	43 (25%)	0	1
2	E	173/182 (95%)	129 (75%)	44 (25%)	0	1
2	H	173/182 (95%)	140 (81%)	33 (19%)	1	4
2	K	173/182 (95%)	139 (80%)	34 (20%)	1	4
3	C	177/190 (93%)	128 (72%)	49 (28%)	0	1
3	F	177/190 (93%)	135 (76%)	42 (24%)	1	2
3	I	177/190 (93%)	130 (73%)	47 (27%)	0	1
3	L	177/190 (93%)	125 (71%)	52 (29%)	0	1
All	All	2160/3292 (66%)	1684 (78%)	476 (22%)	1	2

5 of 476 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	82	GLU
1	G	187	ASP
3	L	75	THR
3	F	118	VAL
3	F	198	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	141	ASN
1	G	129	ASN
3	L	92	GLN
3	F	193	ASN
1	G	193	GLN

### 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 carbohydrates 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	234/518 (45%)	0.03	14 (5%) 21 14	19, 35, 59, 76	8 (3%)
1	D	246/518 (47%)	0.25	26 (10%) 6 3	18, 32, 58, 95	21 (8%)
1	G	227/518 (43%)	-0.27	6 (2%) 56 46	18, 32, 57, 70	2 (0%)
1	J	255/518 (49%)	0.30	31 (12%) 4 2	21, 33, 57, 71	30 (11%)
2	B	208/219 (94%)	-0.37	0 100 100	13, 29, 44, 58	0
2	E	208/219 (94%)	-0.44	0 100 100	19, 30, 42, 62	0
2	H	208/219 (94%)	-0.40	4 (1%) 66 59	20, 29, 43, 53	0
2	K	208/219 (94%)	-0.35	2 (0%) 82 77	21, 29, 41, 48	0
3	C	203/218 (93%)	-0.29	1 (0%) 91 88	13, 28, 50, 68	0
3	F	203/218 (93%)	-0.24	3 (1%) 73 67	20, 29, 49, 58	0
3	I	203/218 (93%)	-0.23	1 (0%) 91 88	20, 30, 53, 71	0
3	L	203/218 (93%)	-0.22	3 (1%) 73 67	18, 30, 57, 75	0
All	All	2606/3820 (68%)	-0.17	91 (3%) 44 34	13, 31, 54, 95	61 (2%)

The worst 5 of 91 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	388	LYS	13.4
1	J	387	GLU	12.2
1	A	389	MET	12.2
1	D	419	LEU	11.1
1	D	389	MET	10.8

### 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 carbohydrates 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.