



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 06:44 AM BST

PDB ID : 6D0U  
Title : Crystal structure of C05 V110P/A117E mutant bound to H3 influenza hemagglutinin, HA1 subunit  
Authors : Wu, N.C.; Wilson, I.A.  
Deposited on : 2018-04-11  
Resolution : 3.25 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

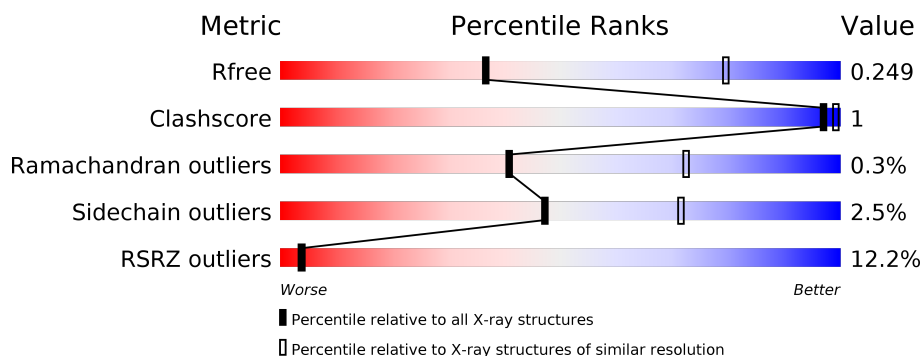
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.25 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	1191 (3.30-3.22)
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (3.30-3.22)

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	274	<div> <div>12%</div> <div>95%</div> <div>..</div> </div>
1	B	274	<div> <div>7%</div> <div>95%</div> <div>..</div> </div>
1	G	274	<div> <div>9%</div> <div>93%</div> <div>...</div> </div>
1	J	274	<div> <div>9%</div> <div>93%</div> <div>...</div> </div>
2	C	247	<div> <div>15%</div> <div>94%</div> <div>..</div> </div>
2	E	247	<div> <div>14%</div> <div>91%</div> <div>5% .</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	247	<div><div></div><div>12%</div><div>92%</div><div></div><div></div></div>
2	K	247	<div><div></div><div>16%</div><div>92%</div><div></div><div></div></div>
3	D	214	<div><div></div><div>20%</div><div>94%</div><div></div><div>6%</div></div>
3	F	214	<div><div></div><div>6%</div><div>95%</div><div></div><div></div></div>
3	I	214	<div><div></div><div>4%</div><div>95%</div><div></div><div></div></div>
3	L	214	<div><div></div><div>22%</div><div>90%</div><div></div><div>9%</div></div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 21964 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
1	A	267	Total	C	N	O	S	0	0	0
			2073	1303	363	396	11			
1	B	267	Total	C	N	O	S	0	0	0
			2073	1303	363	396	11			
1	J	267	Total	C	N	O	S	0	0	0
			2073	1303	363	396	11			
1	G	267	Total	C	N	O	S	0	0	0
			2073	1303	363	396	11			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	310	GLY	-	expression tag	UNP Q91MA7
A	311	HIS	-	expression tag	UNP Q91MA7
A	312	HIS	-	expression tag	UNP Q91MA7
A	313	HIS	-	expression tag	UNP Q91MA7
A	314	HIS	-	expression tag	UNP Q91MA7
A	315	HIS	-	expression tag	UNP Q91MA7
A	316	HIS	-	expression tag	UNP Q91MA7
B	310	GLY	-	expression tag	UNP Q91MA7
B	311	HIS	-	expression tag	UNP Q91MA7
B	312	HIS	-	expression tag	UNP Q91MA7
B	313	HIS	-	expression tag	UNP Q91MA7
B	314	HIS	-	expression tag	UNP Q91MA7
B	315	HIS	-	expression tag	UNP Q91MA7
B	316	HIS	-	expression tag	UNP Q91MA7
J	310	GLY	-	expression tag	UNP Q91MA7
J	311	HIS	-	expression tag	UNP Q91MA7
J	312	HIS	-	expression tag	UNP Q91MA7
J	313	HIS	-	expression tag	UNP Q91MA7
J	314	HIS	-	expression tag	UNP Q91MA7
J	315	HIS	-	expression tag	UNP Q91MA7
J	316	HIS	-	expression tag	UNP Q91MA7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	310	GLY	-	expression tag	UNP Q91MA7
G	311	HIS	-	expression tag	UNP Q91MA7
G	312	HIS	-	expression tag	UNP Q91MA7
G	313	HIS	-	expression tag	UNP Q91MA7
G	314	HIS	-	expression tag	UNP Q91MA7
G	315	HIS	-	expression tag	UNP Q91MA7
G	316	HIS	-	expression tag	UNP Q91MA7

- Molecule 2 is a protein called Antibody C05 V110P/A117E mutant, heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	238	Total	C	N	O	S	0	1	0
			1777	1111	298	360	8			
2	C	238	Total	C	N	O	S	0	0	0
			1771	1108	297	358	8			
2	K	238	Total	C	N	O	S	0	0	0
			1771	1108	297	358	8			
2	H	238	Total	C	N	O	S	0	1	0
			1777	1111	298	360	8			

- Molecule 3 is a protein called Antibody C05, light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	213	Total	C	N	O	S	0	0	0
			1637	1027	277	329	4			
3	D	213	Total	C	N	O	S	0	0	0
			1637	1027	277	329	4			
3	L	213	Total	C	N	O	S	0	0	0
			1637	1027	277	329	4			
3	I	213	Total	C	N	O	S	0	0	0
			1637	1027	277	329	4			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).

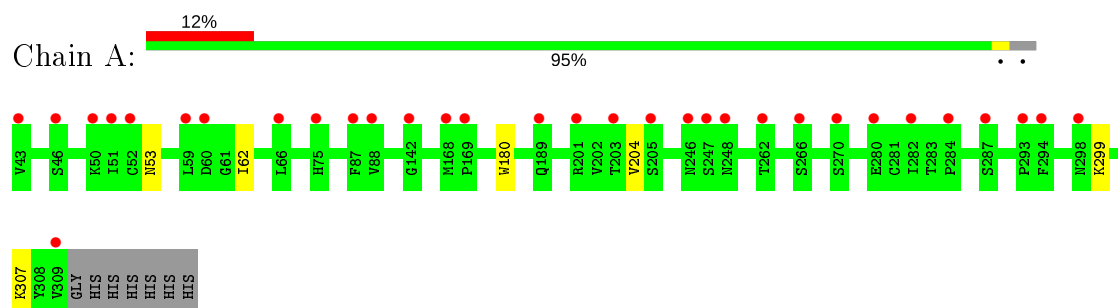


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	J	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		

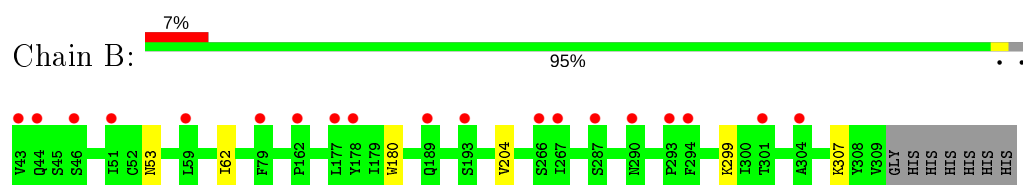
### 3 Residue-property plots [i](#)

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.

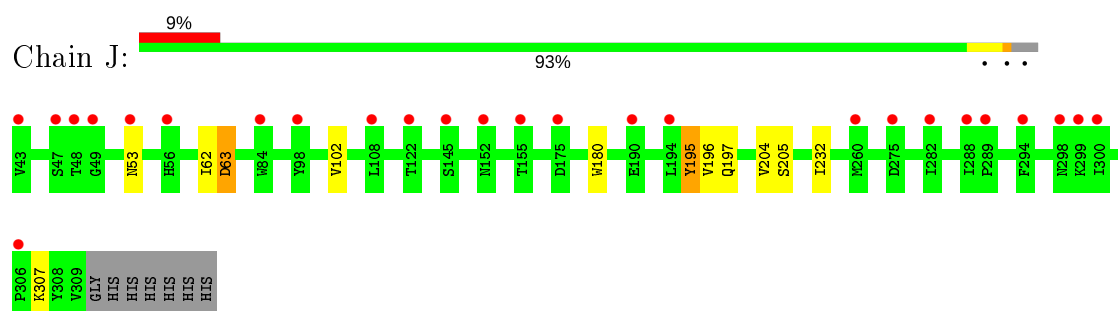
- Molecule 1: Hemagglutinin



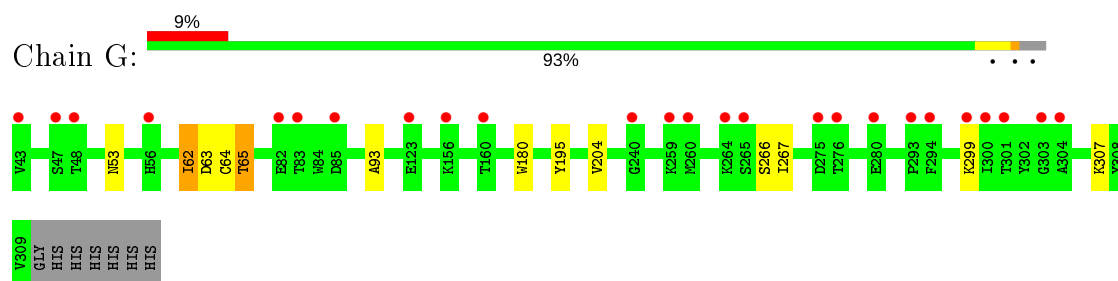
- Molecule 1: Hemagglutinin



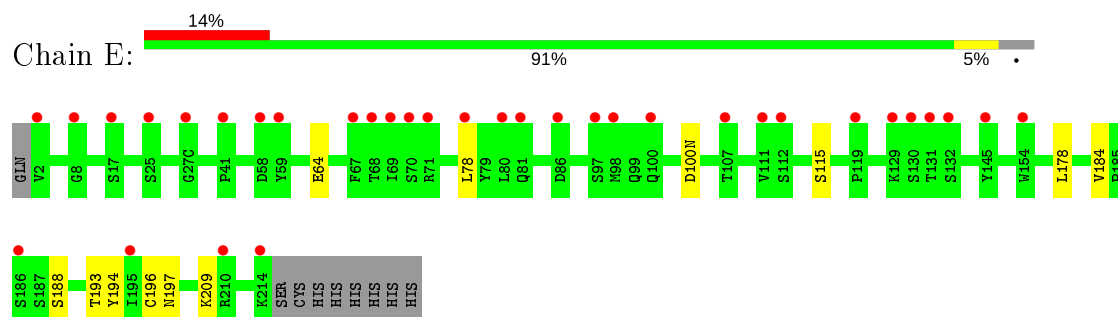
- Molecule 1: Hemagglutinin



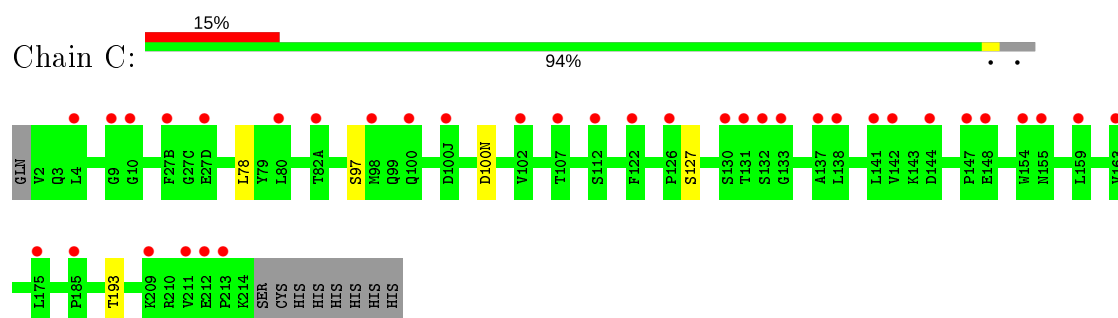
- Molecule 1: Hemagglutinin



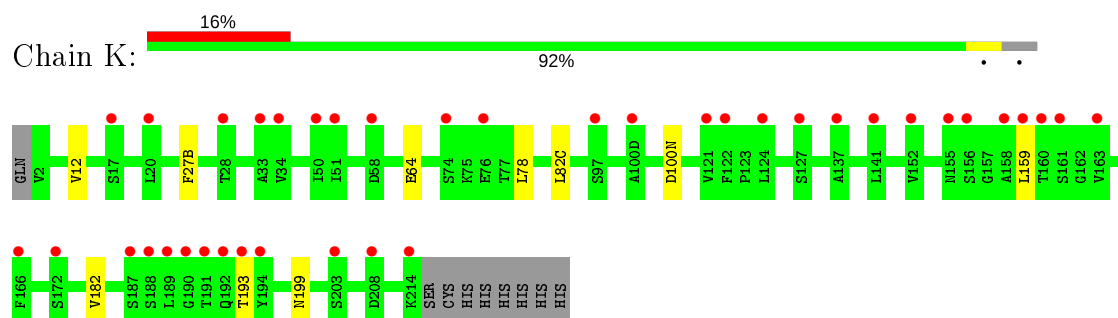
- Molecule 2: Antibody C05 V110P/A117E mutant, heavy chain



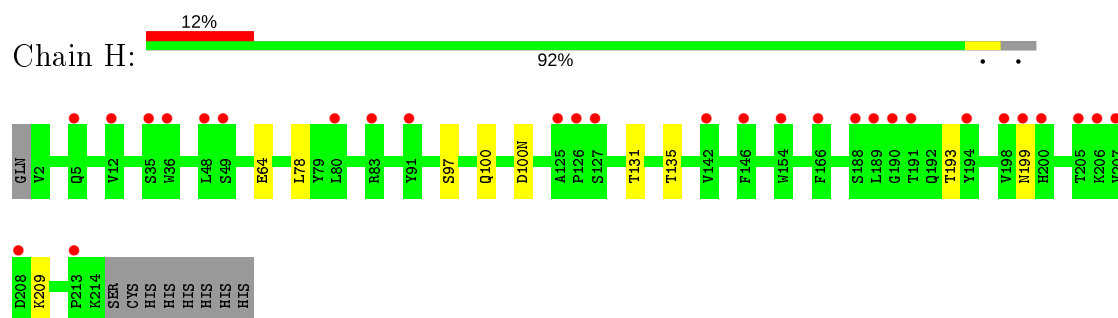
- Molecule 2: Antibody C05 V110P/A117E mutant, heavy chain



- Molecule 2: Antibody C05 V110P/A117E mutant, heavy chain



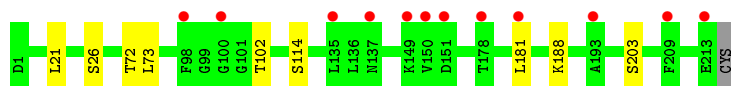
- Molecule 2: Antibody C05 V110P/A117E mutant, heavy chain



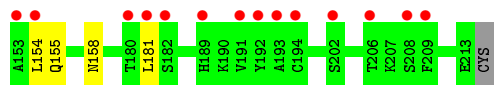
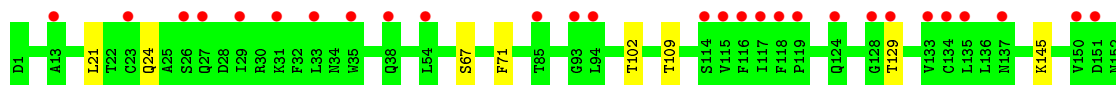
- Molecule 3: Antibody C05, light chain



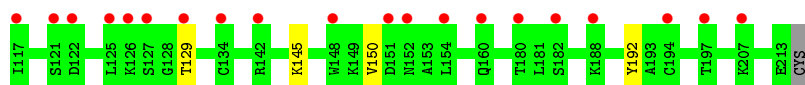
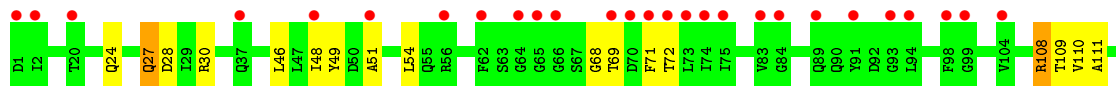
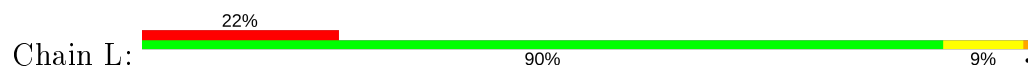




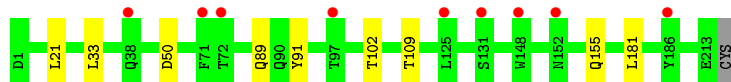
- Molecule 3: Antibody C05, light chain



- Molecule 3: Antibody C05, light chain



- Molecule 3: Antibody C05, light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.61Å 258.42Å 91.88Å 90.00° 90.54° 90.00°	Depositor
Resolution (Å)	50.00 – 3.25 45.94 – 3.23	Depositor EDS
% Data completeness (in resolution range)	94.0 (50.00-3.25) 97.8 (45.94-3.23)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.16	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.91 (at 3.25Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.254 , 0.268 0.238 , 0.249	Depositor DCC
$R_{free}$ test set	3191 reflections (4.77%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	96.0	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 35.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.035 for l,k,-h 0.043 for h,-k,-l 0.032 for l,-k,h	Xtriage
Reported twinning fraction	0.473 for H, K, L 0.527 for -L, K, H	Depositor
Outliers	0 of 66894 reflections	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	21964	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.45% of the height of the origin peak. No significant pseudotranslation is detected.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	0.59	0/2125	0.69	0/2895
1	B	0.58	0/2125	0.70	0/2895
1	G	0.60	0/2125	0.71	0/2895
1	J	0.60	0/2125	0.71	0/2895
2	C	0.57	0/1810	0.71	0/2463
2	E	0.60	0/1816	0.69	0/2471
2	H	0.60	0/1816	0.71	0/2471
2	K	0.60	0/1810	0.72	0/2463
3	D	0.65	0/1671	0.73	0/2266
3	F	0.64	0/1671	0.72	0/2266
3	I	0.64	0/1671	0.71	0/2266
3	L	0.65	0/1671	0.76	0/2266
All	All	0.61	0/22436	0.71	0/30512

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2073	0	2015	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2073	0	2015	2	0
1	G	2073	0	2015	7	0
1	J	2073	0	2014	5	0
2	C	1771	0	1716	0	0
2	E	1777	0	1720	2	0
2	H	1777	0	1720	1	0
2	K	1771	0	1716	2	0
3	D	1637	0	1602	3	0
3	F	1637	0	1602	2	0
3	I	1637	0	1602	3	0
3	L	1637	0	1602	12	0
4	G	14	0	13	0	0
4	J	14	0	13	0	0
All	All	21964	0	21365	40	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 1.

All (40) 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:L:27:GLN:O	3:L:69:THR:HG22	1.93	0.68
1:G:63:ASP:HA	1:G:93:ALA:HA	1.78	0.65
3:D:155:GLN:HE21	3:D:158:ASN:HD21	1.44	0.64
3:D:67:SER:HA	3:D:71:PHE:CZ	2.33	0.64
3:L:108:ARG:NE	3:L:109:THR:O	2.31	0.63
3:L:30:ARG:O	3:L:71:PHE:HZ	1.85	0.58
1:G:64:CYS:O	1:G:65:THR:O	2.23	0.57
3:L:108:ARG:NH2	3:L:111:ALA:HB2	2.24	0.53
3:I:50:ASP:OD1	3:I:91:TYR:OH	2.22	0.52
3:I:33:LEU:HA	3:I:89:GLN:O	2.11	0.51
1:J:180:TRP:CE2	1:J:204:VAL:HG21	2.46	0.50
1:B:180:TRP:CE2	1:B:204:VAL:HG21	2.47	0.49
1:G:64:CYS:O	1:G:65:THR:C	2.50	0.49
2:E:184:VAL:HG11	2:E:194:TYR:CE2	2.47	0.49
1:G:266:SER:OG	1:G:267:ILE:N	2.47	0.47
3:F:21:LEU:HD22	3:F:73:LEU:HD23	1.96	0.47
3:D:21:LEU:HG	3:D:102:THR:HG21	1.96	0.46
1:G:62:ILE:H	1:G:62:ILE:HD12	1.81	0.46
3:L:150:VAL:HG13	3:L:192:TYR:CE1	2.50	0.45
3:F:21:LEU:HG	3:F:102:THR:HG21	1.98	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:196:CYS:O	2:E:196:CYS:SG	2.75	0.45
2:H:131:THR:HA	2:H:135:THR:O	2.17	0.45
3:I:21:LEU:HG	3:I:102:THR:HG21	1.98	0.45
1:G:64:CYS:C	1:G:65:THR:O	2.56	0.44
2:K:159:LEU:HD21	2:K:182:VAL:HG21	1.99	0.44
1:J:102:VAL:HG22	1:J:232:ILE:HB	2.01	0.43
3:L:108:ARG:HH21	3:L:111:ALA:HB2	1.83	0.43
3:L:108:ARG:HH21	3:L:110:VAL:C	2.21	0.43
3:L:108:ARG:NH2	3:L:110:VAL:O	2.53	0.42
1:A:180:TRP:CD2	1:A:204:VAL:HG21	2.54	0.42
1:B:180:TRP:CD2	1:B:204:VAL:HG21	2.55	0.42
2:K:12:VAL:HG21	2:K:82(C):LEU:HD13	2.02	0.41
3:L:46:LEU:HD11	3:L:49:TYR:HB3	2.02	0.41
3:L:48:ILE:HD13	3:L:54:LEU:HA	2.01	0.41
1:J:195:TYR:O	1:J:197:GLN:N	2.53	0.41
1:J:63:ASP:OD1	1:J:63:ASP:N	2.54	0.41
3:L:46:LEU:HD21	3:L:49:TYR:HB3	2.02	0.41
1:G:180:TRP:CE2	1:G:204:VAL:HG21	2.56	0.41
3:L:28:ASP:OD1	3:L:68:GLY:HA2	2.21	0.41
1:J:180:TRP:CD2	1:J:204:VAL:HG21	2.56	0.40

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	265/274 (97%)	251 (95%)	13 (5%)	1 (0%)	34 67
1	B	265/274 (97%)	252 (95%)	12 (4%)	1 (0%)	34 67
1	G	265/274 (97%)	251 (95%)	13 (5%)	1 (0%)	34 67
1	J	265/274 (97%)	254 (96%)	10 (4%)	1 (0%)	34 67

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	236/247 (96%)	227 (96%)	8 (3%)	1 (0%)	34	67
2	E	237/247 (96%)	228 (96%)	8 (3%)	1 (0%)	34	67
2	H	237/247 (96%)	226 (95%)	10 (4%)	1 (0%)	34	67
2	K	236/247 (96%)	228 (97%)	7 (3%)	1 (0%)	34	67
3	D	211/214 (99%)	202 (96%)	9 (4%)	0	100	100
3	F	211/214 (99%)	205 (97%)	6 (3%)	0	100	100
3	I	211/214 (99%)	206 (98%)	5 (2%)	0	100	100
3	L	211/214 (99%)	205 (97%)	5 (2%)	1 (0%)	29	62
All	All	2850/2940 (97%)	2735 (96%)	106 (4%)	9 (0%)	41	72

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	65	THR
3	L	51	ALA
2	C	100(N)	ASP
2	K	100(N)	ASP
2	H	100(N)	ASP
1	A	62	ILE
2	E	100(N)	ASP
1	B	62	ILE
1	J	196	VAL

### 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	236/242 (98%)	233 (99%)	3 (1%)	69	82
1	B	236/242 (98%)	233 (99%)	3 (1%)	69	82
1	G	236/242 (98%)	231 (98%)	5 (2%)	53	75
1	J	236/242 (98%)	230 (98%)	6 (2%)	47	71

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	199/208 (96%)	195 (98%)	4 (2%)	55	76
2	E	200/208 (96%)	192 (96%)	8 (4%)	31	61
2	H	200/208 (96%)	193 (96%)	7 (4%)	36	64
2	K	199/208 (96%)	194 (98%)	5 (2%)	47	71
3	D	186/187 (100%)	180 (97%)	6 (3%)	39	66
3	F	186/187 (100%)	180 (97%)	6 (3%)	39	66
3	I	186/187 (100%)	183 (98%)	3 (2%)	62	79
3	L	186/187 (100%)	180 (97%)	6 (3%)	39	66
All	All	2486/2548 (98%)	2424 (98%)	62 (2%)	47	71

All (62) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	53	ASN
1	A	299	LYS
1	A	307	LYS
2	E	64	GLU
2	E	78	LEU
2	E	115	SER
2	E	178	LEU
2	E	188	SER
2	E	193	THR
2	E	197	ASN
2	E	209	LYS
3	F	26	SER
3	F	72	THR
3	F	114	SER
3	F	181	LEU
3	F	188	LYS
3	F	203	SER
1	B	53	ASN
1	B	299	LYS
1	B	307	LYS
2	C	78	LEU
2	C	97	SER
2	C	127	SER
2	C	193	THR
3	D	24	GLN
3	D	109	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	D	129	THR
3	D	145	LYS
3	D	154	LEU
3	D	181	LEU
1	J	53	ASN
1	J	62	ILE
1	J	63	ASP
1	J	195	TYR
1	J	205	SER
1	J	307	LYS
2	K	27(B)	PHE
2	K	64	GLU
2	K	78	LEU
2	K	193	THR
2	K	199	ASN
3	L	24	GLN
3	L	27	GLN
3	L	72	THR
3	L	108	ARG
3	L	129	THR
3	L	145	LYS
2	H	64	GLU
2	H	78	LEU
2	H	97	SER
2	H	100	GLN
2	H	193	THR
2	H	199	ASN
2	H	209	LYS
3	I	109	THR
3	I	155	GLN
3	I	181	LEU
1	G	53	ASN
1	G	62	ILE
1	G	195	TYR
1	G	299	LYS
1	G	307	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (31) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	53	ASN
1	A	54	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
3	F	53	ASN
3	F	55	GLN
3	F	137	ASN
3	F	138	ASN
3	F	199	GLN
1	B	53	ASN
1	B	54	ASN
3	D	24	GLN
3	D	53	ASN
3	D	55	GLN
3	D	137	ASN
3	D	138	ASN
3	D	155	GLN
1	J	53	ASN
1	J	54	ASN
1	J	189	GLN
3	L	3	GLN
3	L	24	GLN
3	L	27	GLN
3	L	55	GLN
2	H	192	GLN
3	I	24	GLN
3	I	53	ASN
3	I	55	GLN
3	I	137	ASN
3	I	138	ASN
1	G	53	ASN
1	G	54	ASN
1	G	210	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 ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	NAG	J	401	1	14,14,15	0.36	0	17,19,21	0.79	0
4	NAG	G	401	-	14,14,15	0.34	0	17,19,21	0.97	1 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	J	401	1	-	3/6/23/26	0/1/1/1
4	NAG	G	401	-	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	401	NAG	O5-C1-C2	-2.63	107.14	111.29

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	J	401	NAG	C4-C5-C6-O6
4	J	401	NAG	O5-C5-C6-O6
4	J	401	NAG	C1-C2-N2-C7

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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, 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	267/274 (97%)	0.71	32 (11%) 4 4	86, 107, 149, 166	0
1	B	267/274 (97%)	0.57	19 (7%) 16 15	82, 103, 165, 180	0
1	G	267/274 (97%)	0.66	25 (9%) 8 9	84, 112, 141, 160	0
1	J	267/274 (97%)	0.71	26 (9%) 7 8	90, 105, 147, 184	0
2	C	238/247 (96%)	0.82	36 (15%) 2 2	103, 124, 160, 173	0
2	E	238/247 (96%)	0.83	34 (14%) 2 2	93, 107, 132, 165	0
2	H	238/247 (96%)	0.69	29 (12%) 4 4	96, 111, 140, 157	0
2	K	238/247 (96%)	0.92	39 (16%) 1 2	105, 129, 166, 199	0
3	D	213/214 (99%)	1.14	42 (19%) 1 1	109, 126, 183, 249	0
3	F	213/214 (99%)	0.48	12 (5%) 24 22	86, 101, 116, 136	0
3	I	213/214 (99%)	0.39	9 (4%) 36 33	89, 99, 111, 123	0
3	L	213/214 (99%)	1.10	47 (22%) 0 1	120, 134, 184, 211	0
All	All	2872/2940 (97%)	0.75	350 (12%) 4 4	82, 111, 157, 249	0

All (350) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	181	LEU	12.4
3	D	116	PHE	9.7
2	K	189	LEU	9.1
3	D	154	LEU	8.3
2	C	10	GLY	8.1
1	J	47	SER	7.6
2	K	193	THR	7.6
3	D	180	THR	7.6
1	A	51	ILE	7.4
1	J	43	VAL	7.3
1	A	205	SER	7.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	D	192	TYR	7.2
3	D	193	ALA	7.1
2	H	207	VAL	6.9
1	B	46	SER	6.9
2	C	132	SER	6.8
2	K	122	PHE	6.8
1	J	48	THR	6.7
3	L	154	LEU	6.3
3	D	114	SER	6.3
3	D	117	ILE	6.2
1	G	56	HIS	6.2
3	D	150	VAL	5.9
1	J	56	HIS	5.8
2	E	214	LYS	5.8
3	F	181	LEU	5.8
3	D	182	SER	5.5
2	E	131	THR	5.5
3	L	129	THR	5.4
1	A	60	ASP	5.4
2	C	126	PRO	5.3
1	A	189	GLN	5.3
3	D	133	VAL	5.2
2	K	194	TYR	5.2
1	J	155	THR	5.2
1	A	246	ASN	5.0
2	C	27(B)	PHE	5.0
1	A	293	PRO	5.0
3	D	135	LEU	4.9
1	J	306	PRO	4.9
3	L	94	LEU	4.9
3	L	152	ASN	4.8
1	J	299	LYS	4.8
1	A	309	VAL	4.7
3	L	72	THR	4.7
2	H	36	TRP	4.7
1	G	48	THR	4.7
3	L	74	ILE	4.6
3	D	208	SER	4.5
1	J	49	GLY	4.5
2	H	189	LEU	4.4
3	L	122	ASP	4.4
2	K	161	SER	4.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	G	83	THR	4.4
3	L	99	GLY	4.4
1	B	59	LEU	4.3
2	C	137	ALA	4.3
2	H	198	VAL	4.3
2	H	91	TYR	4.3
1	A	298	ASN	4.3
2	E	112	SER	4.2
3	L	134	CYS	4.2
2	C	147	PRO	4.2
1	J	145	SER	4.2
3	D	191	VAL	4.1
2	K	34	VAL	4.1
2	C	107	THR	4.0
3	D	118	PHE	4.0
1	G	47	SER	4.0
2	K	127	SER	4.0
2	H	127	SER	4.0
3	D	137	ASN	4.0
3	L	1	ASP	3.9
1	J	260	MET	3.9
2	K	137	ALA	3.9
3	L	65	GLY	3.9
3	F	150	VAL	3.8
3	L	125	LEU	3.8
2	K	141	LEU	3.8
1	A	203	THR	3.8
1	A	287	SER	3.8
3	D	153	ALA	3.8
1	G	259	LYS	3.8
2	C	122	PHE	3.7
2	K	121	VAL	3.7
2	E	119	PRO	3.7
3	F	178	THR	3.7
1	G	43	VAL	3.7
3	L	127	SER	3.6
3	D	151	ASP	3.6
3	D	119	PRO	3.6
3	I	72	THR	3.6
1	A	59	LEU	3.6
3	L	89	GLN	3.5
1	B	162	PRO	3.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	G	82	GLU	3.5
1	G	294	PHE	3.5
1	B	193	SER	3.5
2	C	9	GLY	3.5
3	F	209	PHE	3.5
1	B	189	GLN	3.5
2	E	186	SER	3.5
3	D	129	THR	3.5
3	I	131	SER	3.4
1	A	169	PRO	3.4
1	B	293	PRO	3.4
2	K	166	PHE	3.4
2	K	33	ALA	3.4
2	K	192	GLN	3.4
2	C	144	ASP	3.4
2	H	206	LYS	3.4
2	C	148	GLU	3.4
3	D	115	VAL	3.4
2	E	69	ILE	3.4
2	E	2	VAL	3.4
2	H	125	ALA	3.4
2	C	131	THR	3.3
2	H	35	SER	3.3
1	A	266	SER	3.3
2	K	190	GLY	3.3
2	C	141	LEU	3.3
3	L	188	LYS	3.3
3	L	2	ILE	3.3
3	D	13	ALA	3.3
3	L	91	TYR	3.3
1	G	260	MET	3.2
2	E	130	SER	3.2
2	C	82(A)	THR	3.2
3	D	206	THR	3.2
1	A	280	GLU	3.2
2	E	68	THR	3.2
2	C	212	GLU	3.2
2	C	211	VAL	3.2
2	C	175	LEU	3.2
2	K	158	ALA	3.2
1	G	301	THR	3.2
1	G	293	PRO	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	J	190	GLU	3.2
2	E	41	PRO	3.2
1	J	53	ASN	3.2
3	F	193	ALA	3.2
2	K	172	SER	3.1
3	L	64	GLY	3.1
1	J	84	TRP	3.1
2	E	78	LEU	3.1
2	K	124	LEU	3.1
2	C	159	LEU	3.1
3	D	134	CYS	3.1
3	L	194	CYS	3.1
2	K	50	ILE	3.1
2	E	107	THR	3.1
2	H	191	THR	3.1
1	G	280	GLU	3.1
1	A	87	PHE	3.1
3	D	189	HIS	3.1
1	J	175	ASP	3.0
2	K	163	VAL	3.0
1	J	98	TYR	3.0
3	L	160	GLN	3.0
2	C	138	LEU	3.0
1	B	304	ALA	3.0
3	F	137	ASN	3.0
1	G	303	GLY	3.0
1	G	85	ASP	3.0
2	C	98	MET	3.0
3	D	124	GLN	3.0
3	L	98	PHE	2.9
1	G	265	SER	2.9
2	H	166	PHE	2.9
3	L	71	PHE	2.9
1	B	287	SER	2.9
2	E	129	LYS	2.9
2	K	74	SER	2.9
2	H	213	PRO	2.9
3	L	93	GLY	2.9
2	K	214	LYS	2.9
2	C	154	TRP	2.9
3	D	128	GLY	2.9
2	H	49	SER	2.8

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	L	126	LYS	2.8
2	H	208	ASP	2.8
2	C	133	GLY	2.8
1	A	142	GLY	2.8
2	K	208	ASP	2.8
2	K	152	VAL	2.8
2	K	155	ASN	2.8
3	I	152	ASN	2.8
1	A	294	PHE	2.8
2	K	187	SER	2.8
3	L	180	THR	2.8
3	I	71	PHE	2.8
1	A	247	SER	2.7
1	G	275	ASP	2.7
1	J	294	PHE	2.7
3	L	73	LEU	2.7
1	J	298	ASN	2.7
3	L	75	ILE	2.7
1	A	52	CYS	2.7
2	K	17	SER	2.7
3	F	149	LYS	2.7
1	A	262	THR	2.7
1	A	282	ILE	2.7
2	E	98	MET	2.7
2	H	154	TRP	2.7
1	B	267	ILE	2.7
3	L	48	ILE	2.6
2	E	86	ASP	2.6
3	L	66	GLY	2.6
2	H	126	PRO	2.6
2	C	130	SER	2.6
2	K	156	SER	2.6
2	C	185	PRO	2.6
3	I	125	LEU	2.6
2	K	159	LEU	2.6
2	E	100	GLN	2.6
1	J	275	ASP	2.6
3	D	202	SER	2.6
3	F	100	GLY	2.6
2	C	80	LEU	2.6
2	E	8	GLY	2.6
3	F	151	ASP	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	K	203	SER	2.5
3	I	148	TRP	2.5
3	L	151	ASP	2.5
1	B	44	GLN	2.5
3	L	20	THR	2.5
3	D	85	THR	2.5
3	L	84	GLY	2.5
1	G	123	GLU	2.5
1	A	75	HIS	2.5
2	K	160	THR	2.5
1	G	160	THR	2.4
3	L	197	THR	2.4
1	B	51	ILE	2.4
1	B	301	THR	2.4
1	B	178	TYR	2.4
1	B	266	SER	2.4
1	A	168	MET	2.4
2	H	200	HIS	2.4
2	E	59	TYR	2.4
1	G	300	ILE	2.4
1	J	108	LEU	2.4
3	D	94	LEU	2.4
1	A	50	LYS	2.4
2	K	100(D)	ALA	2.4
3	D	26	SER	2.4
2	H	199	ASN	2.4
1	J	300	ILE	2.4
1	A	284	PRO	2.4
1	A	43	VAL	2.4
3	L	69	THR	2.4
3	D	209	PHE	2.4
2	H	194	TYR	2.4
1	J	152	ASN	2.4
3	D	35	TRP	2.3
2	H	48	LEU	2.3
2	C	142	VAL	2.3
2	C	163	VAL	2.3
2	C	27(D)	GLU	2.3
3	L	70	ASP	2.3
2	H	83	ARG	2.3
3	D	29	ILE	2.3
1	J	122	THR	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	C	100(J)	ASP	2.3
3	I	186	TYR	2.3
2	K	28	THR	2.3
1	B	177	LEU	2.3
2	E	195	ILE	2.3
3	L	121	SER	2.3
3	L	148	TRP	2.3
3	F	213	GLU	2.3
3	D	31	LYS	2.3
2	K	51	ILE	2.3
2	C	213	PRO	2.3
3	D	23	CYS	2.3
3	L	104	VAL	2.3
1	J	289	PRO	2.3
2	H	190	GLY	2.3
2	E	71	ARG	2.3
3	L	51	ALA	2.3
3	D	54	LEU	2.2
1	J	288	ILE	2.2
1	B	79	PHE	2.2
3	D	38	GLN	2.2
1	J	194	LEU	2.2
1	B	43	VAL	2.2
1	G	264	LYS	2.2
2	C	4	LEU	2.2
2	C	112	SER	2.2
1	B	294	PHE	2.2
1	A	248	ASN	2.2
3	L	56	ARG	2.2
1	A	270	SER	2.2
1	A	201	ARG	2.2
2	E	210	ARG	2.2
2	C	102	VAL	2.2
2	H	142	VAL	2.2
2	E	145	TYR	2.2
1	G	156	LYS	2.2
3	L	62	PHE	2.2
1	G	240	GLY	2.2
3	D	93	GLY	2.2
1	G	304	ALA	2.2
3	L	37	GLN	2.2
1	A	46	SER	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	K	188	SER	2.2
3	L	182	SER	2.2
2	E	67	PHE	2.2
1	G	276	THR	2.2
1	G	299	LYS	2.2
2	K	191	THR	2.2
2	E	132	SER	2.1
3	F	98	PHE	2.1
2	E	111	VAL	2.1
2	H	12	VAL	2.1
3	L	142	ARG	2.1
2	H	188	SER	2.1
2	E	27(C)	GLY	2.1
2	C	155	ASN	2.1
2	E	25	SER	2.1
2	E	70	SER	2.1
2	K	97	SER	2.1
2	C	209	LYS	2.1
2	E	81	GLN	2.1
2	E	97	SER	2.1
3	D	33	LEU	2.1
2	E	80	LEU	2.1
3	D	194	CYS	2.1
3	L	117	ILE	2.1
2	H	146	PHE	2.1
2	C	100	GLN	2.1
2	H	80	LEU	2.1
3	F	135	LEU	2.1
1	J	282	ILE	2.1
1	A	66	LEU	2.1
2	E	154	TRP	2.1
2	H	205	THR	2.1
2	E	17	SER	2.1
2	E	58	ASP	2.1
2	K	58	ASP	2.1
3	L	207	LYS	2.0
1	B	290	ASN	2.0
2	K	20	LEU	2.0
3	I	38	GLN	2.0
2	H	5	GLN	2.0
1	A	88	VAL	2.0
3	I	97	THR	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	K	76	GLU	2.0
3	D	27	GLN	2.0
3	L	83	VAL	2.0

## 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	G	401	14/15	0.81	0.30	110,110,110,110	0
4	NAG	J	401	14/15	0.89	0.14	77,78,79,79	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.