



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 22, 2020 – 04:04 pm GMT

PDB ID : 6Z7W
Title : Human insulin in complex with the analytical antibody HUI-018 Fab
Authors : Johansson, E.
Deposited on : 2020-06-02
Resolution : 2.42 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 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 : 2.16
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) : 2.16

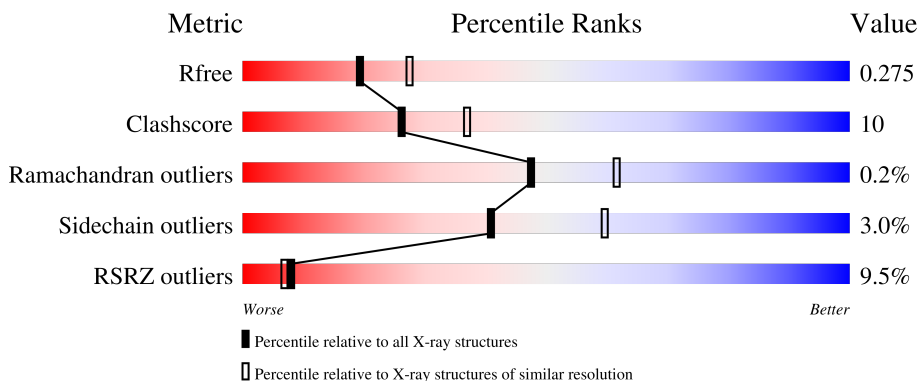
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.42 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	4647 (2.44-2.40)
Clashscore	141614	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)
RSRZ outliers	127900	4543 (2.44-2.40)

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	214	5% (Poor fit) 82% (0-1 outliers) 18% (2-3 outliers)
1	C	214	6% (Poor fit) 78% (0-1 outliers) 22% (2-3 outliers)
1	E	214	14% (Poor fit) 73% (0-1 outliers) 25% (2-3 outliers)
1	G	214	% (Poor fit) 81% (0-1 outliers) 18% (2-3 outliers)
2	B	224	12% (Poor fit) 75% (0-1 outliers) 22% (2-3 outliers)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	224	
2	F	224	
2	H	224	
3	I	21	
3	K	21	
3	M	21	
3	O	21	
4	J	30	
4	L	30	
4	N	30	
4	P	30	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 15292 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 MAb 6H10 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	214	Total 1672	C 1038	N 288	O 338	S 8	0	0	0
1	C	213	Total 1665	C 1035	N 287	O 336	S 7	0	0	0
1	E	214	Total 1672	C 1038	N 288	O 338	S 8	0	0	0
1	G	214	Total 1672	C 1038	N 288	O 338	S 8	0	0	0

- Molecule 2 is a protein called HUI-018 Fab Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	220	Total 1641	C 1032	N 274	O 328	S 7	0	0	0
2	D	213	Total 1593	C 1008	N 263	O 316	S 6	0	0	0
2	F	216	Total 1618	C 1021	N 269	O 321	S 7	0	0	0
2	H	219	Total 1636	C 1030	N 273	O 326	S 7	0	0	0

- Molecule 3 is a protein called Insulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	I	16	Total 126	C 76	N 19	O 27	S 4	0	0	0
3	K	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
3	M	21	Total 163	C 99	N 25	O 35	S 4	0	0	0
3	O	21	Total 163	C 99	N 25	O 35	S 4	0	0	0

- Molecule 4 is a protein called Insulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	J	24	Total	C	N	O	S	0	0	0
			190	125	31	32	2			
4	L	24	Total	C	N	O	S	0	0	0
			190	125	31	32	2			
4	N	26	Total	C	N	O	S	0	0	0
			207	135	34	36	2			
4	P	26	Total	C	N	O	S	0	0	0
			208	136	35	35	2			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	107	Total	O	0	0
			107	107		
5	B	78	Total	O	0	0
			78	78		
5	C	80	Total	O	0	0
			80	80		
5	D	84	Total	O	0	0
			84	84		
5	E	34	Total	O	0	0
			34	34		
5	F	49	Total	O	0	0
			49	49		
5	G	110	Total	O	0	0
			110	110		
5	H	115	Total	O	0	0
			115	115		
5	I	15	Total	O	0	0
			15	15		
5	J	3	Total	O	0	0
			3	3		
5	K	5	Total	O	0	0
			5	5		
5	L	6	Total	O	0	0
			6	6		
5	M	3	Total	O	0	0
			3	3		
5	N	7	Total	O	0	0
			7	7		
5	O	6	Total	O	0	0
			6	6		

Continued on next page...

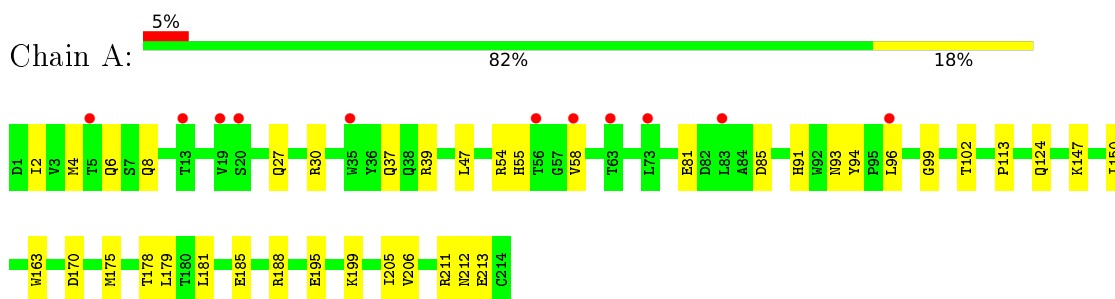
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	P	11	Total	O	0	0
			11	11		

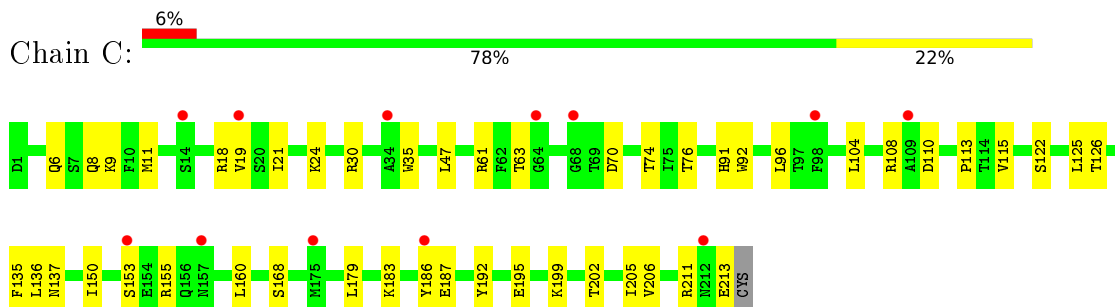
3 Residue-property plots [i](#)

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.

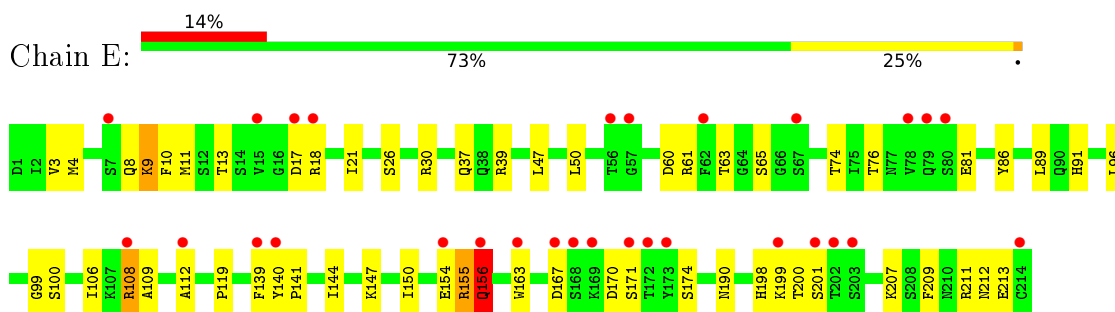
- Molecule 1: MAb 6H10 light chain



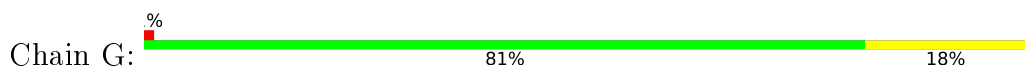
- Molecule 1: MAb 6H10 light chain

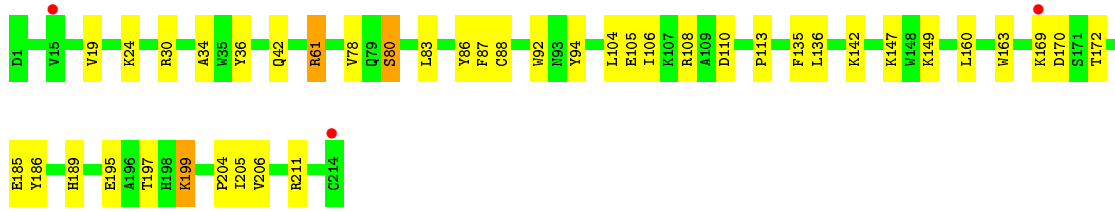


- Molecule 1: MAb 6H10 light chain

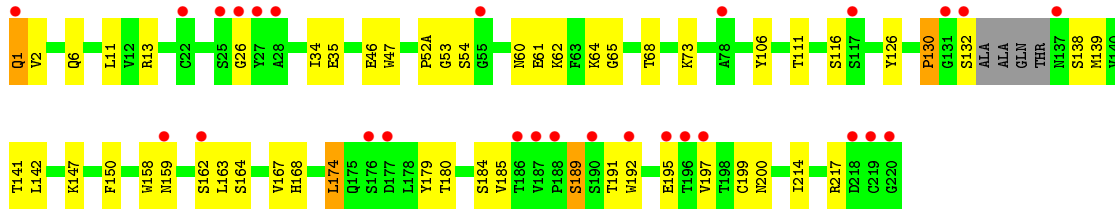
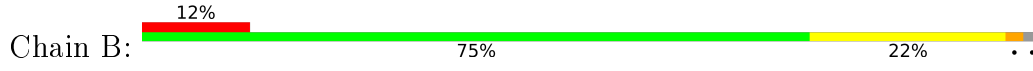


- Molecule 1: MAb 6H10 light chain

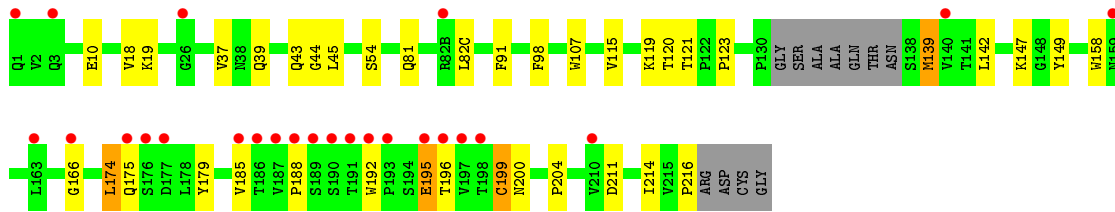
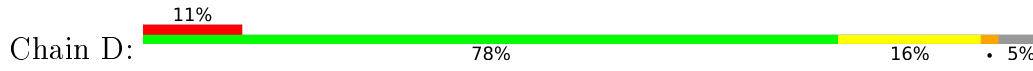




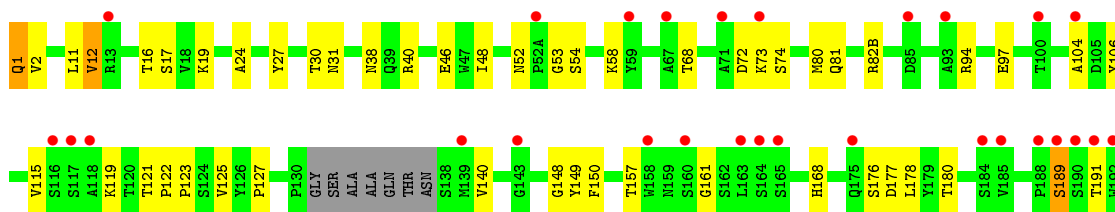
● Molecule 2: HUI-018 Fab Heavy Chain



● Molecule 2: HUI-018 Fab Heavy Chain



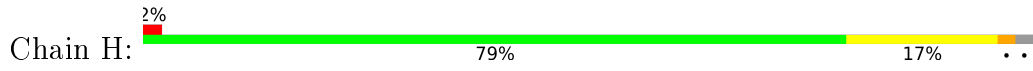
● Molecule 2: HUI-018 Fab Heavy Chain

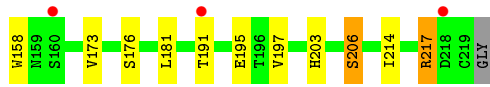


● Molecule 2: HUI-018 Fab Heavy Chain

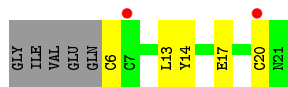


● Molecule 2: HUI-018 Fab Heavy Chain

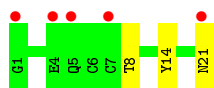
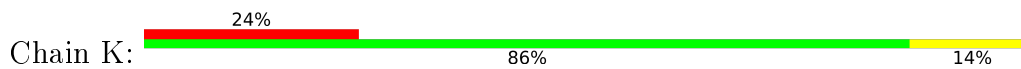




- Molecule 3: Insulin



- Molecule 3: Insulin



- Molecule 3: Insulin



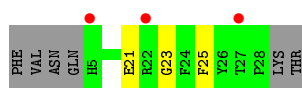
- Molecule 3: Insulin



- Molecule 4: Insulin



- Molecule 4: Insulin



- Molecule 4: Insulin





- Molecule 4: Insulin



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	50.23Å 216.61Å 101.09Å 90.00° 103.84° 90.00°	Depositor
Resolution (Å)	29.87 – 2.42 29.87 – 2.42	Depositor EDS
% Data completeness (in resolution range)	89.8 (29.87-2.42) 89.3 (29.87-2.42)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.69 (at 2.42Å)	Xtrriage
Refinement program	PHENIX 1.18_3845	Depositor
R, R_{free}	0.209 , 0.275 0.209 , 0.275	Depositor DCC
R_{free} test set	1784 reflections (2.51%)	wwPDB-VP
Wilson B-factor (Å ²)	34.7	Xtrriage
Anisotropy	0.440	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 44.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.038 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	15292	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.93% 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 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	0.50	0/1709	0.67	1/2320 (0.0%)
1	C	0.50	1/1702 (0.1%)	0.70	4/2312 (0.2%)
1	E	0.47	0/1709	0.77	5/2320 (0.2%)
1	G	0.52	0/1709	0.64	0/2320
2	B	0.46	0/1680	0.76	2/2293 (0.1%)
2	D	0.54	0/1632	0.74	3/2231 (0.1%)
2	F	0.48	1/1657 (0.1%)	0.75	3/2264 (0.1%)
2	H	0.58	0/1675	0.68	1/2288 (0.0%)
3	I	0.58	0/127	0.70	0/170
3	K	0.63	0/164	0.65	0/220
3	M	0.48	0/164	0.73	0/220
3	O	0.68	0/164	0.94	1/220 (0.5%)
4	J	0.46	0/196	0.84	1/265 (0.4%)
4	L	0.83	1/196 (0.5%)	1.17	4/265 (1.5%)
4	N	0.68	0/213	1.29	4/286 (1.4%)
4	P	0.53	0/214	0.65	0/288
All	All	0.52	3/14911 (0.0%)	0.74	29/20282 (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	E	0	1
4	N	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	L	21	GLU	CB-CG	-7.88	1.37	1.52
1	C	24	LYS	CD-CE	5.60	1.65	1.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	12	VAL	C-N	-5.38	1.21	1.34

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1	GLN	N-CA-C	-10.42	82.86	111.00
2	D	195	GLU	CA-CB-CG	9.96	135.31	113.40
4	N	22	ARG	NE-CZ-NH2	9.61	125.11	120.30
4	N	22	ARG	CG-CD-NE	-9.16	92.56	111.80
2	B	139	MET	CG-SD-CE	-8.71	86.27	100.20

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	213	GLU	Peptide
4	N	22	ARG	Sidechain

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	1672	0	1601	23	0
1	C	1665	0	1597	27	0
1	E	1672	0	1602	44	0
1	G	1672	0	1602	33	0
2	B	1641	0	1609	35	0
2	D	1593	0	1571	30	0
2	F	1618	0	1592	45	0
2	H	1636	0	1607	32	0
3	I	126	0	109	5	0
3	K	163	0	149	4	0
3	M	163	0	149	8	0
3	O	163	0	149	7	0
4	J	190	0	177	12	0
4	L	190	0	177	2	0
4	N	207	0	197	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	P	208	0	198	8	0
5	A	107	0	0	2	0
5	B	78	0	0	9	0
5	C	80	0	0	4	1
5	D	84	0	0	7	0
5	E	34	0	0	4	1
5	F	49	0	0	5	0
5	G	110	0	0	9	1
5	H	115	0	0	9	1
5	I	15	0	0	2	0
5	J	3	0	0	0	0
5	K	5	0	0	0	0
5	L	6	0	0	0	0
5	M	3	0	0	0	0
5	N	7	0	0	2	0
5	O	6	0	0	0	0
5	P	11	0	0	2	0
All	All	15292	0	14086	299	2

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

The worst 5 of 299 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)
2:F:82(B):ARG:NH2	5:F:301:HOH:O	1.83	1.08
1:G:24:LYS:NZ	5:G:302:HOH:O	1.89	0.98
1:G:195:GLU:OE1	5:G:301:HOH:O	1.82	0.97
1:A:54:ARG:HD2	1:A:58:VAL:HG23	1.49	0.94
3:I:6:CYS:N	5:I:101:HOH:O	2.02	0.91

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:E:333:HOH:O	5:G:408:HOH:O[1_556]	1.86	0.34
5:C:674:HOH:O	5:H:401:HOH:O[2_344]	2.15	0.05

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	212/214 (99%)	206 (97%)	6 (3%)	0	100	100
1	C	211/214 (99%)	204 (97%)	7 (3%)	0	100	100
1	E	212/214 (99%)	200 (94%)	12 (6%)	0	100	100
1	G	212/214 (99%)	205 (97%)	7 (3%)	0	100	100
2	B	216/224 (96%)	206 (95%)	9 (4%)	1 (0%)	29	40
2	D	209/224 (93%)	200 (96%)	9 (4%)	0	100	100
2	F	212/224 (95%)	204 (96%)	7 (3%)	1 (0%)	29	40
2	H	215/224 (96%)	208 (97%)	7 (3%)	0	100	100
3	I	14/21 (67%)	12 (86%)	1 (7%)	1 (7%)	1	0
3	K	19/21 (90%)	17 (90%)	2 (10%)	0	100	100
3	M	19/21 (90%)	17 (90%)	2 (10%)	0	100	100
3	O	19/21 (90%)	16 (84%)	3 (16%)	0	100	100
4	J	22/30 (73%)	20 (91%)	1 (4%)	1 (4%)	2	1
4	L	22/30 (73%)	21 (96%)	1 (4%)	0	100	100
4	N	24/30 (80%)	22 (92%)	2 (8%)	0	100	100
4	P	24/30 (80%)	23 (96%)	1 (4%)	0	100	100
All	All	1862/1956 (95%)	1781 (96%)	77 (4%)	4 (0%)	47	61

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	J	6	LEU
2	F	196	THR
2	B	130	PRO
3	I	20	CYS

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	191/191 (100%)	188 (98%)	3 (2%)	62	78
1	C	190/191 (100%)	183 (96%)	7 (4%)	34	51
1	E	191/191 (100%)	182 (95%)	9 (5%)	26	41
1	G	191/191 (100%)	186 (97%)	5 (3%)	46	64
2	B	187/189 (99%)	181 (97%)	6 (3%)	39	57
2	D	182/189 (96%)	178 (98%)	4 (2%)	52	69
2	F	185/189 (98%)	179 (97%)	6 (3%)	39	57
2	H	187/189 (99%)	182 (97%)	5 (3%)	44	63
3	I	16/20 (80%)	16 (100%)	0	100	100
3	K	20/20 (100%)	20 (100%)	0	100	100
3	M	20/20 (100%)	20 (100%)	0	100	100
3	O	20/20 (100%)	19 (95%)	1 (5%)	24	38
4	J	20/26 (77%)	20 (100%)	0	100	100
4	L	20/26 (77%)	20 (100%)	0	100	100
4	N	22/26 (85%)	18 (82%)	4 (18%)	1	1
4	P	22/26 (85%)	22 (100%)	0	100	100
All	All	1664/1704 (98%)	1614 (97%)	50 (3%)	41	59

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	65	SER
1	E	199	LYS
4	N	22	ARG
1	E	108	ARG
1	E	156	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	8	GLN
1	C	8	GLN
2	F	200	ASN
2	H	175	GLN
4	J	10	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	214/214 (100%)	0.45	11 (5%) 28 26	19, 33, 49, 65	0
1	C	213/214 (99%)	0.60	12 (5%) 24 22	22, 43, 57, 70	0
1	E	214/214 (100%)	1.01	29 (13%) 3 2	28, 52, 78, 85	0
1	G	214/214 (100%)	0.15	3 (1%) 75 73	16, 29, 48, 59	0
2	B	220/224 (98%)	0.72	27 (12%) 4 3	25, 41, 65, 85	0
2	D	213/224 (95%)	0.63	25 (11%) 4 4	18, 34, 72, 86	0
2	F	216/224 (96%)	1.03	34 (15%) 2 1	34, 50, 81, 94	0
2	H	219/224 (97%)	0.24	5 (2%) 60 57	16, 30, 47, 67	0
3	I	16/21 (76%)	0.83	2 (12%) 3 3	36, 44, 55, 60	0
3	K	21/21 (100%)	1.36	5 (23%) 0 0	32, 51, 74, 79	0
3	M	21/21 (100%)	1.02	5 (23%) 0 0	45, 59, 69, 72	0
3	O	21/21 (100%)	1.07	6 (28%) 0 0	29, 45, 58, 74	0
4	J	24/30 (80%)	0.81	3 (12%) 3 3	33, 47, 63, 84	0
4	L	24/30 (80%)	0.93	3 (12%) 3 3	40, 48, 75, 81	0
4	N	26/30 (86%)	1.45	8 (30%) 0 0	43, 55, 75, 88	0
4	P	26/30 (86%)	0.62	2 (7%) 13 11	34, 45, 67, 75	0
All	All	1902/1956 (97%)	0.64	180 (9%) 8 7	16, 40, 70, 94	0

The worst 5 of 180 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	168	SER	8.9
2	D	190	SER	7.7
2	B	137	ASN	7.0
1	E	203	SER	6.8
2	D	186	THR	6.6

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.