



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 28, 2021 – 10:19 AM EST

PDB ID : 6X1S  
Title : Structure of pHis Fab (SC1-1) in complex with pHis mimetic peptide  
Authors : Kalagiri, R.; Stanfield, R.L.; Wilson, I.A.; Hunter, T.  
Deposited on : 2020-05-19  
Resolution : 1.65 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 : 2.16  
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

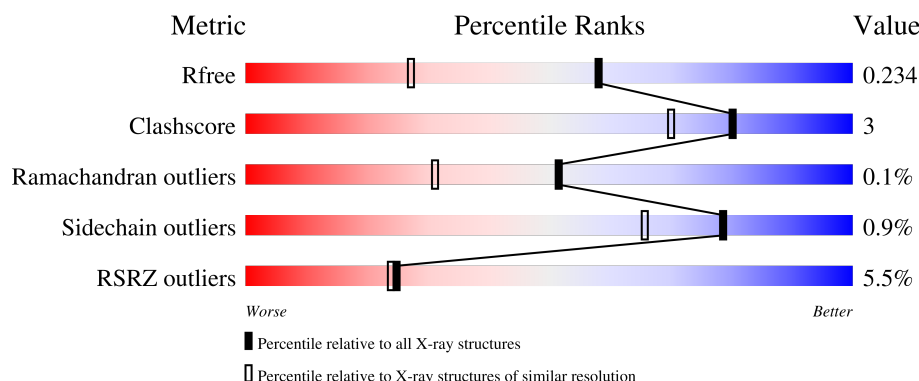
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.65 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	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)
RSRZ outliers	127900	1791 (1.66-1.66)

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  $\geq 3$ , 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	228	<div> <div>8%</div> <div>93%</div> <div>6%</div> <div>• •</div> </div>
1	C	228	<div> <div>11%</div> <div>91%</div> <div>6%</div> <div>•</div> </div>
1	E	228	<div> <div>7%</div> <div>94%</div> <div>6%</div> <div>• •</div> </div>
1	H	228	<div> <div>9%</div> <div>91%</div> <div>6%</div> <div>•</div> </div>
2	B	214	<div> <div>0%</div> <div>93%</div> <div>7%</div> <div></div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	D	214	 97%
2	F	214	 94% 6%
2	L	214	 93% 7%
3	G	9	 33% 22% 22% 11% 44%
3	I	9	 11% 22% 11% 22% 44%
3	J	9	 22% 11% 44% 44%
3	K	9	 33% 22% 33% 44%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MPD	F	301	-	-	X	-
4	MPD	L	301	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 14256 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 SC1-1 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	221	Total	C	N	O	S	0	7	0
			1644	1028	277	330	9			
1	A	220	Total	C	N	O	S	0	5	0
			1628	1018	275	327	8			
1	C	221	Total	C	N	O	S	0	7	0
			1647	1031	279	328	9			
1	E	221	Total	C	N	O	S	0	7	0
			1642	1028	276	330	8			

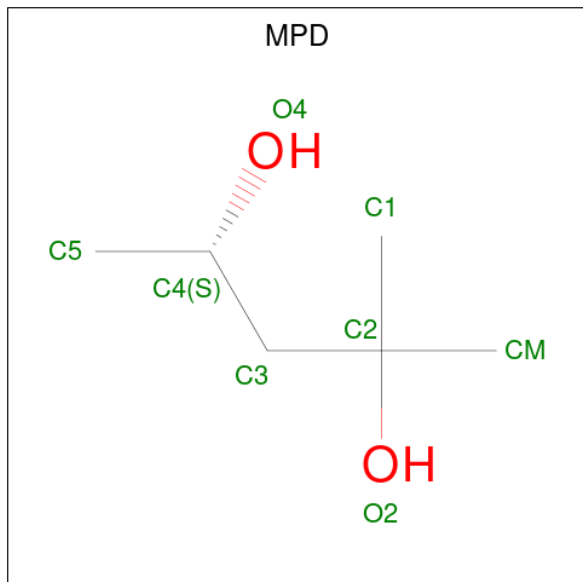
- Molecule 2 is a protein called SC1-1 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	214	Total	C	N	O	S	0	8	0
			1619	1011	259	338	11			
2	B	214	Total	C	N	O	S	0	4	0
			1606	1002	258	337	9			
2	D	214	Total	C	N	O	S	0	3	0
			1600	998	258	335	9			
2	F	214	Total	C	N	O	S	0	4	0
			1604	1001	258	335	10			

- Molecule 3 is a protein called NM23-1-pTza peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	5	Total	C	N	O	P	0	0	0
			40	22	8	9	1			
3	I	5	Total	C	N	O	P	0	0	0
			40	22	8	9	1			
3	J	5	Total	C	N	O	P	0	0	0
			40	22	8	9	1			
3	K	5	Total	C	N	O	P	0	0	0
			40	22	8	9	1			

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula:  $C_6H_{14}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	H	1	Total	C	O	0	0
			8	6	2		
4	L	1	Total	C	O	0	0
			8	6	2		
4	A	1	Total	C	O	0	0
			8	6	2		
4	A	1	Total	C	O	0	0
			8	6	2		
4	B	1	Total	C	O	0	0
			8	6	2		
4	D	1	Total	C	O	0	0
			8	6	2		
4	E	1	Total	C	O	0	0
			8	6	2		
4	F	1	Total	C	O	0	0
			8	6	2		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	H	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	E	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	H	142	Total	O	0	0
			142	142		
6	L	137	Total	O	0	0
			137	137		
6	A	138	Total	O	0	0
			138	138		
6	B	113	Total	O	0	0
			113	113		
6	C	107	Total	O	0	0
			107	107		
6	D	114	Total	O	0	0
			114	114		
6	E	124	Total	O	0	0
			124	124		
6	F	145	Total	O	0	0
			145	145		
6	G	2	Total	O	0	0
			2	2		

*Continued on next page...*

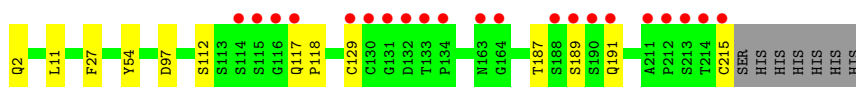
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	1	Total	O	0	0
			1	1		
6	J	2	Total	O	0	0
			2	2		
6	K	2	Total	O	0	0
			2	2		

### 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.

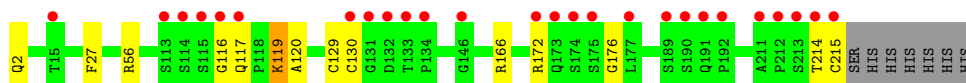
- Molecule 1: SC1-1 Heavy chain



- Molecule 1: SC1-1 Heavy chain



- Molecule 1: SC1-1 Heavy chain



- Molecule 1: SC1-1 Heavy chain

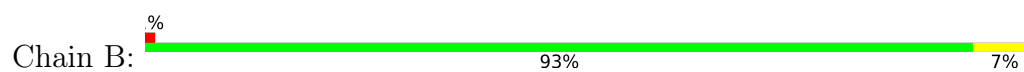


- Molecule 2: SC1-1 Light chain



- Molecule 2: SC1-1 Light chain





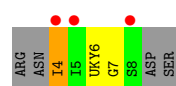
- Molecule 2: SC1-1 Light chain



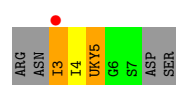
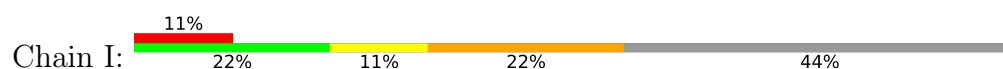
- Molecule 2: SC1-1 Light chain



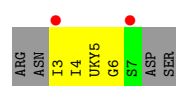
- Molecule 3: NM23-1-pTza peptide



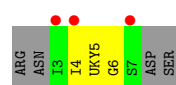
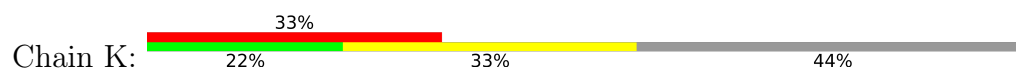
- Molecule 3: NM23-1-pTza peptide



- Molecule 3: NM23-1-pTza peptide



- Molecule 3: NM23-1-pTza peptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.79Å 72.25Å 156.75Å 90.00° 95.83° 90.00°	Depositor
Resolution (Å)	46.72 – 1.65 46.68 – 1.65	Depositor EDS
% Data completeness (in resolution range)	98.8 (46.72-1.65) 98.8 (46.68-1.65)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.17 (at 1.64Å)	Xtrriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.200 , 0.226 0.207 , 0.234	Depositor DCC
$R_{free}$ test set	11404 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.8	Xtrriage
Anisotropy	0.312	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 39.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14256	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	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 57.03 % 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 2.5133e-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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: MPD, PCA, SO4, UKY

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	0.36	0/1673	0.61	0/2292
1	C	0.35	0/1698	0.61	0/2326
1	E	0.37	0/1693	0.62	0/2320
1	H	0.37	0/1695	0.63	0/2322
2	B	0.36	0/1650	0.59	0/2260
2	D	0.37	0/1641	0.60	0/2248
2	F	0.42	0/1648	0.61	0/2258
2	L	0.41	0/1675	0.60	0/2295
3	G	0.32	0/24	0.57	0/29
3	I	0.43	0/24	0.70	0/29
3	J	0.38	0/24	0.72	0/29
3	K	0.29	0/24	0.60	0/29
All	All	0.38	0/13469	0.61	0/18437

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	1628	0	1611	5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1647	0	1638	8	0
1	E	1642	0	1631	4	0
1	H	1644	0	1629	6	0
2	B	1606	0	1558	10	0
2	D	1600	0	1552	6	0
2	F	1604	0	1558	9	0
2	L	1619	0	1580	11	0
3	G	40	0	29	1	0
3	I	40	0	29	2	0
3	J	40	0	29	4	0
3	K	40	0	29	2	0
4	A	16	0	28	2	0
4	B	8	0	14	5	0
4	D	8	0	14	5	0
4	E	8	0	14	1	0
4	F	8	0	14	6	0
4	H	8	0	14	2	0
4	L	8	0	14	6	0
5	C	5	0	0	0	0
5	E	5	0	0	0	0
5	H	5	0	0	0	0
6	A	138	0	0	0	0
6	B	113	0	0	1	0
6	C	107	0	0	1	0
6	D	114	0	0	1	0
6	E	124	0	0	0	0
6	F	145	0	0	1	0
6	G	2	0	0	0	0
6	H	142	0	0	0	0
6	I	1	0	0	0	0
6	J	2	0	0	0	0
6	K	2	0	0	0	0
6	L	137	0	0	1	0
All	All	14256	0	12985	70	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 3.

All (70) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:F:301:MPD:H52	4:F:301:MPD:HM2	1.52	0.91
4:L:301:MPD:HM2	4:L:301:MPD:H52	1.56	0.85
1:H:117:GLN:HG2	1:H:118:PRO:HD2	1.59	0.84
2:L:95(B):ASP:O	4:L:301:MPD:H51	1.80	0.81
2:B:95(B):ASP:O	4:B:301:MPD:H51	1.82	0.80
2:D:95(B):ASP:O	4:D:301:MPD:H51	1.82	0.80
4:D:301:MPD:H52	4:D:301:MPD:HM2	1.63	0.79
2:D:3:ASP:N	4:D:301:MPD:HM1	1.98	0.79
2:F:3:ASP:N	4:F:301:MPD:HM1	2.01	0.75
2:F:95(B):ASP:O	4:F:301:MPD:H51	1.87	0.75
4:B:301:MPD:HM2	4:B:301:MPD:H52	1.69	0.73
2:L:3:ASP:N	4:L:301:MPD:HM1	2.07	0.70
2:B:3:ASP:N	4:B:301:MPD:HM1	2.09	0.68
4:L:301:MPD:H53	6:L:505:HOH:O	2.00	0.61
2:L:3:ASP:N	2:L:26:SER:HG	1.99	0.61
3:K:4:ILE:O	3:K:6:GLY:N	2.34	0.60
2:B:3:ASP:N	2:B:26:SER:HG	1.99	0.60
2:F:3:ASP:N	2:F:26:SER:HG	2.00	0.60
2:D:3:ASP:N	2:D:26:SER:HG	2.01	0.59
1:E:166[B]:ARG:HG3	1:E:183:VAL:HG23	1.84	0.59
4:F:301:MPD:H53	6:F:488:HOH:O	2.03	0.59
1:E:97[B]:ASP:OD2	4:E:301:MPD:O2	2.23	0.57
1:A:173:GLN:NE2	2:B:159[B]:GLU:OE1	2.38	0.57
1:H:97[B]:ASP:OD2	4:H:301:MPD:O2	2.23	0.57
1:H:11:LEU:HD11	1:H:112:SER:HB3	1.85	0.56
3:I:3:ILE:HG22	3:I:4:ILE:HG12	1.88	0.56
1:A:166[B]:ARG:HG3	1:A:183:VAL:HG23	1.87	0.55
4:B:301:MPD:H53	6:B:490:HOH:O	2.06	0.55
2:F:6:GLN:HG3	2:F:23[B]:CYS:SG	2.45	0.55
2:L:143:VAL:HG12	2:L:197:GLN:HG3	1.87	0.55
4:L:301:MPD:CM	4:L:301:MPD:H52	2.34	0.55
1:C:56[B]:ARG:HH12	3:J:6:GLY:H	1.55	0.54
1:E:56:ARG:HD3	2:F:94:SER:O	2.08	0.54
1:C:130:CYS:N	6:C:403:HOH:O	2.41	0.53
1:A:97[B]:ASP:OD2	4:A:301:MPD:O2	2.25	0.53
1:C:56[B]:ARG:CZ	3:J:3:ILE:HG23	2.38	0.53
4:D:301:MPD:H53	6:D:489:HOH:O	2.09	0.52
2:B:37:GLN:HB2	2:B:47:LEU:HD11	1.91	0.51
2:F:6:GLN:CG	2:F:23[B]:CYS:SG	2.98	0.51
2:B:143:VAL:HG12	2:B:197:GLN:HG3	1.93	0.50
1:C:119:LYS:HD2	1:C:120:ALA:N	2.27	0.49
2:D:28:TYR:CE1	3:J:4:ILE:HG13	2.48	0.48

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:166[A]:ARG:NH1	2:D:173:ASN:OD1	2.47	0.48
2:L:6:GLN:CG	2:L:23[B]:CYS:SG	3.01	0.48
2:L:6:GLN:HG3	2:L:23[B]:CYS:SG	2.53	0.48
4:F:301:MPD:CM	4:F:301:MPD:H52	2.31	0.48
1:C:56[B]:ARG:NH1	3:J:6:GLY:H	2.12	0.47
4:L:301:MPD:HM2	4:L:301:MPD:C5	2.39	0.47
1:C:129[B]:CYS:SG	2:D:119:PRO:HG3	2.55	0.46
2:B:107:LYS:HA	2:B:140:PHE:CZ	2.51	0.46
1:E:166[B]:ARG:HG3	1:E:183:VAL:CG2	2.47	0.45
2:L:63:LYS:HB3	2:L:74[A]:THR:OG1	2.16	0.45
2:B:189:LYS:HG2	2:B:207:ASN:OD1	2.16	0.45
3:G:4:ILE:HD11	3:G:7:GLY:H	1.80	0.45
1:A:129[A]:CYS:SG	2:B:119:PRO:HG3	2.58	0.44
1:H:187:THR:HG22	1:H:189:SER:H	1.83	0.43
4:F:301:MPD:CM	4:F:301:MPD:C5	2.97	0.43
2:L:20:THR:HG22	2:L:74[A]:THR:HG22	1.99	0.43
2:B:28:TYR:CE1	3:K:4:ILE:HG13	2.53	0.43
4:H:301:MPD:H53	2:L:49:TYR:CE1	2.53	0.43
4:B:301:MPD:CM	4:B:301:MPD:H52	2.45	0.43
1:C:172:ARG:HE	1:C:176:GLY:HA2	1.84	0.43
4:D:301:MPD:HM2	4:D:301:MPD:C5	2.42	0.42
1:H:54:TYR:HD2	3:I:5:UKY:O	2.01	0.42
2:L:21:ILE:HD13	2:L:86:TYR:HB2	2.02	0.42
1:A:99:GLY:HA3	4:A:302:MPD:H4	2.02	0.41
2:F:158:ILE:HA	2:F:177:THR:O	2.21	0.41
2:F:5[A]:THR:OG1	2:F:24:GLN:HB3	2.19	0.41
1:H:129[A]:CYS:SG	2:L:119:PRO:HG3	2.60	0.41
2:F:111:VAL:HG13	2:F:139:TYR:HA	2.02	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	223/228 (98%)	219 (98%)	4 (2%)	0	100	100
1	C	226/228 (99%)	218 (96%)	7 (3%)	1 (0%)	34	16
1	E	226/228 (99%)	221 (98%)	5 (2%)	0	100	100
1	H	226/228 (99%)	221 (98%)	5 (2%)	0	100	100
2	B	216/214 (101%)	212 (98%)	4 (2%)	0	100	100
2	D	215/214 (100%)	210 (98%)	5 (2%)	0	100	100
2	F	216/214 (101%)	211 (98%)	5 (2%)	0	100	100
2	L	220/214 (103%)	216 (98%)	4 (2%)	0	100	100
3	G	2/9 (22%)	2 (100%)	0	0	100	100
3	I	2/9 (22%)	1 (50%)	1 (50%)	0	100	100
3	J	2/9 (22%)	2 (100%)	0	0	100	100
3	K	2/9 (22%)	2 (100%)	0	0	100	100
All	All	1776/1804 (98%)	1735 (98%)	40 (2%)	1 (0%)	51	31

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	116	GLY

### 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	188/191 (98%)	186 (99%)	2 (1%)	73	57
1	C	191/191 (100%)	186 (97%)	5 (3%)	46	21
1	E	191/191 (100%)	190 (100%)	1 (0%)	88	81
1	H	191/191 (100%)	188 (98%)	3 (2%)	62	41
2	B	190/186 (102%)	190 (100%)	0	100	100
2	D	189/186 (102%)	189 (100%)	0	100	100
2	F	190/186 (102%)	190 (100%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	L	194/186 (104%)	194 (100%)	0	100	100
3	G	3/7 (43%)	2 (67%)	1 (33%)	0	0
3	I	3/7 (43%)	2 (67%)	1 (33%)	0	0
3	J	3/7 (43%)	3 (100%)	0	100	100
3	K	3/7 (43%)	3 (100%)	0	100	100
All	All	1536/1536 (100%)	1523 (99%)	13 (1%)	78	70

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	27	PHE
1	H	191	GLN
1	H	215	CYS
1	A	27	PHE
1	A	61	SER
1	C	27	PHE
1	C	117	GLN
1	C	119	LYS
1	C	214	THR
1	C	215	CYS
1	E	27	PHE
3	G	4	ILE
3	I	3	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	L	188	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

8 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
3	UKY	J	5	3	9,14,15	1.09	0	9,20,22	2.19	2 (22%)
3	UKY	K	5	3	9,14,15	1.09	0	9,20,22	2.18	1 (11%)
1	PCA	A	2	1	7,8,9	1.76	1 (14%)	9,10,12	2.07	4 (44%)
3	UKY	G	6	3	9,14,15	1.17	1 (11%)	9,20,22	1.76	1 (11%)
1	PCA	E	2	1	7,8,9	1.84	1 (14%)	9,10,12	2.18	4 (44%)
1	PCA	C	2	1	7,8,9	1.73	1 (14%)	9,10,12	2.16	5 (55%)
1	PCA	H	2	1	7,8,9	1.74	1 (14%)	9,10,12	2.04	4 (44%)
3	UKY	I	5	3	9,14,15	1.09	0	9,20,22	2.49	3 (33%)

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
3	UKY	J	5	3	-	0/0/12/14	0/1/1/1
3	UKY	K	5	3	-	0/0/12/14	0/1/1/1
1	PCA	A	2	1	-	0/0/11/13	0/1/1/1
3	UKY	G	6	3	-	0/0/12/14	0/1/1/1
1	PCA	E	2	1	-	0/0/11/13	0/1/1/1
1	PCA	C	2	1	-	0/0/11/13	0/1/1/1
1	PCA	H	2	1	-	0/0/11/13	0/1/1/1
3	UKY	I	5	3	-	0/0/12/14	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	2	PCA	CD-N	4.70	1.47	1.34
1	A	2	PCA	CD-N	4.49	1.46	1.34
1	C	2	PCA	CD-N	4.48	1.46	1.34
1	H	2	PCA	CD-N	4.45	1.46	1.34
3	G	6	UKY	NE2-ND2	-2.34	1.31	1.34

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	5	UKY	CB-NG-ND2	6.04	125.53	116.67
3	K	5	UKY	CB-NG-ND2	5.52	124.77	116.67
3	J	5	UKY	CB-NG-ND2	5.34	124.51	116.67
3	G	6	UKY	CB-NG-ND2	4.69	123.55	116.67
1	E	2	PCA	OE-CD-CG	-3.21	121.16	126.76
1	E	2	PCA	CA-N-CD	-3.21	102.59	113.58
1	A	2	PCA	CA-N-CD	-3.18	102.69	113.58
1	A	2	PCA	CB-CA-N	3.15	112.33	103.30
1	H	2	PCA	CA-N-CD	-3.13	102.86	113.58
1	C	2	PCA	CA-N-CD	-3.11	102.93	113.58
1	E	2	PCA	CB-CA-N	3.04	112.04	103.30
1	H	2	PCA	CB-CA-N	2.98	111.86	103.30
1	C	2	PCA	CB-CA-N	2.98	111.84	103.30
3	I	5	UKY	O2-P-CD1	-2.75	104.39	112.56
1	A	2	PCA	OE-CD-CG	-2.74	121.97	126.76
1	C	2	PCA	OE-CD-CG	-2.72	122.02	126.76
1	H	2	PCA	OE-CD-CG	-2.72	122.02	126.76
1	E	2	PCA	CG-CD-N	2.66	115.29	108.39
1	H	2	PCA	CG-CD-N	2.63	115.20	108.39
1	A	2	PCA	CG-CD-N	2.61	115.16	108.39
1	C	2	PCA	CG-CD-N	2.57	115.05	108.39
3	I	5	UKY	CB-NG-CD1	-2.52	125.77	128.81
3	J	5	UKY	O2-P-CD1	-2.18	106.09	112.56
1	C	2	PCA	CB-CA-C	-2.04	109.89	112.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	5	UKY	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

11 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	MPD	D	301	-	7,7,7	0.40	0	9,10,10	0.98	0
5	SO4	C	301	-	4,4,4	0.18	0	6,6,6	0.13	0
4	MPD	H	301	-	7,7,7	0.29	0	9,10,10	0.52	0
4	MPD	A	301	-	7,7,7	0.33	0	9,10,10	0.39	0
4	MPD	L	301	-	7,7,7	0.60	0	9,10,10	1.05	0
4	MPD	E	301	-	7,7,7	0.28	0	9,10,10	0.26	0
4	MPD	A	302	-	7,7,7	0.37	0	9,10,10	0.63	0
5	SO4	H	302	-	4,4,4	0.14	0	6,6,6	0.21	0
4	MPD	B	301	-	7,7,7	0.26	0	9,10,10	0.88	1 (11%)
5	SO4	E	302	-	4,4,4	0.12	0	6,6,6	0.12	0
4	MPD	F	301	-	7,7,7	0.51	0	9,10,10	1.02	0

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	MPD	D	301	-	-	1/5/5/5	-
4	MPD	H	301	-	-	5/5/5/5	-
4	MPD	A	301	-	-	1/5/5/5	-
4	MPD	L	301	-	-	1/5/5/5	-
4	MPD	E	301	-	-	0/5/5/5	-
4	MPD	A	302	-	-	3/5/5/5	-
4	MPD	B	301	-	-	1/5/5/5	-
4	MPD	F	301	-	-	1/5/5/5	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	301	MPD	O2-C2-CM	-2.06	101.47	108.08

There are no chirality outliers.

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	302	MPD	O2-C2-C3-C4
4	A	302	MPD	CM-C2-C3-C4
4	H	301	MPD	O2-C2-C3-C4
4	D	301	MPD	C2-C3-C4-C5
4	L	301	MPD	C2-C3-C4-C5
4	B	301	MPD	C2-C3-C4-C5
4	F	301	MPD	C2-C3-C4-C5
4	H	301	MPD	C2-C3-C4-O4
4	H	301	MPD	C1-C2-C3-C4
4	H	301	MPD	CM-C2-C3-C4
4	A	301	MPD	CM-C2-C3-C4
4	H	301	MPD	C2-C3-C4-C5
4	A	302	MPD	C2-C3-C4-C5

There are no ring outliers.

8 monomers are involved in 27 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	301	MPD	5	0
4	H	301	MPD	2	0
4	A	301	MPD	1	0
4	L	301	MPD	6	0
4	E	301	MPD	1	0
4	A	302	MPD	1	0
4	B	301	MPD	5	0
4	F	301	MPD	6	0

## 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 ⓘ

### 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	219/228 (96%)	0.30	18 (8%) 11 11	14, 22, 56, 86	0
1	C	220/228 (96%)	0.72	26 (11%) 4 3	14, 27, 60, 90	0
1	E	220/228 (96%)	0.46	16 (7%) 15 14	13, 23, 59, 99	0
1	H	220/228 (96%)	0.47	21 (9%) 8 7	13, 20, 52, 103	0
2	B	214/214 (100%)	-0.04	3 (1%) 75 79	17, 26, 38, 47	0
2	D	214/214 (100%)	-0.11	3 (1%) 75 79	16, 26, 35, 48	0
2	F	214/214 (100%)	-0.21	0 100 100	14, 22, 32, 45	0
2	L	214/214 (100%)	-0.19	1 (0%) 91 92	14, 23, 32, 50	0
3	G	4/9 (44%)	2.17	3 (75%) 0 0	27, 32, 43, 43	0
3	I	4/9 (44%)	1.72	1 (25%) 0 0	26, 32, 40, 44	0
3	J	4/9 (44%)	2.07	2 (50%) 0 0	27, 29, 35, 41	0
3	K	4/9 (44%)	2.15	3 (75%) 0 0	33, 37, 47, 47	0
All	All	1751/1804 (97%)	0.20	97 (5%) 25 24	13, 23, 46, 103	0

All (97) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	215	CYS	14.2
1	E	215	CYS	13.7
1	C	215	CYS	12.9
1	E	214	THR	10.1
1	C	214	THR	9.9
1	C	116	GLY	9.1
1	A	214	THR	8.8
1	H	132	ASP	8.5
1	E	132	ASP	8.3
1	E	213	SER	7.8
1	H	214	THR	7.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	114	SER	7.7
1	H	133	THR	7.1
1	C	115	SER	7.1
1	C	134	PRO	6.6
1	H	114	SER	6.4
1	E	131	GLY	6.4
1	C	175	SER	6.1
1	C	113	SER	6.0
1	C	130	CYS	6.0
1	E	133	THR	5.9
1	H	134	PRO	5.8
1	E	134	PRO	5.7
1	A	132	ASP	5.6
1	C	132	ASP	5.4
1	E	211	ALA	5.4
1	C	189	SER	5.2
1	H	191	GLN	5.1
1	A	188	SER	5.1
1	A	134	PRO	5.1
1	C	213	SER	5.0
1	A	130	CYS	4.9
3	I	3	ILE	4.7
1	A	114	SER	4.7
1	H	130	CYS	4.7
1	A	115	SER	4.5
2	D	211	CYS	4.4
1	E	116	GLY	4.3
1	E	190	SER	4.3
3	G	4	ILE	4.3
3	J	3	ILE	4.3
1	C	174	SER	4.2
1	H	213	SER	4.2
1	A	133	THR	4.2
1	A	189	SER	4.2
1	E	191	GLN	4.0
1	A	191	GLN	3.9
1	H	116	GLY	3.9
1	H	188	SER	3.8
2	L	211	CYS	3.6
1	E	130	CYS	3.6
1	H	211	ALA	3.6
1	C	191	GLN	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	115	SER	3.6
1	H	115	SER	3.5
1	C	133	THR	3.5
2	B	211	CYS	3.5
1	E	212	PRO	3.3
3	K	3	ILE	3.3
1	A	213	SER	3.2
1	A	131	GLY	3.1
1	H	190	SER	3.0
1	C	211	ALA	2.9
1	H	189	SER	2.8
1	C	190	SER	2.8
1	H	129[A]	CYS	2.8
1	C	131	GLY	2.7
2	B	167	SER	2.7
1	C	173	GLN	2.7
1	C	117	GLN	2.6
1	E	129	CYS	2.6
1	A	116	GLY	2.6
1	A	211	ALA	2.5
1	H	131	GLY	2.5
1	E	189	SER	2.5
3	J	7	SER	2.5
2	D	68	GLY	2.5
3	K	4	ILE	2.4
3	K	7	SER	2.4
1	C	15	THR	2.4
1	A	190	SER	2.4
3	G	5	ILE	2.3
1	A	192	PRO	2.3
1	A	163	ASN	2.2
1	C	172	ARG	2.2
2	B	68	GLY	2.2
1	H	117	GLN	2.2
1	C	192	PRO	2.2
1	C	212	PRO	2.2
1	H	163	ASN	2.2
3	G	8	SER	2.1
1	A	212	PRO	2.1
1	H	212	PRO	2.1
1	C	146	GLY	2.1
1	C	177	LEU	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	H	164	GLY	2.0
2	D	183	THR	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
1	PCA	A	2	8/9	0.93	0.10	20,23,29,36	0
1	PCA	C	2	8/9	0.93	0.09	22,27,32,37	0
1	PCA	E	2	8/9	0.94	0.07	21,24,33,38	0
1	PCA	H	2	8/9	0.94	0.10	22,24,35,40	0
3	UKY	J	5	14/15	0.97	0.09	15,17,24,29	0
3	UKY	G	6	14/15	0.98	0.08	13,17,25,33	0
3	UKY	K	5	14/15	0.98	0.08	15,18,29,34	0
3	UKY	I	5	14/15	0.98	0.09	13,16,24,28	0

## 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	MPD	A	302	8/8	0.71	0.36	35,37,42,42	0
4	MPD	B	301	8/8	0.82	0.11	17,22,28,33	0
4	MPD	E	301	8/8	0.83	0.21	24,36,41,42	0
4	MPD	F	301	8/8	0.85	0.12	10,15,23,27	0
4	MPD	A	301	8/8	0.86	0.20	27,38,44,45	0
4	MPD	H	301	8/8	0.88	0.15	25,34,41,41	0
4	MPD	D	301	8/8	0.89	0.10	12,17,28,29	0
4	MPD	L	301	8/8	0.89	0.12	10,16,23,28	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	SO4	E	302	5/5	0.90	0.20	49,54,67,73	0
5	SO4	C	301	5/5	0.95	0.09	44,45,48,50	0
5	SO4	H	302	5/5	0.96	0.11	40,43,47,48	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.