



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 05:13 am BST

PDB ID : 5NLT  
Title : CvAA9A  
Authors : Frandsen, K.E.H.; Poulsen, J.-C.N.; Tandrup, T.; Schnorr, K.; Tovborg, M.;  
Lo Leggio, L.  
Deposited on : 2017-04-04  
Resolution : 1.90 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 : 2.11  
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

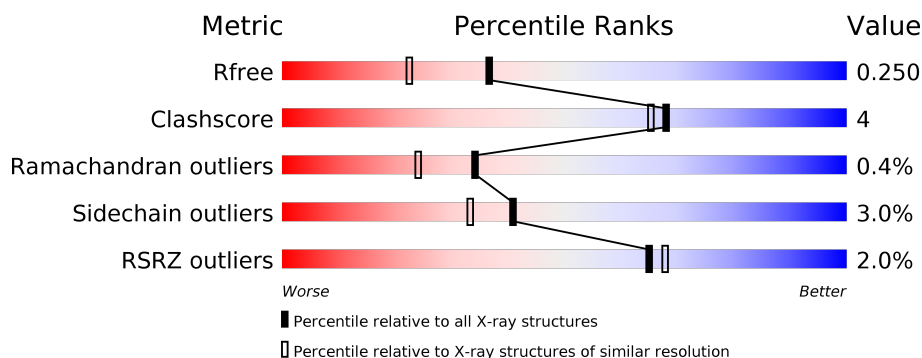
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	252	<div> <div>2%</div> <div> <div></div> <div>81%</div> <div>8%</div> <div>11%</div> </div> </div>
1	B	252	<div> <div>%</div> <div> <div></div> <div>82%</div> <div>7%</div> <div>10%</div> </div> </div>
1	C	252	<div> <div>2%</div> <div> <div></div> <div>78%</div> <div>10%</div> <div>10%</div> </div> </div>
1	D	252	<div> <div>2%</div> <div> <div></div> <div>79%</div> <div>9%</div> <div>11%</div> </div> </div>
1	E	252	<div> <div>2%</div> <div> <div></div> <div>79%</div> <div>10%</div> <div>10%</div> </div> </div>
1	F	252	<div> <div>3%</div> <div> <div></div> <div>80%</div> <div>7%</div> <div>11%</div> </div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 11754 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 CvAA9A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	224	Total	C	N	O	S	0	3	0
			1762	1128	290	337	7			
1	B	227	Total	C	N	O	S	0	3	0
			1791	1142	296	346	7			
1	C	226	Total	C	N	O	S	0	8	0
			1809	1151	299	352	7			
1	D	224	Total	C	N	O	S	0	2	0
			1755	1122	290	336	7			
1	E	226	Total	C	N	O	S	0	4	0
			1791	1144	298	342	7			
1	F	224	Total	C	N	O	S	0	2	0
			1759	1124	293	335	7			

- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total	Cu	0	0
			1	1		
2	E	1	Total	Cu	0	0
			1	1		
2	B	1	Total	Cu	0	0
			1	1		
2	C	1	Total	Cu	0	0
			1	1		
2	A	1	Total	Cu	0	0
			1	1		
2	F	1	Total	Cu	0	0
			1	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	198	Total	O	0	3
			201	201		
4	B	229	Total	O	0	4
			230	230		
4	C	173	Total	O	0	6
			174	174		
4	D	139	Total	O	0	3
			140	140		

*Continued on next page...*

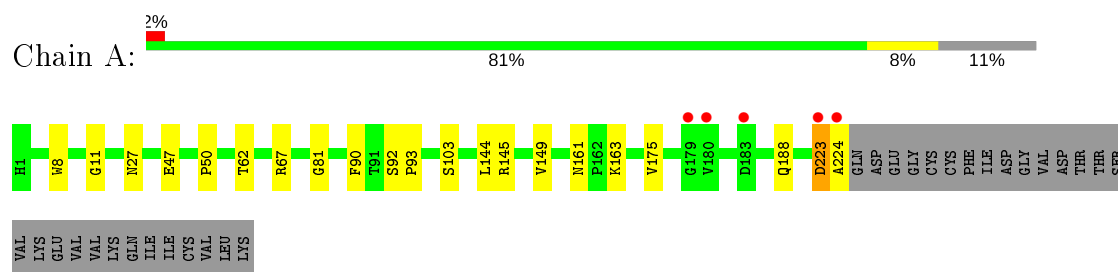
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	E	123	Total 123	O 123	0	0
4	F	169	Total 173	O 173	0	7

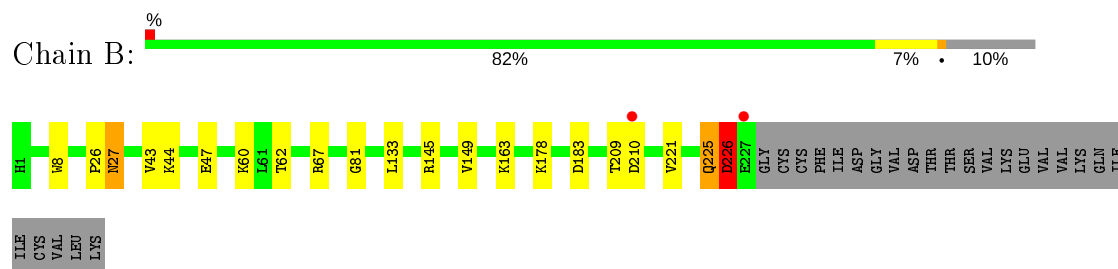
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

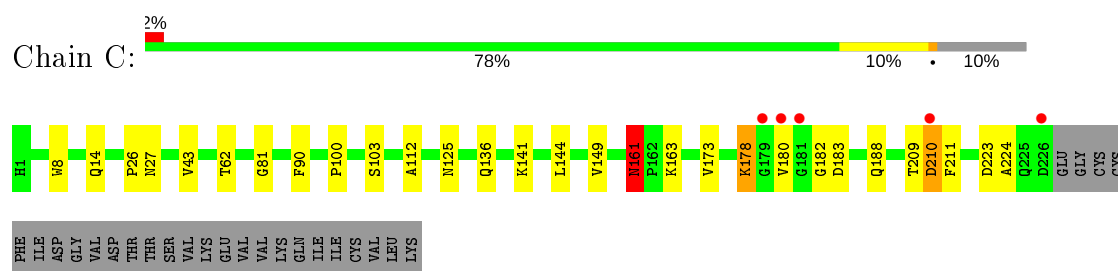
#### • Molecule 1: CvAA9A



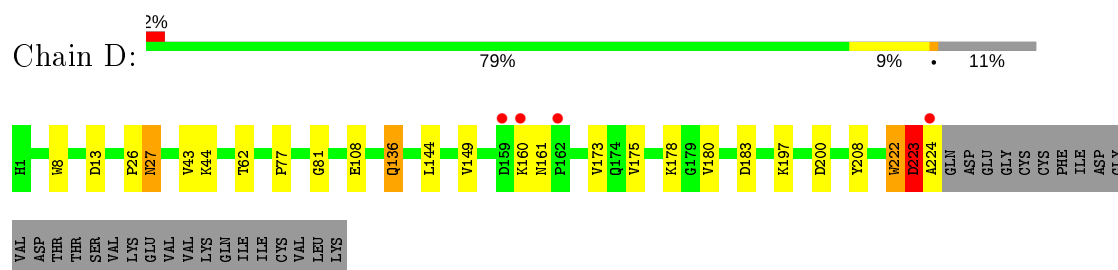
#### • Molecule 1: CvAA9A



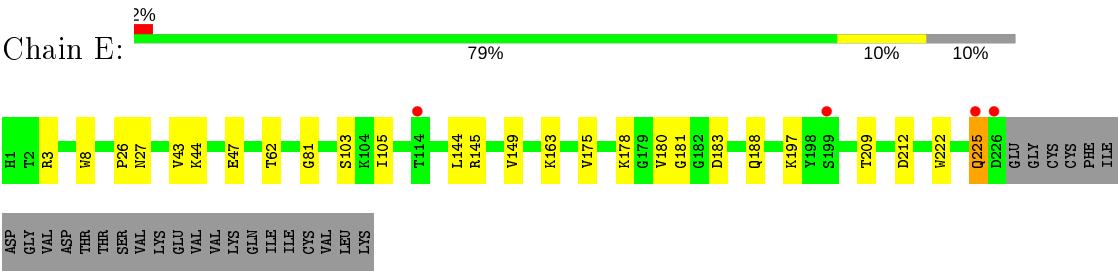
#### • Molecule 1: CvAA9A



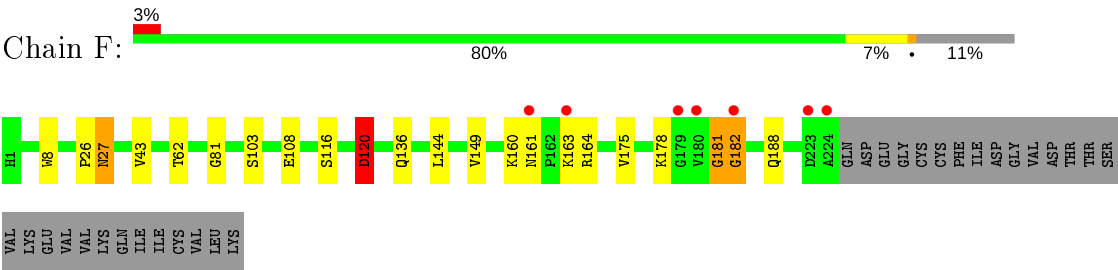
#### • Molecule 1: CvAA9A



● Molecule 1: CvAA9A



● Molecule 1: CvAA9A



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.00Å 59.42Å 115.45Å 102.67° 98.89° 89.54°	Depositor
Resolution (Å)	50.00 – 1.90 46.42 – 1.90	Depositor EDS
% Data completeness (in resolution range)	89.9 (50.00-1.90) 89.9 (46.42-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.04 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.198 , 0.244 0.204 , 0.250	Depositor DCC
$R_{free}$ test set	4376 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.2	Xtriage
Anisotropy	0.447	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 42.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.022 for -h,k,-k-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	11754	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 19.39% 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: HIC, SO4, CU

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	0.70	0/1817	0.85	2/2480 (0.1%)
1	B	0.75	2/1840 (0.1%)	0.86	5/2512 (0.2%)
1	C	0.71	0/1864	0.93	7/2544 (0.3%)
1	D	0.71	0/1807	0.87	6/2468 (0.2%)
1	E	0.70	1/1844 (0.1%)	0.84	3/2516 (0.1%)
1	F	0.71	1/1811 (0.1%)	0.94	2/2472 (0.1%)
All	All	0.71	4/10983 (0.0%)	0.88	25/14992 (0.2%)

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	C	0	2
1	F	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	120	ASP	CB-CG	-6.39	1.38	1.51
1	B	226[A]	ASP	CB-CG	5.35	1.62	1.51
1	B	226[B]	ASP	CB-CG	5.35	1.62	1.51
1	E	47	GLU	CD-OE1	-5.26	1.19	1.25

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	120	ASP	CB-CG-OD1	-16.65	103.31	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	223	ASP	CB-CG-OD2	-7.00	112.00	118.30
1	B	183	ASP	CB-CG-OD2	-6.88	112.11	118.30
1	D	222	TRP	O-C-N	-6.84	111.75	122.70
1	F	120	ASP	CB-CG-OD2	6.39	124.05	118.30
1	D	222	TRP	CA-C-N	6.37	131.22	117.20
1	C	178	LYS	CA-CB-CG	6.21	127.06	113.40
1	D	183	ASP	CB-CG-OD2	-6.20	112.72	118.30
1	C	178	LYS	CD-CE-NZ	5.99	125.48	111.70
1	B	145	ARG	NE-CZ-NH1	5.92	123.26	120.30
1	A	67	ARG	NE-CZ-NH1	-5.89	117.35	120.30
1	A	145	ARG	NE-CZ-NH1	5.85	123.22	120.30
1	E	145	ARG	NE-CZ-NH1	5.72	123.16	120.30
1	D	44	LYS	CD-CE-NZ	5.54	124.44	111.70
1	B	133	LEU	CA-CB-CG	5.46	127.86	115.30
1	C	161[A]	ASN	N-CA-CB	5.45	120.41	110.60
1	C	161[B]	ASN	N-CA-CB	5.45	120.41	110.60
1	B	145	ARG	NE-CZ-NH2	-5.33	117.64	120.30
1	C	161[A]	ASN	CB-CA-C	-5.33	99.75	110.40
1	C	161[B]	ASN	CB-CA-C	-5.33	99.75	110.40
1	E	47	GLU	CG-CD-OE2	5.26	128.81	118.30
1	B	67	ARG	NE-CZ-NH2	-5.25	117.67	120.30
1	D	223	ASP	N-CA-C	-5.24	96.86	111.00
1	D	222	TRP	C-N-CA	5.18	134.64	121.70
1	E	3	ARG	NE-CZ-NH1	-5.04	117.78	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	180	VAL	Peptide
1	F	181	GLY	Peptide

## 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	1762	0	1672	10	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1791	0	1685	9	0
1	C	1809	0	1701	15	0
1	D	1755	0	1662	20	2
1	E	1791	0	1697	14	0
1	F	1759	0	1668	17	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	B	5	0	0	0	0
3	D	10	0	0	0	0
3	E	15	0	0	1	0
3	F	10	0	0	0	0
4	A	201	0	0	0	0
4	B	230	0	0	3	0
4	C	174	0	0	5	0
4	D	140	0	0	4	0
4	E	123	0	0	0	0
4	F	173	0	0	2	0
All	All	11754	0	10085	79	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 4.

All (79) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:164[B]:ARG:HH11	1:F:164[B]:ARG:HG3	0.99	1.10
1:D:222:TRP:CD1	1:D:223:ASP:HB3	2.05	0.92
1:F:181:GLY:HA3	1:F:182:GLY:O	1.70	0.91
1:F:164[B]:ARG:NH1	1:F:164[B]:ARG:HG3	1.73	0.90
1:C:210[B]:ASP:O	1:C:210[B]:ASP:OD1	1.90	0.88
1:D:144:LEU:HD12	1:D:175:VAL:HG21	1.61	0.82
1:E:144[A]:LEU:HD12	1:E:175:VAL:HG21	1.60	0.82
1:B:60:LYS:HE3	4:B:426:HOH:O	1.78	0.82
1:F:144:LEU:HD12	1:F:175:VAL:HG21	1.62	0.81
1:A:144[A]:LEU:HD12	1:A:175:VAL:HG21	1.61	0.80
1:C:210[B]:ASP:CG	1:C:210[B]:ASP:O	2.29	0.71
1:D:180:VAL:HB	1:E:44:LYS:HG2	1.72	0.70

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:125:ASN:ND2	4:C:401:HOH:O	2.03	0.68
1:D:180:VAL:HB	1:E:44:LYS:CG	2.28	0.63
1:F:116:SER:HB2	1:F:120:ASP:OD1	1.99	0.62
1:D:136:GLN:HE21	1:D:224:ALA:HB1	1.64	0.62
1:E:212:ASP:HB2	3:E:303:SO4:O2	2.01	0.59
1:D:222:TRP:HA	1:D:223:ASP:HB2	1.83	0.59
1:D:200:ASP:OD1	4:D:402:HOH:O	2.17	0.57
1:F:164[B]:ARG:NH1	1:F:164[B]:ARG:CG	2.55	0.57
1:F:161[B]:ASN:HB3	4:F:548[B]:HOH:O	2.05	0.56
1:A:11:GLY:O	1:C:100:PRO:HD3	2.06	0.56
1:F:160:LYS:HB2	1:F:161[B]:ASN:HD22	1.71	0.55
1:C:90:PHE:HB3	1:C:141:LYS:HB2	1.88	0.55
1:F:164[B]:ARG:CG	1:F:164[B]:ARG:HH11	1.89	0.55
1:C:209[B]:THR:O	1:C:211:PHE:N	2.40	0.55
1:E:105:ILE:O	1:E:225:GLN:OE1	2.24	0.55
1:D:136:GLN:NE2	1:D:224:ALA:HB1	2.22	0.55
1:D:160:LYS:HB2	1:D:161[B]:ASN:ND2	2.22	0.54
1:F:181:GLY:CA	1:F:182:GLY:O	2.52	0.54
1:A:93:PRO:HD3	1:D:77:PRO:O	2.08	0.53
1:B:221:VAL:CG1	1:B:225:GLN:HG2	2.39	0.53
1:D:161[B]:ASN:ND2	4:D:408[B]:HOH:O	2.41	0.52
1:C:182[A]:GLY:N	4:C:408[A]:HOH:O	2.37	0.52
1:B:210:ASP:OD1	4:B:401:HOH:O	2.19	0.51
1:D:27:ASN:ND2	4:D:409:HOH:O	2.44	0.51
1:B:8:TRP:HB2	1:B:62:THR:HB	1.93	0.50
1:D:222:TRP:HD1	1:D:223:ASP:HB3	1.71	0.50
1:A:8:TRP:HB2	1:A:62:THR:HB	1.94	0.50
1:E:180:VAL:N	1:E:181:GLY:HA2	2.26	0.50
1:C:136[A]:GLN:NE2	1:C:224:ALA:HB3	2.28	0.49
1:E:8:TRP:HB2	1:E:62:THR:HB	1.94	0.49
1:C:8:TRP:HB2	1:C:62:THR:HB	1.94	0.49
1:D:26:PRO:HD2	1:D:43:VAL:CG2	2.43	0.49
1:F:26:PRO:HD2	1:F:43:VAL:CG2	2.43	0.48
1:D:8:TRP:HB2	1:D:62:THR:HB	1.95	0.48
1:C:81:GLY:HA3	1:C:149:VAL:O	2.14	0.48
1:E:222:TRP:O	1:E:225:GLN:HB2	2.14	0.48
1:B:26:PRO:HD2	1:B:43:VAL:CG2	2.45	0.47
1:C:161[B]:ASN:ND2	4:C:404:HOH:O	2.33	0.47
1:C:26:PRO:HD2	1:C:43:VAL:CG2	2.44	0.47
1:C:112:ALA:O	4:C:402:HOH:O	2.20	0.47
1:E:81:GLY:HA3	1:E:149:VAL:O	2.15	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:164[A]:ARG:HG2	4:F:524[A]:HOH:O	2.15	0.47
1:B:44:LYS:NZ	1:B:47[A]:GLU:OE2	2.48	0.46
1:E:26[A]:PRO:HD2	1:E:43:VAL:CG2	2.46	0.46
1:F:8:TRP:HB2	1:F:62:THR:HB	1.98	0.46
1:E:26[B]:PRO:HD2	1:E:43:VAL:CG2	2.46	0.46
1:D:81:GLY:HA3	1:D:149:VAL:O	2.16	0.45
1:A:223:ASP:O	1:A:224:ALA:HB2	2.16	0.45
1:A:50:PRO:HA	1:D:208:TYR:CE1	2.52	0.45
1:F:144:LEU:CD1	1:F:175:VAL:HG21	2.42	0.44
1:D:144:LEU:CD1	1:D:175:VAL:HG21	2.42	0.44
1:E:105:ILE:O	1:E:225:GLN:CD	2.55	0.44
1:F:81:GLY:HA3	1:F:149:VAL:O	2.18	0.44
1:B:81:GLY:HA3	1:B:149:VAL:O	2.18	0.44
1:D:197:LYS:N	4:D:403:HOH:O	2.24	0.43
1:A:90:PHE:HD2	1:D:77:PRO:HG2	1.83	0.43
1:E:26[B]:PRO:O	1:E:27[B]:ASN:CG	2.57	0.43
1:A:81:GLY:HA3	1:A:149:VAL:O	2.19	0.43
1:F:27:ASN:C	1:F:27:ASN:HD22	2.23	0.42
1:A:144[A]:LEU:CD1	1:A:175:VAL:HG21	2.42	0.42
1:C:14:GLN:NE2	4:C:413:HOH:O	2.44	0.42
1:C:103:SER:HA	1:C:188:GLN:HE21	1.85	0.41
1:A:103:SER:HA	1:A:188:GLN:HE21	1.86	0.41
1:E:103:SER:HA	1:E:188:GLN:HE21	1.86	0.41
1:B:60:LYS:CE	4:B:426:HOH:O	2.51	0.40
1:F:103:SER:HA	1:F:188:GLN:HE21	1.87	0.40
1:B:27:ASN:HD22	1:B:27:ASN:C	2.24	0.40

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)
1:D:13:ASP:O	1:D:160:LYS:NZ[1_455]	1.76	0.44
1:D:13:ASP:O	1:D:160:LYS:CE[1_455]	2.09	0.11

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	225/252 (89%)	212 (94%)	12 (5%)	1 (0%)	34	24
1	B	228/252 (90%)	215 (94%)	11 (5%)	2 (1%)	17	7
1	C	232/252 (92%)	217 (94%)	13 (6%)	2 (1%)	17	7
1	D	224/252 (89%)	211 (94%)	12 (5%)	1 (0%)	34	24
1	E	228/252 (90%)	209 (92%)	18 (8%)	1 (0%)	34	24
1	F	224/252 (89%)	211 (94%)	12 (5%)	1 (0%)	34	24
All	All	1361/1512 (90%)	1275 (94%)	78 (6%)	8 (1%)	34	15

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	210[A]	ASP
1	C	210[B]	ASP
1	D	223	ASP
1	F	182	GLY
1	B	226[A]	ASP
1	B	226[B]	ASP
1	A	223	ASP
1	E	225	GLN

### 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	189/212 (89%)	184 (97%)	5 (3%)	46	39
1	B	191/212 (90%)	184 (96%)	7 (4%)	34	25
1	C	194/212 (92%)	186 (96%)	8 (4%)	30	21
1	D	188/212 (89%)	183 (97%)	5 (3%)	44	38
1	E	192/212 (91%)	187 (97%)	5 (3%)	46	39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	188/212 (89%)	182 (97%)	6 (3%)	39	30
All	All	1142/1272 (90%)	1106 (97%)	36 (3%)	41	30

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	ASN
1	A	47	GLU
1	A	92	SER
1	A	161	ASN
1	A	163	LYS
1	B	27	ASN
1	B	163	LYS
1	B	178	LYS
1	B	209	THR
1	B	225	GLN
1	B	226[A]	ASP
1	B	226[B]	ASP
1	C	27	ASN
1	C	144	LEU
1	C	161[A]	ASN
1	C	161[B]	ASN
1	C	163	LYS
1	C	173	VAL
1	C	178	LYS
1	C	183	ASP
1	D	27	ASN
1	D	108	GLU
1	D	136	GLN
1	D	173	VAL
1	D	178	LYS
1	E	163	LYS
1	E	178	LYS
1	E	183	ASP
1	E	197	LYS
1	E	209	THR
1	F	27	ASN
1	F	108	GLU
1	F	120	ASP
1	F	136	GLN
1	F	163	LYS
1	F	178	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (24) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	14	GLN
1	A	27	ASN
1	A	161	ASN
1	A	188	GLN
1	A	193	ASN
1	B	27	ASN
1	B	188	GLN
1	B	193	ASN
1	C	14	GLN
1	C	27	ASN
1	C	51	GLN
1	C	188	GLN
1	C	193	ASN
1	D	27	ASN
1	D	51	GLN
1	D	136	GLN
1	D	188	GLN
1	D	193	ASN
1	E	188	GLN
1	E	193	ASN
1	F	27	ASN
1	F	136	GLN
1	F	188	GLN
1	F	193	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

6 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
1	HIC	C	1	1,2	8,11,12	1.03	0	6,14,16	0.85	0
1	HIC	A	1	1,2	8,11,12	1.17	1 (12%)	6,14,16	0.64	0
1	HIC	D	1	1,2	8,11,12	1.46	1 (12%)	6,14,16	0.56	0
1	HIC	E	1	1,2	8,11,12	1.11	0	6,14,16	0.71	0
1	HIC	B	1	1,2	8,11,12	1.14	1 (12%)	6,14,16	0.55	0
1	HIC	F	1	1,2	8,11,12	0.99	0	6,14,16	0.94	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
1	HIC	C	1	1,2	-	1/5/6/8	0/1/1/1
1	HIC	A	1	1,2	-	0/5/6/8	0/1/1/1
1	HIC	D	1	1,2	-	0/5/6/8	0/1/1/1
1	HIC	E	1	1,2	-	0/5/6/8	0/1/1/1
1	HIC	B	1	1,2	-	1/5/6/8	0/1/1/1
1	HIC	F	1	1,2	-	0/5/6/8	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	1	HIC	CD2-CG	3.34	1.41	1.36
1	A	1	HIC	CD2-CG	2.42	1.39	1.36
1	B	1	HIC	CD2-CG	2.09	1.39	1.36

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	1	HIC	CA-CB-CG-ND1
1	B	1	HIC	CA-CB-CG-ND1

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 6 are monoatomic - leaving 8 for Mogul analysis.

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	SO4	D	302	-	4,4,4	0.54	0	6,6,6	0.35	0
3	SO4	E	302	-	4,4,4	0.51	0	6,6,6	0.40	0
3	SO4	F	303	-	4,4,4	0.47	0	6,6,6	0.35	0
3	SO4	D	303	-	4,4,4	0.52	0	6,6,6	0.55	0
3	SO4	F	302	-	4,4,4	0.43	0	6,6,6	0.29	0
3	SO4	E	304	-	4,4,4	0.36	0	6,6,6	0.56	0
3	SO4	E	303	-	4,4,4	0.30	0	6,6,6	0.50	0
3	SO4	B	302	-	4,4,4	0.53	0	6,6,6	0.77	0

There are no bond length outliers.

There are no bond angle outliers.

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	E	303	SO4	1	0

## 5.7 Other polymers [i](#)

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	223/252 (88%)	-0.23	5 (2%) 62 64	12, 20, 38, 70	0
1	B	226/252 (89%)	-0.28	2 (0%) 84 85	13, 20, 37, 56	0
1	C	225/252 (89%)	-0.12	5 (2%) 62 64	14, 23, 44, 75	0
1	D	223/252 (88%)	-0.00	4 (1%) 68 71	15, 27, 44, 70	0
1	E	225/252 (89%)	0.03	4 (1%) 68 71	15, 28, 48, 78	0
1	F	223/252 (88%)	0.09	7 (3%) 49 51	15, 27, 45, 79	0
All	All	1345/1512 (88%)	-0.09	27 (2%) 65 68	12, 24, 45, 79	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	226	ASP	4.9
1	F	180	VAL	3.9
1	C	226	ASP	3.6
1	F	223	ASP	3.4
1	B	210	ASP	3.4
1	A	180	VAL	3.3
1	F	224	ALA	3.3
1	F	179	GLY	3.1
1	F	182	GLY	3.0
1	D	224	ALA	2.9
1	C	180	VAL	2.9
1	A	224	ALA	2.9
1	D	162	PRO	2.8
1	A	179	GLY	2.7
1	C	179	GLY	2.7
1	C	181[A]	GLY	2.6
1	E	225	GLN	2.4
1	A	223	ASP	2.4
1	E	199	SER	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	160	LYS	2.3
1	F	163	LYS	2.2
1	A	183	ASP	2.2
1	D	159	ASP	2.1
1	E	114	THR	2.1
1	B	227	GLU	2.1
1	C	210[A]	ASP	2.0
1	F	161[A]	ASN	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	HIC	F	1	11/12	0.86	0.16	33,38,49,52	0
1	HIC	D	1	11/12	0.90	0.13	25,28,33,35	0
1	HIC	C	1	11/12	0.90	0.14	34,42,46,48	0
1	HIC	E	1	11/12	0.94	0.13	23,31,38,39	0
1	HIC	B	1	11/12	0.95	0.08	24,27,28,28	0
1	HIC	A	1	11/12	0.96	0.09	18,22,29,29	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates 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
3	SO4	F	303	5/5	0.63	0.24	87,93,95,100	0
3	SO4	F	302	5/5	0.78	0.24	94,95,99,106	0
3	SO4	E	304	5/5	0.86	0.21	65,74,77,78	0
3	SO4	E	303	5/5	0.90	0.14	70,75,82,83	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	D	303	5/5	0.91	0.16	51,54,60,63	0
3	SO4	E	302	5/5	0.91	0.15	55,56,62,64	0
3	SO4	D	302	5/5	0.92	0.15	44,53,55,62	0
3	SO4	B	302	5/5	0.94	0.13	42,47,52,55	0
2	CU	B	301	1/1	0.99	0.02	28,28,28,28	0
2	CU	C	301	1/1	0.99	0.05	38,38,38,38	0
2	CU	D	301	1/1	0.99	0.03	27,27,27,27	0
2	CU	E	301	1/1	0.99	0.04	27,27,27,27	0
2	CU	F	301	1/1	0.99	0.04	37,37,37,37	0
2	CU	A	301	1/1	1.00	0.03	21,21,21,21	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.