Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 03:43 PM GMT

PDB ID : 4D5T
Title : Structure of N-terminally truncated A49 from Vaccinia Virus Western Reserve
Authors : Neidel, S.; Maluquer de Motes, C.; Mansur, D.S.; Strnadova, P.; Smith, G.L.; Graham, S.C.
Deposited on : 2014-11-07
Resolution : 1.84 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

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 1.84 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.
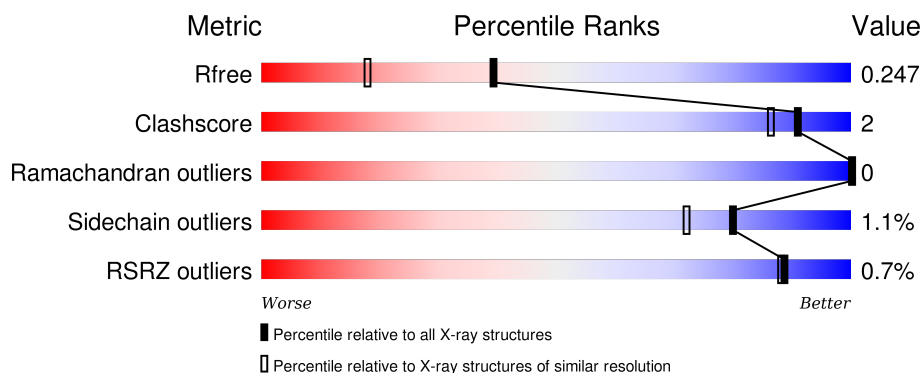
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2634 (1.86-1.82)
Clashscore	102246	2862 (1.86-1.82)
Ramachandran outliers	100387	2831 (1.86-1.82)
Sidechain outliers	100360	2832 (1.86-1.82)
RSRZ outliers	91569	2639 (1.86-1.82)

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 >=3, 2, 1 and 0 types of geometric quality criteria. 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 <=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	160	
1	B	160	
1	C	160	
1	D	160	
1	E	160	

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	160	85% · 8% · 6%
1	G	160	86% · 8% · 6%
1	H	160	89% · 5% · 6%

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
2	SO4	D	1000	-	-	-	X
2	SO4	G	1000	-	-	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 11015 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 PROTEIN A49R.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	151	Total 1290	C 812	N 205	O 265	S 8	0	11	0
1	B	151	Total 1281	C 805	N 205	O 263	S 8	0	9	0
1	C	150	Total 1297	C 816	N 206	O 267	S 8	0	11	0
1	D	151	Total 1276	C 804	N 205	O 260	S 7	0	8	0
1	E	151	Total 1248	C 791	N 198	O 252	S 7	0	4	0
1	F	150	Total 1238	C 784	N 194	O 254	S 6	0	4	0
1	G	150	Total 1242	C 787	N 196	O 252	S 7	0	5	0
1	H	150	Total 1224	C 777	N 196	O 244	S 7	0	3	0

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	MET	-	EXPRESSION TAG	UNP P31037
A	163	GLY	-	EXPRESSION TAG	UNP P31037
A	164	SER	-	EXPRESSION TAG	UNP P31037
A	165	LYS	-	EXPRESSION TAG	UNP P31037
A	166	HIS	-	EXPRESSION TAG	UNP P31037
A	167	HIS	-	EXPRESSION TAG	UNP P31037
A	168	HIS	-	EXPRESSION TAG	UNP P31037
A	169	HIS	-	EXPRESSION TAG	UNP P31037
A	170	HIS	-	EXPRESSION TAG	UNP P31037
A	171	HIS	-	EXPRESSION TAG	UNP P31037
B	12	MET	-	EXPRESSION TAG	UNP P31037
B	163	GLY	-	EXPRESSION TAG	UNP P31037
B	164	SER	-	EXPRESSION TAG	UNP P31037

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	165	LYS	-	EXPRESSION TAG	UNP P31037
B	166	HIS	-	EXPRESSION TAG	UNP P31037
B	167	HIS	-	EXPRESSION TAG	UNP P31037
B	168	HIS	-	EXPRESSION TAG	UNP P31037
B	169	HIS	-	EXPRESSION TAG	UNP P31037
B	170	HIS	-	EXPRESSION TAG	UNP P31037
B	171	HIS	-	EXPRESSION TAG	UNP P31037
C	12	MET	-	EXPRESSION TAG	UNP P31037
C	163	GLY	-	EXPRESSION TAG	UNP P31037
C	164	SER	-	EXPRESSION TAG	UNP P31037
C	165	LYS	-	EXPRESSION TAG	UNP P31037
C	166	HIS	-	EXPRESSION TAG	UNP P31037
C	167	HIS	-	EXPRESSION TAG	UNP P31037
C	168	HIS	-	EXPRESSION TAG	UNP P31037
C	169	HIS	-	EXPRESSION TAG	UNP P31037
C	170	HIS	-	EXPRESSION TAG	UNP P31037
C	171	HIS	-	EXPRESSION TAG	UNP P31037
D	12	MET	-	EXPRESSION TAG	UNP P31037
D	163	GLY	-	EXPRESSION TAG	UNP P31037
D	164	SER	-	EXPRESSION TAG	UNP P31037
D	165	LYS	-	EXPRESSION TAG	UNP P31037
D	166	HIS	-	EXPRESSION TAG	UNP P31037
D	167	HIS	-	EXPRESSION TAG	UNP P31037
D	168	HIS	-	EXPRESSION TAG	UNP P31037
D	169	HIS	-	EXPRESSION TAG	UNP P31037
D	170	HIS	-	EXPRESSION TAG	UNP P31037
D	171	HIS	-	EXPRESSION TAG	UNP P31037
E	12	MET	-	EXPRESSION TAG	UNP P31037
E	163	GLY	-	EXPRESSION TAG	UNP P31037
E	164	SER	-	EXPRESSION TAG	UNP P31037
E	165	LYS	-	EXPRESSION TAG	UNP P31037
E	166	HIS	-	EXPRESSION TAG	UNP P31037
E	167	HIS	-	EXPRESSION TAG	UNP P31037
E	168	HIS	-	EXPRESSION TAG	UNP P31037
E	169	HIS	-	EXPRESSION TAG	UNP P31037
E	170	HIS	-	EXPRESSION TAG	UNP P31037
E	171	HIS	-	EXPRESSION TAG	UNP P31037
F	12	MET	-	EXPRESSION TAG	UNP P31037
F	163	GLY	-	EXPRESSION TAG	UNP P31037
F	164	SER	-	EXPRESSION TAG	UNP P31037
F	165	LYS	-	EXPRESSION TAG	UNP P31037
F	166	HIS	-	EXPRESSION TAG	UNP P31037

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	167	HIS	-	EXPRESSION TAG	UNP P31037
F	168	HIS	-	EXPRESSION TAG	UNP P31037
F	169	HIS	-	EXPRESSION TAG	UNP P31037
F	170	HIS	-	EXPRESSION TAG	UNP P31037
F	171	HIS	-	EXPRESSION TAG	UNP P31037
G	12	MET	-	EXPRESSION TAG	UNP P31037
G	163	GLY	-	EXPRESSION TAG	UNP P31037
G	164	SER	-	EXPRESSION TAG	UNP P31037
G	165	LYS	-	EXPRESSION TAG	UNP P31037
G	166	HIS	-	EXPRESSION TAG	UNP P31037
G	167	HIS	-	EXPRESSION TAG	UNP P31037
G	168	HIS	-	EXPRESSION TAG	UNP P31037
G	169	HIS	-	EXPRESSION TAG	UNP P31037
G	170	HIS	-	EXPRESSION TAG	UNP P31037
G	171	HIS	-	EXPRESSION TAG	UNP P31037
H	12	MET	-	EXPRESSION TAG	UNP P31037
H	163	GLY	-	EXPRESSION TAG	UNP P31037
H	164	SER	-	EXPRESSION TAG	UNP P31037
H	165	LYS	-	EXPRESSION TAG	UNP P31037
H	166	HIS	-	EXPRESSION TAG	UNP P31037
H	167	HIS	-	EXPRESSION TAG	UNP P31037
H	168	HIS	-	EXPRESSION TAG	UNP P31037
H	169	HIS	-	EXPRESSION TAG	UNP P31037
H	170	HIS	-	EXPRESSION TAG	UNP P31037
H	171	HIS	-	EXPRESSION TAG	UNP P31037

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 5	O 4	S 1	0	0
2	C	1	Total 5	O 4	S 1	0	0
2	D	1	Total 5	O 4	S 1	0	0
2	F	1	Total 5	O 4	S 1	0	0
2	G	1	Total 5	O 4	S 1	0	0

- Molecule 3 is water.

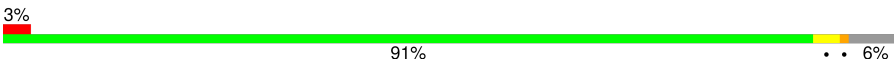| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	152	Total 152	O 152	0	0
3	B	113	Total 113	O 113	0	0
3	C	126	Total 126	O 126	0	0
3	D	87	Total 87	O 87	0	0
3	E	129	Total 129	O 129	0	0
3	F	95	Total 95	O 95	0	0
3	G	110	Total 110	O 110	0	0
3	H	82	Total 82	O 82	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: PROTEIN A49R

- Molecule 1: PROTEIN A49R

- Molecule 1: PROTEIN A49R

- Molecule 1: PROTEIN A49R

- Molecule 1: PROTEIN A49R

- Molecule 1: PROTEIN A49R

MET V13 L14 G15 Y16 V31 I32 N49 C53 R54 S55 N56 L57 D58 L81 F98 K103 R136 D139 I150 L151 L162 GLY SER LYS HIS HIS HIS HIS HIS HIS

• Molecule 1: PROTEIN A49R

Chain G:

86% 8% • 6%

MET V13 I32 E36 D39 N40 N49 C53 R54 F98 D124 T127 R136 D139 L151 L162 GLY SER LYS HIS HIS HIS HIS HIS HIS

• Molecule 1: PROTEIN A49R

Chain H:

89% 5% 6%

MET V13 S28 Q29 I32 E36 N49 R54 L151 S154 L162 GLY SER LYS HIS HIS HIS HIS HIS HIS
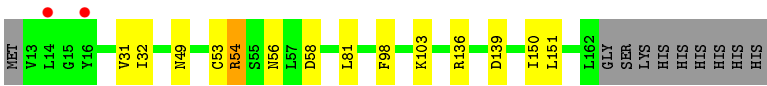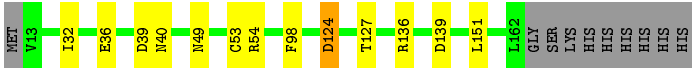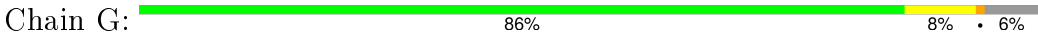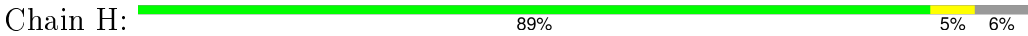
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	92.29Å 45.63Å 160.26Å 90.00° 98.74° 90.00°	Depositor
Resolution (Å)	45.71 – 1.84 45.71 – 1.84	Depositor EDS
% Data completeness (in resolution range)	97.6 (45.71-1.84) 97.6 (45.71-1.84)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 1.84Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.206 , 0.239 0.215 , 0.247	Depositor DCC
R_{free} test set	5667 reflections (5.29%)	DCC
Wilson B-factor (Å^2)	29.8	Xtriage
Anisotropy	0.289	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 43.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Outliers	4 of 112791 reflections (0.004%)	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	11015	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 35.69 % 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 5.6125e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: SO4

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	0.88	0/1308	0.95	4/1775 (0.2%)
1	B	0.86	0/1296	0.87	2/1758 (0.1%)
1	C	0.90	0/1312	0.98	6/1779 (0.3%)
1	D	0.86	0/1291	0.87	4/1750 (0.2%)
1	E	0.91	2/1263 (0.2%)	0.90	0/1714
1	F	0.81	0/1253	0.85	2/1702 (0.1%)
1	G	0.89	0/1263	0.89	2/1715 (0.1%)
1	H	0.82	0/1239	0.84	0/1683
All	All	0.87	2/10225 (0.0%)	0.90	20/13876 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	148	GLU	CD-OE2	-5.75	1.19	1.25
1	E	36	GLU	CD-OE2	5.49	1.31	1.25

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	147	ARG	NE-CZ-NH1	9.17	124.89	120.30
1	A	141	ASP	CB-CG-OD2	-8.76	110.42	118.30
1	C	141	ASP	CB-CG-OD2	-8.29	110.84	118.30
1	G	124	ASP	CB-CG-OD2	6.58	124.23	118.30
1	A	98	PHE	CB-CG-CD1	-6.44	116.29	120.80
1	A	147	ARG	CB-CA-C	-6.12	98.16	110.40
1	F	136	ARG	NE-CZ-NH2	-6.11	117.25	120.30
1	C	147	ARG	CB-CA-C	-6.08	98.25	110.40
1	A	147	ARG	NE-CZ-NH1	5.75	123.18	120.30
1	G	136	ARG	NE-CZ-NH2	-5.70	117.45	120.30
1	B	98	PHE	CB-CG-CD1	-5.54	116.92	120.80

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	98	PHE	CB-CG-CD1	-5.50	116.95	120.80
1	D	98	PHE	CB-CG-CD1	-5.45	116.99	120.80
1	F	54	ARG	NE-CZ-NH1	5.45	123.03	120.30
1	D	36	GLU	CG-CD-OE1	5.37	129.04	118.30
1	C	147	ARG	NE-CZ-NH2	-5.28	117.66	120.30
1	B	54	ARG	NE-CZ-NH1	5.25	122.92	120.30
1	D	54	ARG	NE-CZ-NH1	5.10	122.85	120.30
1	D	36	GLU	CG-CD-OE2	-5.05	108.20	118.30
1	C	147	ARG	CD-NE-CZ	5.02	130.63	123.60

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	1290	0	1250	15	0
1	B	1281	0	1242	3	1
1	C	1297	0	1261	7	0
1	D	1276	0	1247	4	1
1	E	1248	0	1231	4	1
1	F	1238	0	1210	12	1
1	G	1242	0	1222	6	0
1	H	1224	0	1206	4	0
2	A	5	0	0	0	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
2	F	5	0	0	0	0
2	G	5	0	0	0	0
3	A	152	0	0	5	0
3	B	113	0	0	2	0
3	C	126	0	0	3	0
3	D	87	0	0	1	0
3	E	129	0	0	1	0
3	F	95	0	0	1	0
3	G	110	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	82	0	0	1	0
All	All	11015	0	9869	45	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 2.

All (45) 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:A:32[A]:ILE:HG22	1:F:151:LEU:HD11	1.54	0.89
1:A:32[A]:ILE:HG22	1:F:151:LEU:CD1	2.10	0.81
1:C:103:LYS:NZ	1:G:139[B]:ASP:OD2	2.16	0.78
1:C:104[B]:ASP:OD1	3:C:2088:HOH:O	2.04	0.75
1:A:41[B]:ASP:OD1	3:A:2040:HOH:O	2.09	0.70
1:A:32[A]:ILE:CG2	1:F:151:LEU:HD11	2.22	0.69
1:C:41[A]:ASP:OD1	3:C:2036:HOH:O	2.11	0.69
1:B:74[A]:GLN:NE2	3:B:2049:HOH:O	2.28	0.65
1:A:32[B]:ILE:HD11	1:F:32:ILE:HG21	1.79	0.63
1:A:123:PRO:HG3	1:G:124:ASP:OD1	2.01	0.60
1:A:104[B]:ASP:OD1	3:A:2103:HOH:O	2.16	0.60
1:A:32[B]:ILE:HD11	1:F:32:ILE:CG2	2.31	0.60
1:E:136:ARG:NH1	3:E:2109:HOH:O	2.35	0.60
1:G:49:ASN:OD1	1:G:54:ARG:NH1	2.37	0.57
1:H:49:ASN:OD1	1:H:54:ARG:NH1	2.39	0.56
1:D:100[A]:SER:OG	3:D:2047:HOH:O	2.18	0.55
1:B:49:ASN:OD1	1:B:54:ARG:NH1	2.39	0.54
1:A:74[B]:GLN:CD	3:A:2073:HOH:O	2.45	0.54
1:D:49:ASN:OD1	1:D:54:ARG:NH1	2.40	0.53
1:A:49:ASN:OD1	1:A:54:ARG:NH1	2.41	0.52
1:C:49:ASN:OD1	1:C:54:ARG:NH1	2.42	0.52
1:F:49:ASN:OD1	1:F:54:ARG:NH1	2.42	0.51
1:H:28[B]:SER:OG	1:H:154:SER:HB2	2.11	0.51
1:B:74[A]:GLN:CD	3:B:2049:HOH:O	2.49	0.51
3:A:2114:HOH:O	1:D:15:GLY:N	2.42	0.50
1:E:32[A]:ILE:HD12	1:E:147:ARG:HG3	1.93	0.50
1:E:49:ASN:OD1	1:E:54:ARG:NH1	2.44	0.50
1:A:74[B]:GLN:NE2	3:A:2073:HOH:O	2.45	0.49
1:A:123:PRO:HB2	1:G:127:THR:OG1	2.13	0.48
1:A:32[B]:ILE:CD1	1:F:32:ILE:HG21	2.43	0.48
1:A:32[A]:ILE:HG13	1:A:151:LEU:HD21	1.99	0.45
1:C:131:GLU:HB3	3:C:2100:HOH:O	2.15	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:32[A]:ILE:HG13	1:C:151:LEU:HD21	2.00	0.44
1:H:29:GLN:NE2	3:H:2015:HOH:O	2.50	0.43
1:F:56:ASN:OD1	1:F:58[B]:ASP:HB2	2.18	0.43
1:F:81:LEU:HD23	1:F:81:LEU:C	2.39	0.43
1:F:103:LYS:HE2	3:F:2069:HOH:O	2.19	0.42
1:H:32:ILE:HG13	1:H:151:LEU:HD21	2.02	0.42
1:G:32:ILE:HG13	1:G:151:LEU:HD21	2.03	0.41
1:A:32[A]:ILE:HG22	1:F:151:LEU:HD13	2.00	0.41
1:G:39[A]:ASP:OD1	1:G:40:ASN:N	2.53	0.41
1:D:81:LEU:C	1:D:81:LEU:HD23	2.41	0.41
1:F:31:VAL:HG11	1:F:150:ILE:HD13	2.02	0.41
1:C:81:LEU:HD23	1:C:81:LEU:C	2.41	0.41
1:E:32[B]:ILE:HG13	1:E:151:LEU:HD21	2.03	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:B:103:LYS:NZ	1:F:139[B]:ASP:OD2[1_545]	2.00	0.20
1:D:103:LYS:NZ	1:E:139[A]:ASP:OD2[1_445]	2.14	0.06

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	160/160 (100%)	158 (99%)	2 (1%)	0	100	100
1	B	158/160 (99%)	156 (99%)	2 (1%)	0	100	100
1	C	159/160 (99%)	157 (99%)	2 (1%)	0	100	100
1	D	157/160 (98%)	155 (99%)	2 (1%)	0	100	100
1	E	153/160 (96%)	151 (99%)	2 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	152/160 (95%)	149 (98%)	3 (2%)	0	100	100
1	G	153/160 (96%)	150 (98%)	3 (2%)	0	100	100
1	H	151/160 (94%)	149 (99%)	2 (1%)	0	100	100
All	All	1243/1280 (97%)	1225 (99%)	18 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	151/152 (99%)	151 (100%)	0	100	100
1	B	150/152 (99%)	149 (99%)	1 (1%)	88	84
1	C	153/152 (101%)	153 (100%)	0	100	100
1	D	149/152 (98%)	146 (98%)	3 (2%)	63	47
1	E	146/152 (96%)	143 (98%)	3 (2%)	61	45
1	F	145/152 (95%)	143 (99%)	2 (1%)	74	63
1	G	146/152 (96%)	142 (97%)	4 (3%)	52	34
1	H	142/152 (93%)	141 (99%)	1 (1%)	88	84
All	All	1182/1216 (97%)	1168 (99%)	14 (1%)	80	69

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	36	GLU
1	D	36	GLU
1	D	53	CYS
1	D	88	ASN
1	E	53[A]	CYS
1	E	53[B]	CYS
1	E	98	PHE
1	F	53	CYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	98	PHE
1	G	36	GLU
1	G	53[A]	CYS
1	G	53[B]	CYS
1	G	98	PHE
1	H	36	GLU

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

Mol	Chain	Res	Type
1	H	29	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

5 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	#$|Z| > 2$
2	SO4	A	1000	-	4,4,4	0.51	0	6,6,6	0.61	0
2	SO4	C	1000	-	4,4,4	0.83	0	6,6,6	0.40	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z|$ > 2	Counts	RMSZ	#$|Z|$ > 2
2	SO4	D	1000	-	4,4,4	0.42	0	6,6,6	0.69	0
2	SO4	F	1000	-	4,4,4	0.69	0	6,6,6	0.89	0
2	SO4	G	1000	-	4,4,4	0.09	0	6,6,6	0.28	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
2	SO4	A	1000	-	-	0/0/0/0	0/0/0/0
2	SO4	C	1000	-	-	0/0/0/0	0/0/0/0
2	SO4	D	1000	-	-	0/0/0/0	0/0/0/0
2	SO4	F	1000	-	-	0/0/0/0	0/0/0/0
2	SO4	G	1000	-	-	0/0/0/0	0/0/0/0

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.

No monomer is involved in short contacts.

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, 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(Å²)	Q<0.9
1	A	151/160 (94%)	-0.26	0	100	100	20, 31, 52, 66	0
1	B	151/160 (94%)	-0.15	4 (2%)	59	56	21, 37, 58, 81	0
1	C	150/160 (93%)	-0.22	0	100	100	21, 32, 54, 61	0
1	D	151/160 (94%)	-0.15	3 (1%)	68	66	20, 37, 59, 70	0
1	E	151/160 (94%)	-0.19	0	100	100	19, 34, 54, 66	0
1	F	150/160 (93%)	-0.01	2 (1%)	79	78	25, 38, 63, 85	0
1	G	150/160 (93%)	-0.19	0	100	100	22, 36, 58, 72	0
1	H	150/160 (93%)	-0.12	0	100	100	24, 39, 58, 66	0
All	All	1204/1280 (94%)	-0.16	9 (0%)	89	88	19, 35, 58, 85	0

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	13	VAL	5.2
1	D	13	VAL	3.8
1	F	14	LEU	3.3
1	D	14	LEU	3.1
1	B	14	LEU	3.1
1	B	33	ALA	2.8
1	B	15	GLY	2.5
1	F	16	TYR	2.2
1	D	33	ALA	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	LLDF	B-factors(Å^2)	Q<0.9
2	SO4	D	1000	5/5	0.91	0.17	4.81	35,37,41,42	5
2	SO4	G	1000	5/5	0.93	0.23	2.03	48,48,50,56	5
2	SO4	C	1000	5/5	0.97	0.07	-1.91	36,41,45,46	5
2	SO4	A	1000	5/5	0.99	0.05	-5.09	34,36,40,42	5
2	SO4	F	1000	5/5	0.89	0.17	-	39,40,49,50	5

6.5 Other polymers (i)

There are no such residues in this entry.