Full wwPDB X-ray Structure Validation Report (i)

Oct 11, 2023 – 01:48 PM EDT

PDB ID	:	7S6M
Title	:	Human PARP1 deltaV687-E688 bound to a DNA double strand break.
Authors	:	Rouleau-Turcotte, E.; Pascal, J.M.
Deposited on	:	2021-09-14
Resolution	:	3.20 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

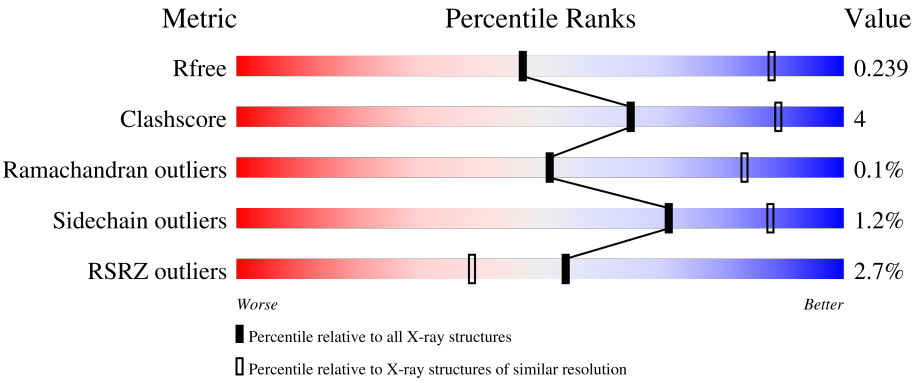
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.20 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.



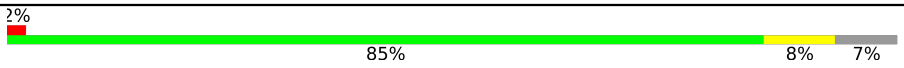
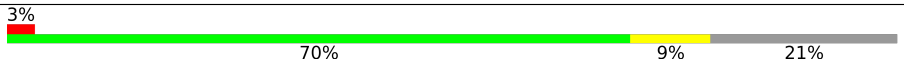
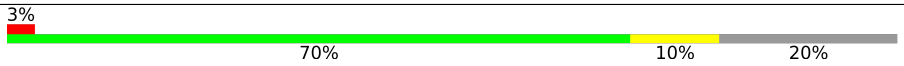
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 <=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	N	5	80% 20%
1	O	5	80% 20%
2	M	5	80% 20%
2	P	5	100%
3	B	504	2% 84% 9% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	504	
4	A	276	
4	C	276	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 22401 atoms, of which 11056 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 DNA chain called DNA (5'-D(*CP*GP*AP*CP*G)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	O	5	Total	C	H	N	O	P	0	0	0
			157	48	57	21	27	4			
1	N	5	Total	C	H	N	O	P	0	0	0
			157	48	57	21	27	4			

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*GP*TP*CP*G)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	P	5	Total	C	H	N	O	P	0	0	0
			157	48	58	18	29	4			
2	M	5	Total	C	H	N	O	P	0	0	0
			157	48	58	18	29	4			

- Molecule 3 is a protein called Poly [ADP-ribose] polymerase 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	D	471	Total	C	H	N	O	S	0	0	0
			7479	2391	3739	634	702	13			
3	B	467	Total	C	H	N	O	S	0	0	0
			7444	2379	3725	630	697	13			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	517	MET	-	initiating methionine	UNP P09874
D	?	-	VAL	deletion	UNP P09874
D	?	-	GLU	deletion	UNP P09874
D	762	ALA	VAL	variant	UNP P09874
D	1015	LEU	-	expression tag	UNP P09874
D	1016	GLU	-	expression tag	UNP P09874
D	1017	HIS	-	expression tag	UNP P09874
D	1018	HIS	-	expression tag	UNP P09874

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	1019	HIS	-	expression tag	UNP P09874
D	1020	HIS	-	expression tag	UNP P09874
D	1021	HIS	-	expression tag	UNP P09874
D	1022	HIS	-	expression tag	UNP P09874
B	517	MET	-	initiating methionine	UNP P09874
B	?	-	VAL	deletion	UNP P09874
B	?	-	GLU	deletion	UNP P09874
B	762	ALA	VAL	variant	UNP P09874
B	1015	LEU	-	expression tag	UNP P09874
B	1016	GLU	-	expression tag	UNP P09874
B	1017	HIS	-	expression tag	UNP P09874
B	1018	HIS	-	expression tag	UNP P09874
B	1019	HIS	-	expression tag	UNP P09874
B	1020	HIS	-	expression tag	UNP P09874
B	1021	HIS	-	expression tag	UNP P09874
B	1022	HIS	-	expression tag	UNP P09874

- Molecule 4 is a protein called Fusion of human PARP1 zinc fingers 1 and 3 (Zn1, Zn3).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	A	219	Total	C	H	N	O	S	0	0	0
			3379	1091	1658	295	322	13			
4	C	222	Total	C	H	N	O	S	0	0	0
			3434	1106	1689	300	326	13			

There are 40 discrepancies between the modelled and reference sequences:

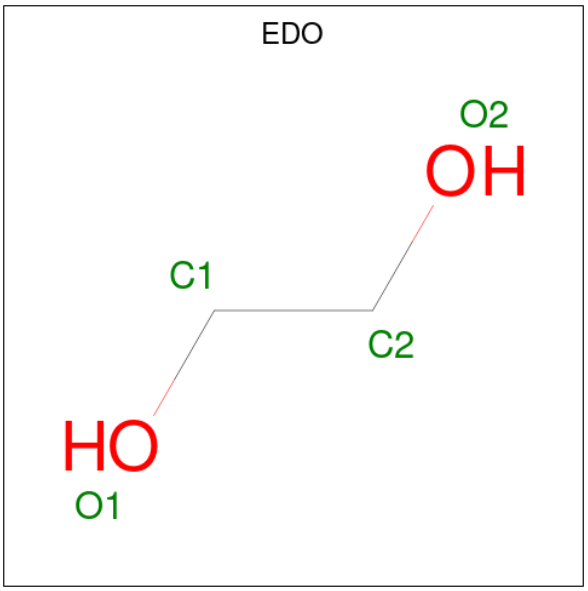
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP P09874
A	-18	GLY	-	expression tag	UNP P09874
A	-17	SER	-	expression tag	UNP P09874
A	-16	SER	-	expression tag	UNP P09874
A	-15	HIS	-	expression tag	UNP P09874
A	-14	HIS	-	expression tag	UNP P09874
A	-13	HIS	-	expression tag	UNP P09874
A	-12	HIS	-	expression tag	UNP P09874
A	-11	HIS	-	expression tag	UNP P09874
A	-10	HIS	-	expression tag	UNP P09874
A	-9	SER	-	expression tag	UNP P09874
A	-8	SER	-	expression tag	UNP P09874
A	-7	GLY	-	expression tag	UNP P09874
A	-6	LEU	-	expression tag	UNP P09874

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	VAL	-	expression tag	UNP P09874
A	-4	PRO	-	expression tag	UNP P09874
A	-3	ARG	-	expression tag	UNP P09874
A	-2	GLY	-	expression tag	UNP P09874
A	-1	SER	-	expression tag	UNP P09874
A	0	HIS	-	expression tag	UNP P09874
C	-19	MET	-	initiating methionine	UNP P09874
C	-18	GLY	-	expression tag	UNP P09874
C	-17	SER	-	expression tag	UNP P09874
C	-16	SER	-	expression tag	UNP P09874
C	-15	HIS	-	expression tag	UNP P09874
C	-14	HIS	-	expression tag	UNP P09874
C	-13	HIS	-	expression tag	UNP P09874
C	-12	HIS	-	expression tag	UNP P09874
C	-11	HIS	-	expression tag	UNP P09874
C	-10	HIS	-	expression tag	UNP P09874
C	-9	SER	-	expression tag	UNP P09874
C	-8	SER	-	expression tag	UNP P09874
C	-7	GLY	-	expression tag	UNP P09874
C	-6	LEU	-	expression tag	UNP P09874
C	-5	VAL	-	expression tag	UNP P09874
C	-4	PRO	-	expression tag	UNP P09874
C	-3	ARG	-	expression tag	UNP P09874
C	-2	GLY	-	expression tag	UNP P09874
C	-1	SER	-	expression tag	UNP P09874
C	0	HIS	-	expression tag	UNP P09874

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	D	1	Total C H O 9 2 5 2	0	0
5	B	1	Total C H O 9 2 5 2	0	0
5	A	1	Total C H O 9 2 5 2	0	0

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	2	Total Zn 2 2	0	0
6	C	2	Total Zn 2 2	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	2	Total O 2 2	0	0
7	B	2	Total O 2 2	0	0
7	A	1	Total O 1 1	0	0
7	C	1	Total O 1 1	0	0

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: DNA (5'-D(*CP*GP*AP*CP*G)-3')

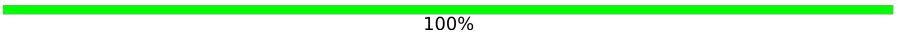
Chain O:

• Molecule 1: DNA (5'-D(*CP*GP*AP*CP*G)-3')

Chain N:

• Molecule 2: DNA (5'-D(*CP*GP*TP*CP*G)-3')

Chain P:



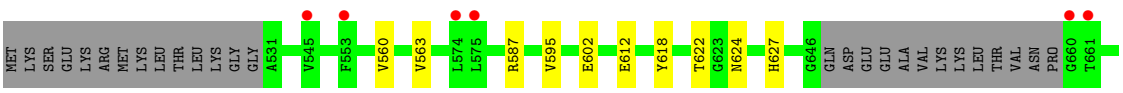
There are no outlier residues recorded for this chain.

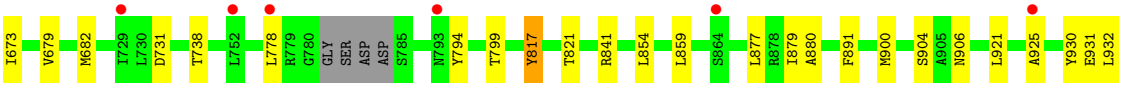
• Molecule 2: DNA (5'-D(*CP*GP*TP*CP*G)-3')

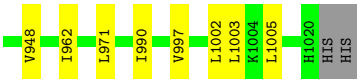
Chain M:

• Molecule 3: Poly [ADP-ribose] polymerase 1

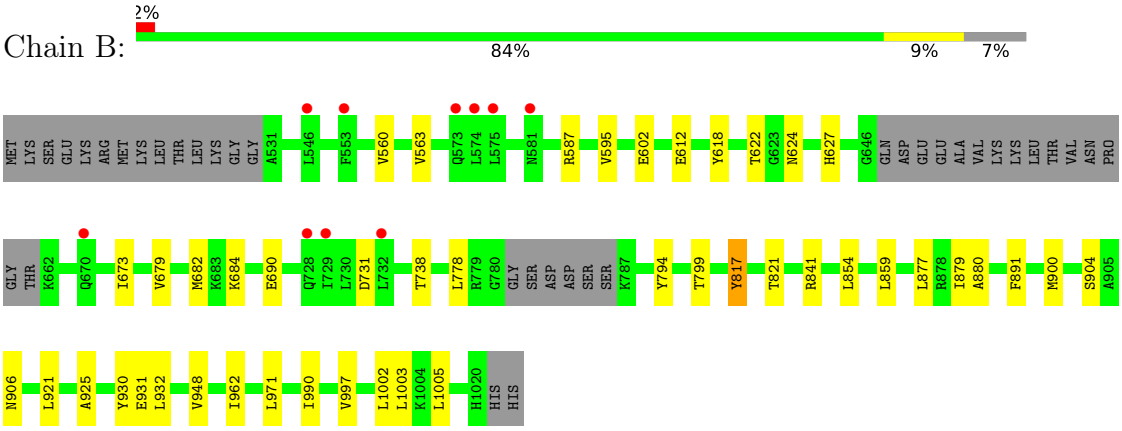
Chain D:



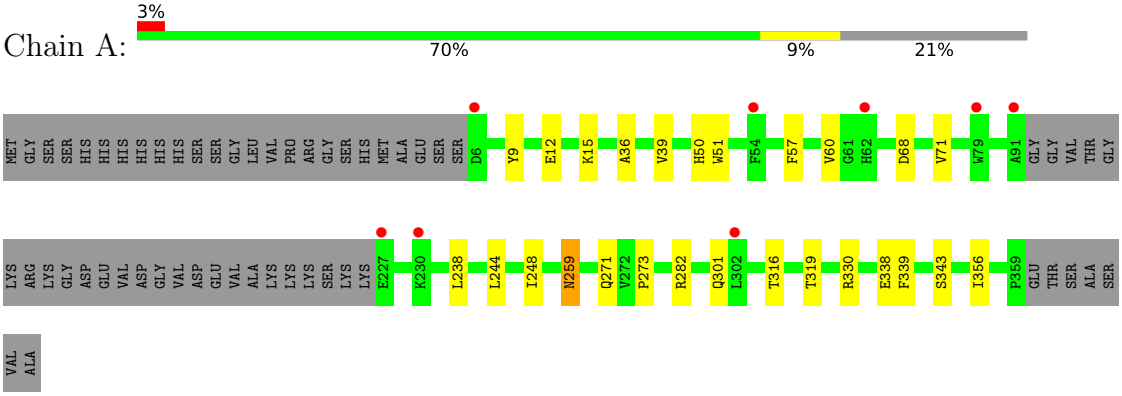




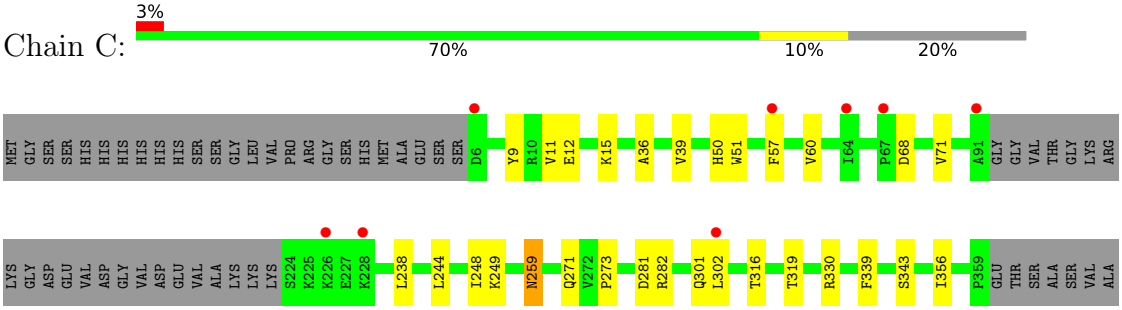
• Molecule 3: Poly [ADP-ribose] polymerase 1



• Molecule 4: Fusion of human PARP1 zinc fingers 1 and 3 (Zn1, Zn3)



• Molecule 4: Fusion of human PARP1 zinc fingers 1 and 3 (Zn1, Zn3)



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	109.70Å 112.07Å 116.48Å 90.00° 114.83° 90.00°	Depositor
Resolution (Å)	39.15 – 3.20 39.15 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.6 (39.15-3.20) 99.6 (39.15-3.20)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.17 (at 3.18Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.206 , 0.239 0.206 , 0.239	Depositor DCC
R_{free} test set	1990 reflections (4.71%)	wwPDB-VP
Wilson B-factor (Å^2)	126.3	Xtriage
Anisotropy	0.174	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 98.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	22401	wwPDB-VP
Average B, all atoms (Å^2)	142.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.84% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	N	0.48	0/112	0.72	0/171
1	O	0.45	0/112	0.72	0/171
2	M	0.56	0/110	0.88	0/168
2	P	0.55	0/110	0.89	0/168
3	B	0.26	0/3801	0.42	0/5130
3	D	0.25	0/3822	0.42	0/5159
4	A	0.26	0/1759	0.42	0/2372
4	C	0.26	0/1783	0.42	0/2402
All	All	0.27	0/11609	0.45	0/15741

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

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	N	100	57	57	1	0
1	O	100	57	57	1	0
2	M	99	58	58	1	0
2	P	99	58	58	0	0
3	B	3719	3725	3725	26	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	3740	3739	3739	26	0
4	A	1721	1658	1658	21	0
4	C	1745	1689	1689	22	0
5	A	4	5	6	0	0
5	B	4	5	6	0	0
5	D	4	5	6	0	0
6	A	2	0	0	0	0
6	C	2	0	0	0	0
7	A	1	0	0	0	0
7	B	2	0	0	0	0
7	C	1	0	0	0	0
7	D	2	0	0	0	0
All	All	11345	11056	11059	92	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 4.

All (92) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:618:TYR:O	3:D:622:THR:HG22	1.86	0.75
3:B:618:TYR:O	3:B:622:THR:HG22	1.86	0.75
3:D:731:ASP:OD1	4:C:316:THR:HG21	1.86	0.73
4:C:36:ALA:HB2	4:C:51:TRP:CE3	2.31	0.65
4:A:36:ALA:HB2	4:A:51:TRP:CE3	2.32	0.65
4:A:12:GLU:OE2	4:A:15:LYS:NZ	2.31	0.63
4:C:12:GLU:OE2	4:C:15:LYS:NZ	2.32	0.62
3:D:587:ARG:NH2	3:D:602:GLU:OE2	2.35	0.60
3:B:587:ARG:NH2	3:B:602:GLU:OE2	2.37	0.58
4:A:271:GLN:N	4:A:301:GLN:OE1	2.31	0.58
3:D:622:THR:HG23	3:D:624:ASN:H	1.69	0.57
3:D:731:ASP:HA	4:C:319:THR:HG21	1.87	0.57
3:B:682:MET:HG3	3:B:778:LEU:HD11	1.87	0.57
3:D:682:MET:HG3	3:D:778:LEU:HD11	1.86	0.57
3:B:859:LEU:HD21	3:B:921:LEU:HB3	1.87	0.56
3:D:673:ILE:HD11	3:D:794:TYR:HD1	1.70	0.56
4:C:271:GLN:N	4:C:301:GLN:OE1	2.32	0.56
3:B:622:THR:HG23	3:B:624:ASN:H	1.69	0.56
4:C:316:THR:HG23	4:C:319:THR:H	1.70	0.55
4:A:316:THR:HG23	4:A:319:THR:H	1.71	0.55
3:D:859:LEU:HD21	3:D:921:LEU:HB3	1.87	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:673:ILE:HD11	3:B:794:TYR:HD1	1.71	0.55
4:A:9:TYR:HB2	4:A:71:VAL:HG22	1.89	0.55
3:B:731:ASP:OD1	4:A:316:THR:HG21	2.08	0.54
3:B:854:LEU:HD23	3:B:925:ALA:HB1	1.89	0.54
3:D:854:LEU:HD23	3:D:925:ALA:HB1	1.89	0.54
3:D:931:GLU:O	3:D:932:LEU:HD23	2.08	0.53
4:C:9:TYR:HB2	4:C:71:VAL:HG22	1.90	0.53
4:A:57:PHE:O	4:A:60:VAL:HG22	2.09	0.53
3:B:931:GLU:O	3:B:932:LEU:HD23	2.09	0.52
4:C:57:PHE:O	4:C:60:VAL:HG22	2.09	0.52
3:B:731:ASP:HA	4:A:319:THR:HG21	1.91	0.52
4:C:39:VAL:HG21	4:C:50:HIS:NE2	2.25	0.52
4:A:39:VAL:HG21	4:A:50:HIS:NE2	2.25	0.51
3:D:891:PHE:HB2	3:D:990:ILE:CD1	2.41	0.51
3:D:618:TYR:CE2	3:D:622:THR:HG21	2.46	0.50
3:B:618:TYR:CE2	3:B:622:THR:HG21	2.46	0.50
4:C:238:LEU:HD11	4:C:356:ILE:HG21	1.93	0.50
3:B:560:VAL:HG12	3:B:738:THR:OG1	2.11	0.49
3:B:891:PHE:HB2	3:B:990:ILE:CD1	2.43	0.49
4:A:238:LEU:HD11	4:A:356:ILE:HG21	1.94	0.49
3:D:962:ILE:HG13	3:D:971:LEU:HD21	1.94	0.48
3:B:684:LYS:O	3:B:690:GLU:HB2	2.12	0.48
3:D:560:VAL:HG12	3:D:738:THR:OG1	2.13	0.48
3:B:962:ILE:HG13	3:B:971:LEU:HD21	1.95	0.48
3:D:821:THR:HB	3:D:900:MET:HA	1.95	0.48
3:B:821:THR:HB	3:B:900:MET:HA	1.96	0.48
3:B:595:VAL:HG22	3:B:595:VAL:O	2.16	0.45
4:C:339:PHE:O	4:C:343:SER:N	2.50	0.45
3:D:1003:LEU:HB3	3:D:1005:LEU:HD11	1.98	0.45
3:B:1002:LEU:C	3:B:1002:LEU:HD23	2.37	0.45
4:C:249:LYS:NZ	4:C:281:ASP:OD1	2.47	0.45
3:D:1002:LEU:HD23	3:D:1003:LEU:N	2.31	0.44
1:O:5:DG:H4'	4:A:338:GLU:OE2	2.17	0.44
3:D:877:LEU:HD12	3:D:997:VAL:HG21	1.99	0.44
3:B:1002:LEU:HD23	3:B:1003:LEU:N	2.31	0.44
3:D:1002:LEU:HD23	3:D:1002:LEU:C	2.38	0.44
3:B:799:THR:HG22	3:B:841:ARG:HG2	2.00	0.44
3:B:877:LEU:HD12	3:B:997:VAL:HG21	2.00	0.44
3:B:1003:LEU:HB3	3:B:1005:LEU:HD11	1.98	0.44
4:C:39:VAL:HG21	4:C:50:HIS:CD2	2.53	0.44
3:D:799:THR:HG22	3:D:841:ARG:HG2	2.00	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:244:LEU:HD11	4:A:248:ILE:HD11	2.00	0.43
4:C:36:ALA:HB2	4:C:51:TRP:CD2	2.54	0.43
4:A:39:VAL:HG21	4:A:50:HIS:CD2	2.53	0.43
4:A:36:ALA:HB2	4:A:51:TRP:CD2	2.54	0.43
3:D:679:VAL:HG22	3:D:778:LEU:HD13	2.01	0.42
4:A:9:TYR:CB	4:A:71:VAL:HG22	2.48	0.42
3:D:879:ILE:HG22	3:D:880:ALA:N	2.34	0.42
3:B:879:ILE:HG22	3:B:880:ALA:N	2.34	0.42
4:C:9:TYR:CB	4:C:71:VAL:HG22	2.49	0.42
3:D:595:VAL:HG22	3:D:595:VAL:O	2.18	0.42
3:D:930:TYR:HB3	3:D:948:VAL:HG22	2.02	0.42
4:A:238:LEU:CD1	4:A:356:ILE:HG21	2.50	0.42
1:N:1:DC:H42	2:M:26:DG:H1	1.68	0.42
3:B:679:VAL:HG22	3:B:778:LEU:HD13	2.01	0.42
3:D:817:TYR:CE1	3:D:821:THR:HG21	2.55	0.42
4:A:339:PHE:O	4:A:343:SER:N	2.50	0.42
3:B:817:TYR:CE1	3:B:821:THR:HG21	2.55	0.42
3:B:930:TYR:HB3	3:B:948:VAL:HG22	2.01	0.42
4:C:244:LEU:HD11	4:C:248:ILE:HD11	2.02	0.41
4:C:68:ASP:N	4:C:68:ASP:OD1	2.52	0.41
4:C:238:LEU:CD1	4:C:356:ILE:HG21	2.50	0.41
4:C:259:ASN:N	4:C:259:ASN:OD1	2.54	0.41
4:C:273:PRO:HB3	4:C:282:ARG:CZ	2.51	0.41
4:C:302:LEU:N	4:C:302:LEU:HD12	2.36	0.41
4:A:273:PRO:HB3	4:A:282:ARG:CZ	2.51	0.41
4:A:36:ALA:HA	4:A:50:HIS:O	2.21	0.41
4:A:68:ASP:N	4:A:68:ASP:OD1	2.53	0.41
4:A:259:ASN:N	4:A:259:ASN:OD1	2.54	0.41
3:D:891:PHE:HB2	3:D:990:ILE:HD13	2.02	0.40
4:C:11:VAL:HG12	4:C:71:VAL:CG1	2.52	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
3	B	461/504 (92%)	435 (94%)	25 (5%)	1 (0%)	47 79
3	D	465/504 (92%)	436 (94%)	28 (6%)	1 (0%)	47 79
4	A	215/276 (78%)	207 (96%)	8 (4%)	0	100 100
4	C	218/276 (79%)	209 (96%)	9 (4%)	0	100 100
All	All	1359/1560 (87%)	1287 (95%)	70 (5%)	2 (0%)	51 83

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	563	VAL
3	B	563	VAL

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
3	B	408/441 (92%)	403 (99%)	5 (1%)	71 88
3	D	409/441 (93%)	404 (99%)	5 (1%)	71 88
4	A	184/243 (76%)	182 (99%)	2 (1%)	73 88
4	C	187/243 (77%)	185 (99%)	2 (1%)	73 88
All	All	1188/1368 (87%)	1174 (99%)	14 (1%)	71 88

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

Mol	Chain	Res	Type
3	D	612	GLU
3	D	627	HIS
3	D	817	TYR
3	D	904	SER
3	D	906	ASN
3	B	612	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	B	627	HIS
3	B	817	TYR
3	B	904	SER
3	B	906	ASN
4	A	259	ASN
4	A	330	ARG
4	C	259	ASN
4	C	330	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

Of 7 ligands modelled in this entry, 4 are monoatomic - leaving 3 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	$\#\|Z\| > 2$
5	EDO	A	402	-	3,3,3	0.34	0	2,2,2	0.33	0
5	EDO	B	1101	-	3,3,3	0.37	0	2,2,2	0.13	0
5	EDO	D	1101	-	3,3,3	0.38	0	2,2,2	0.14	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
5	EDO	A	402	-	-	0/1/1/1	-
5	EDO	B	1101	-	-	0/1/1/1	-
5	EDO	D	1101	-	-	0/1/1/1	-

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, 95^{th} 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	N	5/5 (100%)	0.29	0 100 100	98, 98, 110, 111	0
1	O	5/5 (100%)	0.26	0 100 100	102, 102, 113, 115	0
2	M	5/5 (100%)	0.30	0 100 100	107, 111, 111, 120	0
2	P	5/5 (100%)	0.10	0 100 100	111, 115, 117, 120	0
3	B	467/504 (92%)	0.24	10 (2%) 63 49	92, 123, 167, 195	0
3	D	471/504 (93%)	0.27	12 (2%) 57 43	97, 124, 170, 211	0
4	A	219/276 (79%)	0.11	8 (3%) 41 26	96, 130, 194, 235	0
4	C	222/276 (80%)	0.11	8 (3%) 42 27	96, 129, 197, 247	0
All	All	1399/1580 (88%)	0.21	38 (2%) 54 39	92, 125, 177, 247	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	C	67	PRO	3.8
4	C	91	ALA	3.7
4	A	6	ASP	3.3
4	A	79	TRP	3.3
3	D	575	LEU	3.3
3	D	661	THR	3.1
4	C	57	PHE	3.1
4	A	91	ALA	3.1
3	B	553	PHE	3.1
4	A	230	LYS	3.0
3	B	546	LEU	3.0
3	D	660	GLY	3.0
3	D	574	LEU	3.0
3	B	728	GLN	2.8
4	C	6	ASP	2.8
4	A	54	PHE	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	B	729	ILE	2.7
3	B	575	LEU	2.6
3	D	553	PHE	2.6
4	A	62	HIS	2.5
4	C	228	LYS	2.4
3	D	545	VAL	2.4
4	A	227	GLU	2.3
4	C	302	LEU	2.3
3	D	864	SER	2.3
3	B	581	ASN	2.3
3	B	670	GLN	2.3
3	D	752	LEU	2.2
3	D	778	LEU	2.2
3	B	573	GLN	2.2
3	D	925	ALA	2.2
4	C	226	LYS	2.2
3	D	729	ILE	2.2
3	D	793	ASN	2.1
4	C	64	ILE	2.1
4	A	302	LEU	2.1
3	B	732	LEU	2.0
3	B	574	LEU	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} 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^2)	Q<0.9
5	EDO	A	402	4/4	0.67	0.16	102,122,174,174	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	ZN	C	401	1/1	0.84	0.16	229,229,229,229	0
5	EDO	B	1101	4/4	0.85	0.51	63,93,130,130	0
5	EDO	D	1101	4/4	0.92	0.49	72,87,123,123	0
6	ZN	A	403	1/1	0.95	0.15	255,255,255,255	0
6	ZN	A	401	1/1	0.97	0.14	125,125,125,125	0
6	ZN	C	402	1/1	0.98	0.23	187,187,187,187	0

6.5 Other polymers

There are no such residues in this entry.