



Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 02:46 AM GMT

PDB ID : 3HT3
Title : Crystal structure of fragment DNA polymerase I from *Bacillus stearothermophilus* V713P mutant bound to G:dCTP
Authors : Wu, E.Y.; Beese, L.S.
Deposited on : 2009-06-11
Resolution : 1.70 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

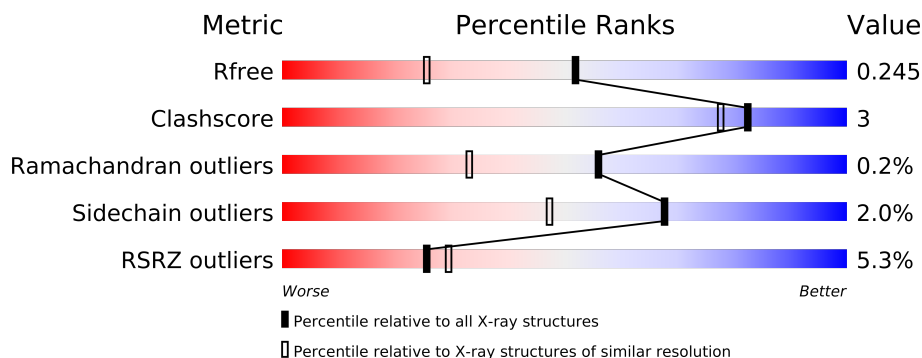
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 1.70 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}	66092	2456 (1.70-1.70)
Clashscore	79885	2929 (1.70-1.70)
Ramachandran outliers	78287	2878 (1.70-1.70)
Sidechain outliers	78261	2878 (1.70-1.70)
RSRZ outliers	66119	2456 (1.70-1.70)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	579	
1	D	579	
2	B	9	
2	E	9	
3	C	12	
3	F	12	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	DCP	A	203	-	X
5	MG	A	200	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 11370 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA polymerase I, large fragment.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	579	Total	C	N	O	S	0	5	0
			4678	2975	811	875	17			
1	D	579	Total	C	N	O	S	0	6	0
			4678	2978	810	872	18			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	598	ALA	ASP	ENGINEERED	PDB 3HT3
A	713	PRO	VAL	ENGINEERED	PDB 3HT3
D	598	ALA	ASP	ENGINEERED	PDB 3HT3
D	713	PRO	VAL	ENGINEERED	PDB 3HT3

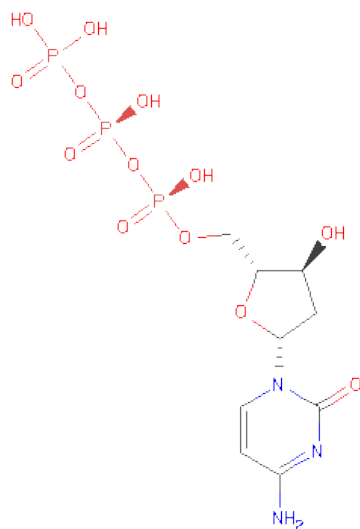
- Molecule 2 is a DNA chain called 5'-D(*CP*GP*AP*TP*CP*AP*CP*GP*(DOC))-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	9	Total	C	N	O	P	0	0	0
			178	86	34	50	8			
2	E	9	Total	C	N	O	P	0	0	0
			178	86	34	50	8			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	12	Total	C	N	O	P	0	0	0
			246	117	48	70	11			
3	F	12	Total	C	N	O	P	0	0	0
			246	117	48	70	11			

- Molecule 4 is 2'-DEOXYCYTIDINE-5'-TRIPHOSPHATE (three-letter code: DCP) (formula: C₉H₁₆N₃O₁₃P₃).

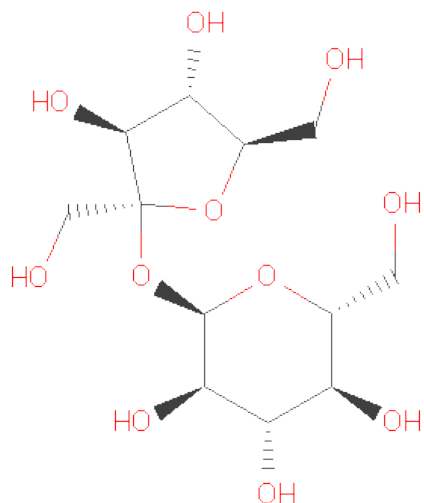


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			28	9	3	13	3		
4	A	1	Total	C	N	O	P	0	0
			28	9	3	13	3		
4	D	1	Total	C	N	O	P	0	0
			28	9	3	13	3		

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

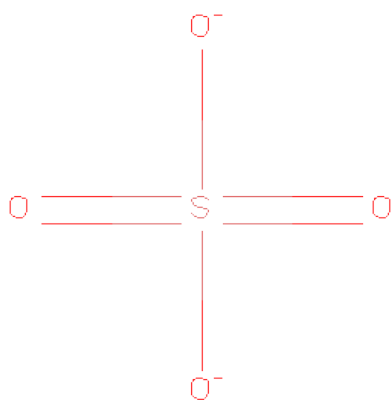
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Mg	0	0
			1	1		
5	D	1	Total	Mg	0	0
			1	1		

- Molecule 6 is SUGAR (SUCROSE) (three-letter code: SUC) (formula: C₁₂H₂₂O₁₁).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			23	12	11		
6	D	1	Total	C	O	0	0
			23	12	11		
6	D	1	Total	C	O	0	0
			23	12	11		

- Molecule 7 is SULFATE ION (three-letter code: SO₄) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	D	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	D	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		

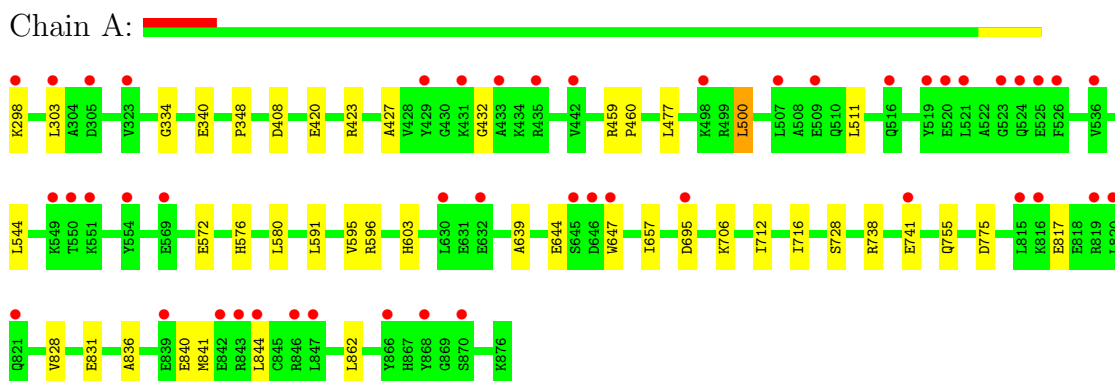
- Molecule 8 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	365	Total	O		0	0
			365	365			
8	D	521	Total	O		0	0
			521	521			
8	B	21	Total	O		0	0
			21	21			
8	C	32	Total	O		0	0
			32	32			
8	E	16	Total	O		0	0
			16	16			
8	F	41	Total	O		0	0
			41	41			

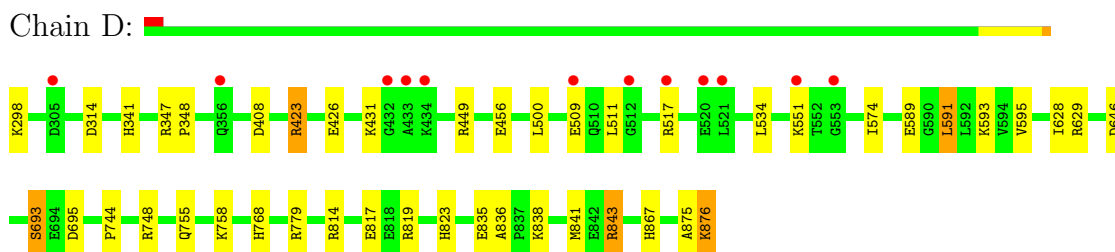
3 Residue-property plots

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.

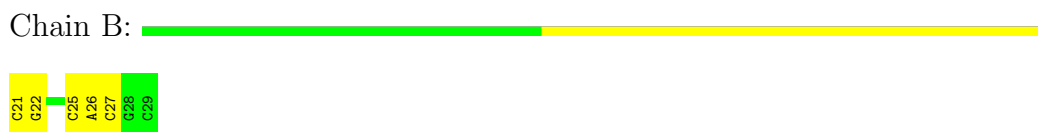
- Molecule 1: DNA polymerase I, large fragment



- Molecule 1: DNA polymerase I, large fragment



- Molecule 2: 5'-D(*CP*GP*AP*TP*CP*AP*CP*GP*(DOC))-3'

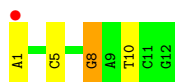


- Molecule 2: 5'-D(*CP*GP*AP*TP*CP*AP*CP*GP*(DOC))-3'



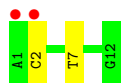
- Molecule 3: 5'-D(*AP*CP*GP*GP*CP*GP*TP*GP*AP*TP*CP*G)-3'





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

Chain F:



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	93.28Å 108.79Å 152.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.98 – 1.70 46.99 – 1.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.98-1.70) 99.8 (46.99-1.70)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.57 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.5.0070	Depositor
R, R_{free}	0.210 , 0.247 0.209 , 0.245	Depositor DCC
R_{free} test set	7528 reflections (4.64%)	DCC
Wilson B-factor (Å ²)	17.9	Xtriage
Anisotropy	0.056	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 32.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 169685 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	11370	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.19% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DOC, MG, SUC, SO4, DCP

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	0.45	0/4775	0.57	0/6452
1	D	0.52	0/4781	0.65	2/6459 (0.0%)
2	B	0.80	0/179	1.51	5/274 (1.8%)
2	E	0.82	0/179	1.61	2/274 (0.7%)
3	C	0.89	0/276	1.55	4/425 (0.9%)
3	F	0.88	0/276	1.46	3/425 (0.7%)
All	All	0.53	0/10466	0.75	16/14309 (0.1%)

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	25	DC	O4'-C4'-C3'	-10.07	99.96	106.00
1	D	843	ARG	NE-CZ-NH2	-10.01	115.30	120.30
2	B	26	DA	O4'-C1'-N9	8.76	114.13	108.00
1	D	843	ARG	NE-CZ-NH1	8.34	124.47	120.30
3	C	1	DA	O4'-C4'-C3'	-8.11	101.14	106.00
3	F	2	DC	O4'-C1'-N1	7.43	113.20	108.00
3	C	8	DG	O4'-C1'-N9	6.88	112.82	108.00
2	B	27	DC	C4'-C3'-C2'	-6.76	97.01	103.10
2	B	25	DC	O4'-C1'-N1	6.30	112.41	108.00
3	C	10	DT	C1'-O4'-C4'	-6.13	103.97	110.10
3	F	7	DT	N3-C4-O4	5.80	123.38	119.90
2	B	27	DC	O4'-C4'-C3'	-5.76	102.19	104.50
2	B	27	DC	P-O3'-C3'	5.54	126.34	119.70
3	C	5	DC	C2-N3-C4	5.29	122.55	119.90
2	E	25	DC	C4'-C3'-C2'	-5.06	98.55	103.10
3	F	7	DT	C6-C5-C7	-5.02	119.89	122.90

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4678	0	4736	26	0
1	D	4678	0	4751	36	0
2	B	178	0	102	1	0
2	E	178	0	102	2	0
3	C	246	0	136	1	0
3	F	246	0	136	0	0
4	A	56	0	24	1	0
4	D	28	0	12	0	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
6	A	23	0	22	1	0
6	D	46	0	44	2	0
7	D	15	0	0	0	0
8	A	365	0	0	7	0
8	B	21	0	0	0	0
8	C	32	0	0	1	0
8	D	521	0	0	9	0
8	E	16	0	0	0	0
8	F	41	0	0	0	0
All	All	11370	0	10065	64	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

All (64) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:758:LYS:HE2	8:D:972:HOH:O	1.56	1.04
1:D:423:ARG:HG2	1:D:423:ARG:HH11	1.31	0.96
1:D:876:LYS:HE3	8:D:14:HOH:O	1.74	0.86
1:D:843:ARG:HD2	8:D:180:HOH:O	1.75	0.86
1:D:456:GLU:HG2	8:D:295:HOH:O	1.74	0.86
1:D:819:ARG:HH11	1:D:819:ARG:HG3	1.40	0.85

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:408:ASP:HB2	6:D:877:SUC:H1'1	1.65	0.78
1:D:693:SER:HB2	1:D:695:ASP:OD1	1.92	0.70
1:A:657:ILE:HD13	1:A:862:LEU:HD22	1.78	0.66
1:D:867:HIS:CG	1:D:876:LYS:HA	2.32	0.65
1:D:875:ALA:O	1:D:876:LYS:HG2	1.96	0.65
1:A:423:ARG:HH11	1:A:427:ALA:HB1	1.64	0.63
1:D:534:LEU:HD11	1:D:574:ILE:HD13	1.80	0.63
1:A:596:ARG:HD3	1:A:603:HIS:CD2	2.34	0.62
1:D:298:LYS:HB3	1:D:449:ARG:NH2	2.15	0.62
1:D:875:ALA:C	1:D:876:LYS:HG2	2.22	0.60
1:A:695:ASP:HB3	8:A:996:HOH:O	2.02	0.59
1:D:819:ARG:NH1	1:D:819:ARG:HG3	2.13	0.58
1:A:459:ARG:HG3	8:A:1107:HOH:O	2.04	0.57
1:A:596:ARG:HD3	1:A:603:HIS:HD2	1.70	0.57
1:A:840:GLU:HB3	8:A:1052:HOH:O	2.04	0.57
1:D:426:GLU:OE2	1:D:431:LYS:HD3	2.05	0.56
1:D:744:PRO:HB2	1:D:748[B]:ARG:HH12	1.73	0.54
1:D:629:ARG:HH22	2:E:29:DOC:H5	1.73	0.54
1:D:423:ARG:HG2	1:D:423:ARG:NH1	2.08	0.53
1:D:456:GLU:CG	8:D:295:HOH:O	2.44	0.53
1:D:836:ALA:HB3	1:D:841:MET:CE	2.40	0.52
1:A:591:LEU:O	1:A:595:VAL:HG23	2.10	0.52
1:D:591:LEU:O	1:D:595:VAL:HG23	2.11	0.51
1:A:596:ARG:NH2	8:A:121:HOH:O	2.43	0.51
1:A:408:ASP:HB2	6:A:877:SUC:H1'1	1.92	0.51
1:D:744:PRO:HB2	1:D:748[B]:ARG:NH1	2.26	0.50
1:D:408:ASP:HB2	6:D:877:SUC:C1'	2.39	0.50
1:D:779:ARG:HD3	8:D:1115:HOH:O	2.12	0.48
1:A:706:LYS:HE3	4:A:201:DCP:O3B	2.14	0.48
1:D:646:ASP:HB3	1:D:838:LYS:HE3	1.96	0.48
1:A:738:ARG:HA	1:A:741:GLU:HG2	1.96	0.48
1:A:459:ARG:NH2	1:D:314:ASP:O	2.47	0.48
1:D:589:GLU:O	1:D:593:LYS:HG3	2.13	0.47
2:B:21:DC:H2''	2:B:22:DG:C8	2.49	0.47
2:E:21:DC:H2'	2:E:22:DG:C8	2.50	0.46
1:D:814:ARG:HA	1:D:817:GLU:HG2	1.98	0.46
1:A:459:ARG:HB3	1:A:460:PRO:HD3	1.98	0.45
1:A:334:GLY:HA2	1:A:348:PRO:HD3	2.00	0.44
1:A:828:VAL:HB	1:A:831:GLU:CG	2.48	0.44
1:D:341:HIS:HE1	8:D:930:HOH:O	1.99	0.44
1:D:768:HIS:HD2	8:D:1152:HOH:O	2.01	0.43
1:D:423:ARG:CG	1:D:423:ARG:NH1	2.79	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:755[A]:GLN:HG2	8:A:1053:HOH:O	2.17	0.43
1:D:823:HIS:HE1	1:D:835:GLU:OE2	2.02	0.43
1:D:867:HIS:CB	1:D:876:LYS:HA	2.48	0.43
1:D:755:GLN:HG2	8:D:126:HOH:O	2.19	0.42
1:A:738:ARG:HA	1:A:741:GLU:CG	2.50	0.42
1:A:500:LEU:HD13	1:A:639:ALA:CB	2.50	0.42
1:A:817:GLU:HG2	8:A:1143:HOH:O	2.19	0.42
1:A:576:HIS:CE1	1:A:580:LEU:HD11	2.55	0.42
1:A:775:ASP:HB2	8:A:67:HOH:O	2.20	0.41
1:A:712:ILE:HA	1:A:716:ILE:HG22	2.02	0.41
1:D:534:LEU:HD11	1:D:574:ILE:CD1	2.50	0.40
1:A:836:ALA:HB3	1:A:841:MET:CE	2.50	0.40
1:A:644:GLU:HB2	1:A:647:TRP:CD1	2.57	0.40
1:D:347:ARG:HA	1:D:348:PRO:HD3	1.98	0.40
3:C:8:DG:H4'	8:C:587:HOH:O	2.22	0.40

There are no symmetry-related clashes.

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	582/579 (100%)	569 (98%)	12 (2%)	1 (0%)	56	33
1	D	583/579 (101%)	572 (98%)	10 (2%)	1 (0%)	56	33
All	All	1165/1158 (101%)	1141 (98%)	22 (2%)	2 (0%)	56	33

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	432	GLY
1	D	628	ILE

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	500/495 (101%)	488 (98%)	12 (2%)	61	39
1	D	501/495 (101%)	492 (98%)	9 (2%)	71	53
All	All	1001/990 (101%)	980 (98%)	21 (2%)	68	45

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	298	LYS
1	A	303	LEU
1	A	340	GLU
1	A	420[A]	GLU
1	A	420[B]	GLU
1	A	477	LEU
1	A	500	LEU
1	A	511	LEU
1	A	544	LEU
1	A	572	GLU
1	A	728	SER
1	A	844	LEU
1	D	423	ARG
1	D	500	LEU
1	D	509	GLU
1	D	511	LEU
1	D	517	ARG
1	D	551	LYS
1	D	591	LEU
1	D	693	SER
1	D	876	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	579	GLN
1	A	867	HIS
1	D	341	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	405	GLN
1	D	502	GLN
1	D	755	GLN
1	D	768	HIS
1	D	823	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

2 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	$\# Z > 2$
2	DOC	B	29	3,2	17,19,20	1.17	2 (11%)	20,26,29	1.60	2 (10%)
2	DOC	E	29	3,2	17,19,20	1.25	2 (11%)	20,26,29	2.10	3 (15%)

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	DOC	B	29	3,2	-	0/5/18/19	0/2/2/2
2	DOC	E	29	3,2	-	0/5/18/19	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	29	DOC	C2-N1	3.56	1.42	1.38
2	B	29	DOC	P-OP1	3.31	1.50	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	29	DOC	P-OP1	3.20	1.50	1.46
2	B	29	DOC	C2-N1	2.85	1.41	1.38

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	29	DOC	C6-C5-C4	7.30	120.50	117.47
2	B	29	DOC	C2-N3-C4	3.90	121.22	115.57
2	B	29	DOC	C6-C5-C4	3.81	119.05	117.47
2	E	29	DOC	C2-N3-C4	3.15	120.14	115.57
2	E	29	DOC	C4'-O4'-C1'	-3.14	108.74	110.05

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 11 ligands modelled in this entry, 2 are monoatomic - leaving 9 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
4	DCP	A	201	5	29,29,29	1.03	1 (3%)	42,45,45	1.42	3 (7%)
4	DCP	A	203	-	29,29,29	0.88	1 (3%)	42,45,45	1.57	5 (11%)
6	SUC	A	877	-	24,24,24	0.36	0	36,36,36	0.66	1 (2%)
7	SO4	D	1	-	4,4,4	0.41	0	6,6,6	0.31	0
7	SO4	D	2	-	4,4,4	0.20	0	6,6,6	0.07	0
4	DCP	D	202	5	29,29,29	0.94	1 (3%)	42,45,45	1.59	6 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
7	SO4	D	3	-	4,4,4	0.07	0	6,6,6	0.14	0
6	SUC	D	877	-	24,24,24	0.33	0	36,36,36	0.58	0
6	SUC	D	878	-	24,24,24	0.43	0	36,36,36	0.57	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	DCP	A	201	5	-	0/19/34/34	0/2/2/2
4	DCP	A	203	-	-	0/19/34/34	0/2/2/2
6	SUC	A	877	-	-	0/12/51/51	0/2/2/2
7	SO4	D	1	-	-	0/0/0/0	0/0/0/0
7	SO4	D	2	-	-	0/0/0/0	0/0/0/0
4	DCP	D	202	5	-	0/19/34/34	0/2/2/2
7	SO4	D	3	-	-	0/0/0/0	0/0/0/0
6	SUC	D	877	-	-	0/12/51/51	0/2/2/2
6	SUC	D	878	-	-	0/12/51/51	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	201	DCP	C2-N1	4.22	1.42	1.38
4	A	203	DCP	C2-N1	3.31	1.41	1.38
4	D	202	DCP	C2-N1	3.20	1.41	1.38

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	203	DCP	C6-C5-C4	6.17	120.03	117.47
4	D	202	DCP	C6-C5-C4	5.52	119.76	117.47
4	A	201	DCP	C6-C5-C4	5.33	119.68	117.47
4	A	203	DCP	PB-O3B-PG	-3.95	120.09	131.68
4	A	203	DCP	C2-N3-C4	3.57	120.74	115.57
4	A	201	DCP	C2-N3-C4	3.55	120.71	115.57
4	D	202	DCP	C2-N3-C4	3.25	120.27	115.57
4	D	202	DCP	C2'-C1'-N1	-3.15	105.89	114.08
4	A	201	DCP	PB-O3B-PG	-2.68	123.83	131.68
4	D	202	DCP	PB-O3B-PG	-2.48	124.40	131.68
6	A	877	SUC	C1-O5-C5	2.21	118.03	113.73
4	D	202	DCP	C2-N1-C1'	-2.14	116.75	119.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	202	DCP	O3G-PG-O2G	2.09	115.75	107.61
4	A	203	DCP	N4-C4-N3	2.08	120.59	116.59
4	A	203	DCP	PB-O3A-PA	-2.07	125.61	131.68

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	579/579 (100%)	0.59	47 (8%) 12 15	10, 26, 42, 50	0
1	D	579/579 (100%)	0.23	12 (2%) 60 66	6, 16, 38, 45	0
2	B	9/9 (100%)	-0.15	0 100 100	14, 25, 45, 49	0
2	E	9/9 (100%)	0.10	1 (11%) 6 8	11, 24, 50, 57	0
3	C	12/12 (100%)	0.28	1 (8%) 11 14	11, 26, 43, 61	0
3	F	12/12 (100%)	0.64	2 (16%) 2 3	8, 20, 47, 67	0
All	All	1200/1200 (100%)	0.40	63 (5%) 25 29	6, 21, 41, 67	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	1	DA	5.7
1	A	844	LEU	5.2
1	A	819	ARG	5.1
1	D	433	ALA	4.8
1	A	820	LEU	4.7
1	A	551	LYS	4.4
1	A	523	GLY	4.2
1	A	550	THR	3.9
1	A	298	LYS	3.8
1	A	846	ARG	3.7
1	A	647	TRP	3.5
1	A	847	LEU	3.5
3	C	1	DA	3.4
1	A	741	GLU	3.1
1	A	519	TYR	3.1
1	D	509	GLU	3.1
1	D	551	LYS	3.0
1	A	305	ASP	2.9
1	D	305	ASP	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	432	GLY	2.9
1	A	630	LEU	2.7
1	A	525	GLU	2.7
1	A	646	ASP	2.7
1	A	842	GLU	2.7
1	A	303	LEU	2.6
1	A	507	LEU	2.6
1	A	442	VAL	2.6
1	A	516	GLN	2.6
1	D	521	LEU	2.5
3	F	2	DC	2.5
1	A	509	GLU	2.5
1	A	816	LYS	2.5
1	A	429	TYR	2.4
1	A	520	GLU	2.4
1	A	526	PHE	2.4
1	A	498	LYS	2.4
1	D	434	LYS	2.4
1	D	512	GLY	2.4
1	A	645	SER	2.4
1	A	821	GLN	2.4
1	D	517	ARG	2.4
1	A	549	LYS	2.3
1	A	524	GLN	2.3
1	A	843	ARG	2.3
1	D	553	GLY	2.3
1	A	815	LEU	2.2
1	A	569	GLU	2.2
1	D	356	GLN	2.2
1	D	520	GLU	2.2
1	A	839	GLU	2.2
1	A	431	LYS	2.2
1	A	632	GLU	2.1
1	A	695	ASP	2.1
1	A	554	TYR	2.1
1	A	866	TYR	2.1
2	E	21	DC	2.1
1	A	536	VAL	2.0
1	A	868	TYR	2.0
1	A	435	ARG	2.0
1	A	433	ALA	2.0
1	A	870	SER	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	323	VAL	2.0
1	A	521	LEU	2.0

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

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	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	DOC	E	29	18/19	0.08	-0.54	8,10,16,18	0
2	DOC	B	29	18/19	0.07	-0.96	13,14,18,20	0

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	DCP	A	203	28/28	0.33	11.47	23,38,51,51	8
5	MG	A	200	1/1	0.26	3.53	39,39,39,39	0
7	SO4	D	3	5/5	0.21	1.74	38,39,40,40	0
4	DCP	A	201	28/28	0.12	0.98	25,28,38,39	0
7	SO4	D	2	5/5	0.19	0.65	64,64,64,64	0
7	SO4	D	1	5/5	0.13	0.64	26,28,31,32	0
6	SUC	A	877	23/23	0.11	-0.09	23,30,31,32	0
4	DCP	D	202	28/28	0.08	-0.75	16,18,21,24	0
6	SUC	D	878	23/23	0.09	-0.93	15,16,18,20	0
6	SUC	D	877	23/23	0.09	-0.94	11,15,17,20	0
5	MG	D	200	1/1	0.05	-2.76	17,17,17,17	0

6.5 Other polymers ⓘ

There are no such residues in this entry.