



# Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 08:53 PM GMT

PDB ID : 3L7L  
Title : Structure of the Wall Teichoic Acid Polymerase TagF, H444N + CDPG (30 minute soak)  
Authors : Lovering, A.L.; Strynadka, N.C.J.  
Deposited on : 2009-12-28  
Resolution : 2.95 Å(reported)

This is a full wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

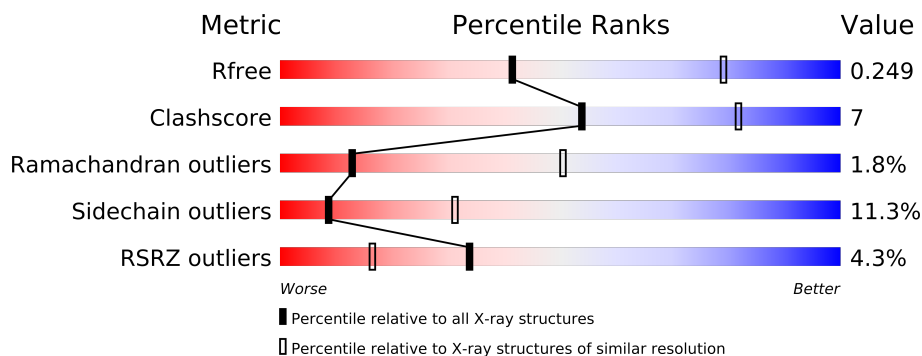
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.95 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	1587 (3.00-2.92)
Clashscore	79885	2029 (3.00-2.92)
Ramachandran outliers	78287	1955 (3.00-2.92)
Sidechain outliers	78261	1958 (3.00-2.92)
RSRZ outliers	66119	1588 (3.00-2.92)

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. 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	729	
1	B	729	
1	C	729	
1	D	729	

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

Mol	Type	Chain	Res	Geometry	Electron density
3	CL	A	733	-	X
3	CL	A	734	-	X
3	CL	A	736	-	X
3	CL	B	732	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
3	CL	B	733	-	X
3	CL	B	735	-	X
3	CL	C	734	-	X
5	EDO	D	734	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 13868 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 Teichoic acid biosynthesis protein F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	410	Total	C	N	O	S	0	0	0
			3462	2224	577	650	11			
1	B	411	Total	C	N	O	S	0	0	0
			3467	2227	578	651	11			
1	C	411	Total	C	N	O	S	0	0	0
			3467	2227	578	651	11			
1	D	401	Total	C	N	O	S	0	0	0
			3383	2169	568	635	11			

There are 36 discrepancies between the modelled and reference sequences:

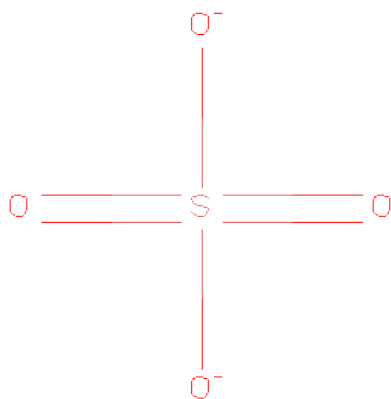
Chain	Residue	Modelled	Actual	Comment	Reference
A	444	ASN	HIS	ENGINEERED	UNP Q5HLM5
A	722	LEU	-	EXPRESSION TAG	UNP Q5HLM5
A	723	GLU	-	EXPRESSION TAG	UNP Q5HLM5
A	724	HIS	-	EXPRESSION TAG	UNP Q5HLM5
A	725	HIS	-	EXPRESSION TAG	UNP Q5HLM5
A	726	HIS	-	EXPRESSION TAG	UNP Q5HLM5
A	727	HIS	-	EXPRESSION TAG	UNP Q5HLM5
A	728	HIS	-	EXPRESSION TAG	UNP Q5HLM5
A	729	HIS	-	EXPRESSION TAG	UNP Q5HLM5
B	444	ASN	HIS	ENGINEERED	UNP Q5HLM5
B	722	LEU	-	EXPRESSION TAG	UNP Q5HLM5
B	723	GLU	-	EXPRESSION TAG	UNP Q5HLM5
B	724	HIS	-	EXPRESSION TAG	UNP Q5HLM5
B	725	HIS	-	EXPRESSION TAG	UNP Q5HLM5
B	726	HIS	-	EXPRESSION TAG	UNP Q5HLM5
B	727	HIS	-	EXPRESSION TAG	UNP Q5HLM5
B	728	HIS	-	EXPRESSION TAG	UNP Q5HLM5
B	729	HIS	-	EXPRESSION TAG	UNP Q5HLM5
C	444	ASN	HIS	ENGINEERED	UNP Q5HLM5
C	722	LEU	-	EXPRESSION TAG	UNP Q5HLM5
C	723	GLU	-	EXPRESSION TAG	UNP Q5HLM5

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	724	HIS	-	EXPRESSION TAG	UNP Q5HLM5
C	725	HIS	-	EXPRESSION TAG	UNP Q5HLM5
C	726	HIS	-	EXPRESSION TAG	UNP Q5HLM5
C	727	HIS	-	EXPRESSION TAG	UNP Q5HLM5
C	728	HIS	-	EXPRESSION TAG	UNP Q5HLM5
C	729	HIS	-	EXPRESSION TAG	UNP Q5HLM5
D	444	ASN	HIS	ENGINEERED	UNP Q5HLM5
D	722	LEU	-	EXPRESSION TAG	UNP Q5HLM5
D	723	GLU	-	EXPRESSION TAG	UNP Q5HLM5
D	724	HIS	-	EXPRESSION TAG	UNP Q5HLM5
D	725	HIS	-	EXPRESSION TAG	UNP Q5HLM5
D	726	HIS	-	EXPRESSION TAG	UNP Q5HLM5
D	727	HIS	-	EXPRESSION TAG	UNP Q5HLM5
D	728	HIS	-	EXPRESSION TAG	UNP Q5HLM5
D	729	HIS	-	EXPRESSION TAG	UNP Q5HLM5

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

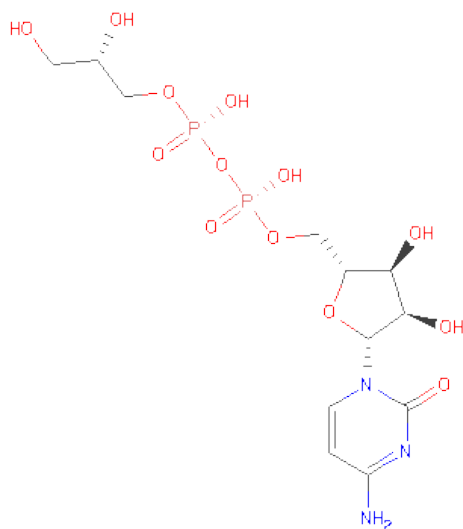
Continued from previous page...

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

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

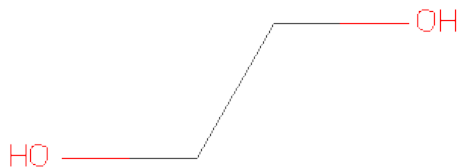
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	5	Total	Cl	0	0
			5	5		
3	A	8	Total	Cl	0	0
			8	8		
3	D	3	Total	Cl	0	0
			3	3		
3	C	5	Total	Cl	0	0
			5	5		

- Molecule 4 is [CYTIDINE-5'-PHOSPHATE]GLYCERYLPHOSPHORIC ACID ESTER (three-letter code: C2G) (formula: C<sub>12</sub>H<sub>21</sub>N<sub>3</sub>O<sub>13</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	D	1	Total	C	N	O	P	
			30	12	3	13	2	0
								0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is water.

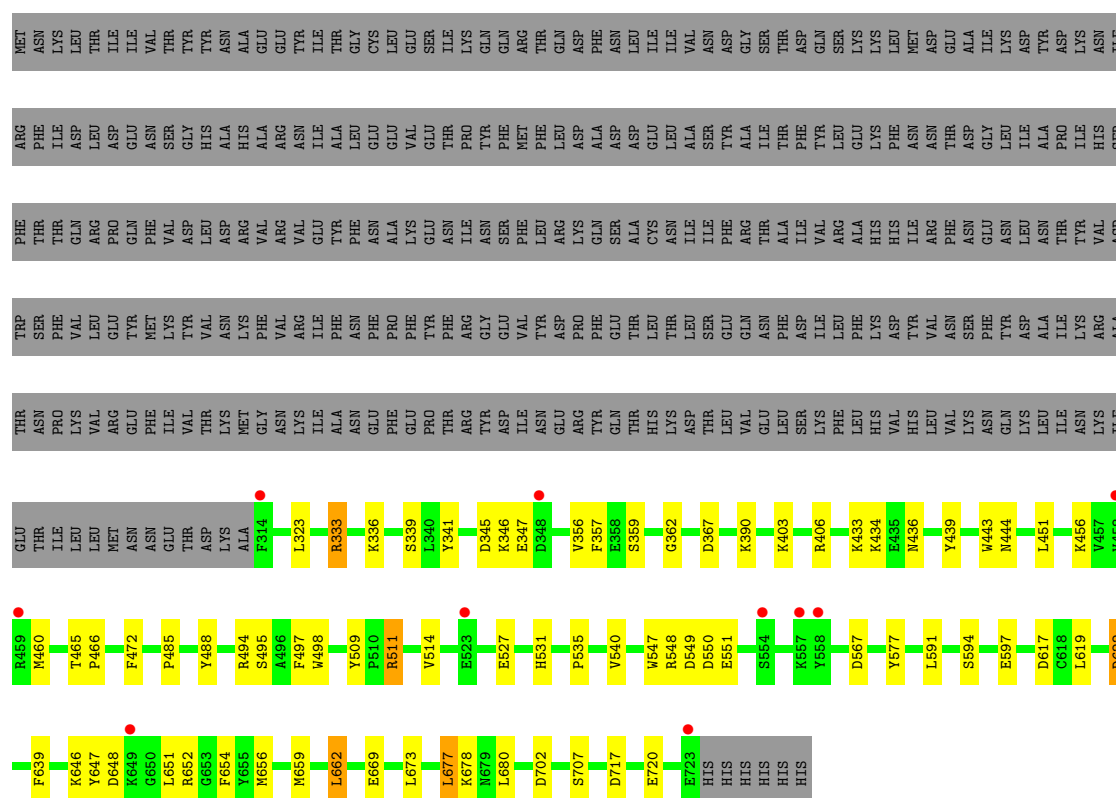
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	O	0	0
			2	2		
6	D	2	Total	O	0	0
			2	2		

### 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.

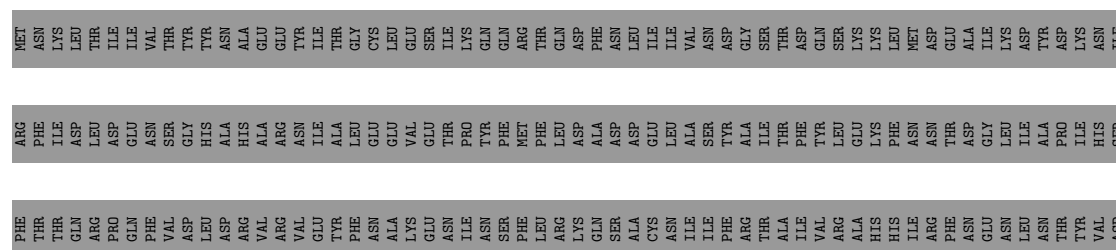
#### • Molecule 1: Teichoic acid biosynthesis protein F

Chain A: 



#### • Molecule 1: Teichoic acid biosynthesis protein F

Chain B: 









## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	222.88Å 222.88Å 100.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.01 – 2.95 20.01 – 2.95	Depositor EDS
% Data completeness (in resolution range)	98.7 (20.01-2.95) 98.7 (20.01-2.95)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.57 (at 2.93Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.193 , 0.254 0.188 , 0.249	Depositor DCC
$R_{free}$ test set	2646 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	88.4	Xtriage
Anisotropy	0.014	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 44.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 52937 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13868	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.04% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities 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: SO4, CL, C2G, EDO

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	0.52	0/3551	0.65	0/4798
1	B	0.52	0/3556	0.66	1/4805 (0.0%)
1	C	0.47	0/3556	0.64	0/4805
1	D	0.48	0/3469	0.62	0/4688
All	All	0.50	0/14132	0.64	1/19096 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	662	LEU	CA-CB-CG	5.66	128.32	115.30

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	3462	0	0	21	0
1	B	3467	0	0	23	0
1	C	3467	0	0	27	0
1	D	3383	0	0	24	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	10	0	0	0	0
2	B	10	0	0	0	0
2	C	10	0	0	0	0
3	A	8	0	0	0	0
3	B	5	0	0	2	0
3	C	5	0	0	0	0
3	D	3	0	0	1	0
4	D	30	0	19	2	0
5	D	4	0	6	0	0
6	A	2	0	0	0	0
6	D	2	0	0	0	0
All	All	13868	0	25	96	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 7.

All (96) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:347:GLU:O	1:B:436:ASN:ND2	2.20	0.74
1:D:547:TRP:CD1	1:D:547:TRP:N	2.59	0.69
1:B:382:ASN:ND2	1:B:382:ASN:N	2.46	0.64
1:C:662:LEU:CD2	1:C:662:LEU:N	2.61	0.64
1:D:325:HIS:CE1	1:D:329:ILE:CD1	2.81	0.62
1:A:333:ARG:NH1	1:A:333:ARG:CG	2.61	0.62
1:D:439:TYR:N	1:D:480:ASP:OD2	2.33	0.62
1:A:548:ARG:NH2	1:A:622:ASP:OD1	2.33	0.61
1:B:333:ARG:CG	1:B:333:ARG:NH1	2.64	0.61
1:A:662:LEU:N	1:A:662:LEU:CD2	2.66	0.58
1:D:410:GLU:OE1	1:D:410:GLU:N	2.37	0.57
1:B:475:GLU:OE1	1:B:478:ARG:NH1	2.38	0.57
1:D:353:LYS:CG	3:D:732:CL:CL	2.91	0.56
1:D:333:ARG:NH1	1:D:333:ARG:CG	2.68	0.56
1:D:470:ARG:NH1	1:D:470:ARG:CG	2.68	0.56
1:C:657:ASN:O	1:C:661:ASP:OD1	2.24	0.56
1:D:565:ASP:OD1	1:D:671:TYR:OH	2.26	0.54
1:C:467:LYS:NZ	1:C:470:ARG:NH2	2.56	0.54
1:B:461:PRO:O	1:B:463:THR:N	2.41	0.53
1:C:450:ARG:NH1	1:C:652:ARG:O	2.41	0.53
1:D:548:ARG:NH2	1:D:622:ASP:OD1	2.42	0.53
1:A:347:GLU:O	1:A:436:ASN:ND2	2.43	0.52
1:C:470:ARG:CG	1:C:470:ARG:NH1	2.71	0.52
1:C:444:ASN:O	1:C:511:ARG:NH1	2.43	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:380:TYR:CE1	1:C:715:HIS:ND1	2.77	0.52
1:A:547:TRP:CG	1:A:548:ARG:N	2.78	0.51
1:A:345:ASP:OD1	1:A:434:LYS:CE	2.58	0.51
1:B:470:ARG:NH1	1:B:470:ARG:CG	2.74	0.51
1:B:548:ARG:NH2	1:B:622:ASP:OD1	2.44	0.50
1:B:561:GLU:CB	3:B:734:CL:CL	2.97	0.49
1:D:367:ASP:CG	1:D:368:SER:N	2.65	0.49
1:D:434:LYS:C	1:D:436:ASN:N	2.65	0.49
1:B:714:ILE:O	1:B:718:ILE:CG2	2.60	0.48
1:B:403:LYS:CE	3:B:735:CL:CL	2.98	0.48
1:C:383:TYR:CZ	1:C:718:ILE:CD1	2.96	0.48
1:D:412:TYR:O	1:D:413:GLN:C	2.52	0.48
1:D:375:TYR:C	1:D:375:TYR:CD2	2.87	0.48
1:D:669:GLU:OE2	1:D:672:GLY:N	2.47	0.47
1:C:350:VAL:CG2	1:C:436:ASN:ND2	2.78	0.47
1:A:547:TRP:CD1	1:A:548:ARG:N	2.82	0.47
1:D:546:THR:CG2	1:D:547:TRP:N	2.77	0.47
1:C:559:LEU:CD1	1:C:559:LEU:C	2.83	0.47
1:B:560:PHE:O	1:B:561:GLU:C	2.53	0.46
1:C:621:THR:OG1	1:C:622:ASP:N	2.48	0.46
1:C:371:TYR:OH	1:C:608:ASP:OD1	2.33	0.46
1:A:639:PHE:CD2	1:A:639:PHE:N	2.84	0.46
1:A:549:ASP:O	1:A:551:GLU:N	2.49	0.45
1:B:329:ILE:CG2	1:B:330:VAL:N	2.78	0.45
1:B:695:TYR:CD2	1:B:695:TYR:C	2.90	0.45
1:A:362:GLY:O	1:A:390:LYS:N	2.50	0.44
1:A:497:PHE:O	1:A:498:TRP:C	2.55	0.44
1:A:443:TRP:CG	1:A:444:ASN:N	2.85	0.44
1:C:347:GLU:O	1:C:436:ASN:ND2	2.51	0.44
1:B:722:LEU:O	1:B:723:GLU:CG	2.66	0.44
1:C:475:GLU:OE1	1:C:478:ARG:NH1	2.50	0.44
1:C:661:ASP:N	1:C:661:ASP:OD1	2.50	0.44
1:D:447:PRO:CA	4:D:730:C2G:H1G2	2.48	0.44
1:A:444:ASN:O	1:A:511:ARG:NH1	2.50	0.43
1:B:411:TYR:CE1	1:B:415:TYR:CE2	3.05	0.43
1:B:570:TYR:O	1:B:574:GLY:N	2.51	0.43
1:D:317:ASN:ND2	1:D:320:ARG:NH2	2.65	0.43
1:C:367:ASP:O	1:C:370:LYS:N	2.51	0.43
1:A:357:PHE:CD2	1:A:357:PHE:N	2.87	0.43
1:B:718:ILE:O	1:B:718:ILE:CG1	2.67	0.43
1:B:642:TYR:CD1	1:B:642:TYR:N	2.87	0.43
1:C:380:TYR:CE1	1:C:715:HIS:CE1	3.06	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:493:PHE:O	1:C:494:ARG:C	2.57	0.43
4:D:730:C2G:HC6	4:D:730:C2G:O5'	2.19	0.42
1:C:346:LYS:C	1:C:348:ASP:N	2.73	0.42
1:C:718:ILE:O	1:C:718:ILE:CG1	2.67	0.42
1:D:470:ARG:NH1	1:D:471:ASN:OD1	2.52	0.42
1:D:450:ARG:CA	1:D:655:TYR:CE2	3.02	0.42
1:D:466:PRO:O	1:D:470:ARG:N	2.52	0.42
1:B:531:HIS:O	1:B:531:HIS:CD2	2.72	0.42
1:B:434:LYS:C	1:B:436:ASN:N	2.74	0.42
1:A:540:VAL:N	1:A:617:ASP:OD2	2.53	0.42
1:D:569:LEU:O	1:D:573:LEU:N	2.53	0.41
1:C:597:GLU:O	1:C:598:ASN:CB	2.68	0.41
1:A:465:THR:N	1:A:466:PRO:CD	2.83	0.41
1:B:380:TYR:N	1:B:381:PRO:CD	2.83	0.41
1:C:541:ILE:O	1:C:580:LEU:N	2.53	0.41
1:C:321:LYS:O	1:C:322:THR:C	2.57	0.41
1:A:527:GLU:O	1:A:531:HIS:N	2.54	0.41
1:D:652:ARG:C	1:D:654:PHE:N	2.74	0.41
1:B:413:GLN:O	1:B:417:GLU:CG	2.69	0.41
1:A:485:PRO:CG	1:A:509:TYR:CE2	3.03	0.41
1:C:380:TYR:CE1	1:C:715:HIS:CG	3.09	0.41
1:A:652:ARG:C	1:A:654:PHE:N	2.73	0.41
1:A:646:LYS:O	1:A:648:ASP:N	2.54	0.40
1:C:434:LYS:C	1:C:436:ASN:N	2.75	0.40
1:B:380:TYR:CE1	1:B:715:HIS:CG	3.10	0.40
1:D:638:PHE:CE2	1:D:677:LEU:CD1	3.04	0.40
1:A:577:TYR:OH	1:A:677:LEU:O	2.40	0.40
1:C:316:VAL:CG1	1:C:317:ASN:N	2.84	0.40
1:C:439:TYR:N	1:C:480:ASP:OD2	2.55	0.40
1:D:588:SER:OG	1:D:602:ASP:OD2	2.40	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	408/729 (56%)	376 (92%)	28 (7%)	4 (1%)	22	68
1	B	409/729 (56%)	362 (88%)	38 (9%)	9 (2%)	10	43
1	C	409/729 (56%)	356 (87%)	46 (11%)	7 (2%)	14	52
1	D	397/729 (54%)	346 (87%)	42 (11%)	9 (2%)	10	41
All	All	1623/2916 (56%)	1440 (89%)	154 (10%)	29 (2%)	13	50

All (29) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	651	LEU
1	C	367	ASP
1	D	367	ASP
1	A	647	TYR
1	B	462	GLY
1	C	647	TYR
1	C	648	ASP
1	D	332	ARG
1	D	521	ASP
1	D	655	TYR
1	A	550	ASP
1	B	345	ASP
1	B	461	PRO
1	B	561	GLU
1	B	649	LYS
1	D	515	LEU
1	C	565	ASP
1	D	458	VAL
1	B	443	TRP
1	B	459	ARG
1	B	565	ASP
1	C	535	PRO
1	C	614	LEU
1	D	443	TRP
1	C	452	ALA
1	B	603	VAL
1	D	545	PRO
1	D	603	VAL
1	A	535	PRO



### 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	379/675 (56%)	339 (89%)	40 (11%)	10	34
1	B	379/675 (56%)	328 (86%)	51 (14%)	6	22
1	C	379/675 (56%)	341 (90%)	38 (10%)	11	38
1	D	370/675 (55%)	329 (89%)	41 (11%)	9	32
All	All	1507/2700 (56%)	1337 (89%)	170 (11%)	9	31

All (170) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	323	LEU
1	A	333	ARG
1	A	336	LYS
1	A	339	SER
1	A	341	TYR
1	A	346	LYS
1	A	356	VAL
1	A	359	SER
1	A	367	ASP
1	A	403	LYS
1	A	406	ARG
1	A	433	LYS
1	A	439	TYR
1	A	451	LEU
1	A	456	LYS
1	A	460	MET
1	A	472	PHE
1	A	488	TYR
1	A	494	ARG
1	A	495	SER
1	A	511	ARG
1	A	514	VAL
1	A	567	ASP
1	A	591	LEU
1	A	594	SER
1	A	597	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	619	LEU
1	A	622	ASP
1	A	656	MET
1	A	659	MET
1	A	662	LEU
1	A	669	GLU
1	A	673	LEU
1	A	677	LEU
1	A	678	LYS
1	A	680	LEU
1	A	702	ASP
1	A	707	SER
1	A	717	ASP
1	A	720	GLU
1	B	316	VAL
1	B	323	LEU
1	B	329	ILE
1	B	333	ARG
1	B	343	LEU
1	B	356	VAL
1	B	367	ASP
1	B	382	ASN
1	B	402	GLU
1	B	404	VAL
1	B	407	ASN
1	B	433	LYS
1	B	439	TYR
1	B	449	LYS
1	B	450	ARG
1	B	451	LEU
1	B	456	LYS
1	B	470	ARG
1	B	472	PHE
1	B	477	SER
1	B	487	ARG
1	B	491	GLU
1	B	514	VAL
1	B	533	ASN
1	B	548	ARG
1	B	549	ASP
1	B	559	LEU
1	B	563	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	578	VAL
1	B	579	ILE
1	B	582	ARG
1	B	589	ASN
1	B	597	GLU
1	B	603	VAL
1	B	607	ASN
1	B	609	VAL
1	B	622	ASP
1	B	645	ASP
1	B	647	TYR
1	B	651	LEU
1	B	661	ASP
1	B	662	LEU
1	B	668	THR
1	B	673	LEU
1	B	677	LEU
1	B	678	LYS
1	B	702	ASP
1	B	703	ASN
1	B	717	ASP
1	B	718	ILE
1	B	722	LEU
1	C	315	LYS
1	C	323	LEU
1	C	332	ARG
1	C	339	SER
1	C	367	ASP
1	C	382	ASN
1	C	403	LYS
1	C	439	TYR
1	C	450	ARG
1	C	451	LEU
1	C	457	VAL
1	C	463	THR
1	C	470	ARG
1	C	472	PHE
1	C	487	ARG
1	C	500	ASP
1	C	536	SER
1	C	559	LEU
1	C	589	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	591	LEU
1	C	597	GLU
1	C	603	VAL
1	C	622	ASP
1	C	623	TYR
1	C	648	ASP
1	C	656	MET
1	C	660	GLU
1	C	661	ASP
1	C	662	LEU
1	C	669	GLU
1	C	673	LEU
1	C	681	ASP
1	C	682	LYS
1	C	683	VAL
1	C	707	SER
1	C	714	ILE
1	C	718	ILE
1	C	722	LEU
1	D	315	LYS
1	D	316	VAL
1	D	317	ASN
1	D	323	LEU
1	D	333	ARG
1	D	337	GLU
1	D	353	LYS
1	D	356	VAL
1	D	367	ASP
1	D	388	SER
1	D	403	LYS
1	D	435	GLU
1	D	439	TYR
1	D	450	ARG
1	D	451	LEU
1	D	456	LYS
1	D	460	MET
1	D	470	ARG
1	D	472	PHE
1	D	487	ARG
1	D	489	SER
1	D	500	ASP
1	D	511	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	514	VAL
1	D	518	ARG
1	D	547	TRP
1	D	548	ARG
1	D	576	ASP
1	D	585	TYR
1	D	609	VAL
1	D	622	ASP
1	D	623	TYR
1	D	625	SER
1	D	645	ASP
1	D	647	TYR
1	D	660	GLU
1	D	662	LEU
1	D	669	GLU
1	D	673	LEU
1	D	702	ASP
1	D	718	ILE

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

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 29 ligands modelled in this entry, 21 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
2	SO4	A	730	-	4,4,4	0.18	0	6,6,6	0.21	0
2	SO4	A	731	-	4,4,4	0.18	0	6,6,6	0.10	0
2	SO4	B	730	-	4,4,4	0.09	0	6,6,6	0.29	0
2	SO4	B	731	-	4,4,4	0.13	0	6,6,6	0.14	0
2	SO4	C	730	-	4,4,4	0.17	0	6,6,6	0.17	0
2	SO4	C	731	-	4,4,4	0.22	0	6,6,6	0.21	0
4	C2G	D	730	-	31,31,31	4.28	13 (41%)	43,46,46	1.59	6 (13%)
5	EDO	D	734	-	3,3,3	0.74	0	2,2,2	0.13	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	730	-	-	0/0/0/0	0/0/0/0
2	SO4	A	731	-	-	0/0/0/0	0/0/0/0
2	SO4	B	730	-	-	0/0/0/0	0/0/0/0
2	SO4	B	731	-	-	0/0/0/0	0/0/0/0
2	SO4	C	730	-	-	0/0/0/0	0/0/0/0
2	SO4	C	731	-	-	0/0/0/0	0/0/0/0
4	C2G	D	730	-	-	0/22/40/40	0/2/2/2
5	EDO	D	734	-	-	0/1/1/1	0/0/0/0

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	730	C2G	C2-N1	15.02	1.54	1.38
4	D	730	C2G	O2-C2	7.99	1.40	1.23
4	D	730	C2G	C6-N1	7.21	1.47	1.35
4	D	730	C2G	C4-N3	6.66	1.48	1.35
4	D	730	C2G	C6-C5	5.67	1.51	1.38
4	D	730	C2G	C2-N3	5.54	1.50	1.35
4	D	730	C2G	O4'-C1'	4.50	1.48	1.41
4	D	730	C2G	C5-C4	4.35	1.51	1.40
4	D	730	C2G	PB-O1B	3.85	1.66	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	730	C2G	PA-O1A	3.80	1.65	1.51
4	D	730	C2G	PB-O3A	3.38	1.66	1.59
4	D	730	C2G	C4-N4	3.33	1.45	1.35
4	D	730	C2G	PA-O3A	3.27	1.65	1.59

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	730	C2G	C6-C5-C4	5.42	119.72	117.47
4	D	730	C2G	C2-N1-C1'	4.03	124.52	119.03
4	D	730	C2G	C2-N3-C4	3.74	120.98	115.57
4	D	730	C2G	PB-O3A-PA	-3.18	122.36	131.68
4	D	730	C2G	C3'-C2'-C1'	3.10	105.76	100.91
4	D	730	C2G	N4-C4-N3	2.82	122.00	116.59

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, 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	410/729 (56%)	-0.14	10 (2%) 56 25	61, 83, 149, 220	0
1	B	411/729 (56%)	-0.06	17 (4%) 35 17	64, 85, 159, 242	0
1	C	411/729 (56%)	-0.01	17 (4%) 35 17	69, 94, 166, 263	0
1	D	401/729 (55%)	0.04	23 (5%) 23 12	66, 92, 151, 279	0
All	All	1633/2916 (56%)	-0.04	67 (4%) 34 17	61, 88, 156, 279	0

All (67) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	652	ARG	7.7
1	C	558	TYR	5.8
1	D	649	LYS	5.5
1	D	590	ALA	5.3
1	C	555	LYS	5.2
1	D	650	GLY	5.0
1	D	724	HIS	5.0
1	A	649	LYS	4.8
1	D	652	ARG	4.7
1	B	555	LYS	4.4
1	C	652	ARG	4.2
1	C	457	VAL	3.8
1	C	556	GLY	3.6
1	C	554	SER	3.6
1	B	649	LYS	3.5
1	D	723	GLU	3.4
1	C	560	PHE	3.4
1	C	557	LYS	3.4
1	B	560	PHE	3.3
1	C	459	ARG	3.2
1	B	558	TYR	3.2

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	557	LYS	3.1
1	B	554	SER	2.9
1	B	459	ARG	2.9
1	B	313	ALA	2.9
1	D	651	LEU	2.8
1	A	458	VAL	2.8
1	C	553	VAL	2.8
1	D	356	VAL	2.8
1	D	484	SER	2.7
1	C	313	ALA	2.7
1	D	547	TRP	2.7
1	D	721	GLN	2.7
1	B	720	GLU	2.7
1	D	491	GLU	2.7
1	A	459	ARG	2.7
1	C	484	SER	2.6
1	B	537	ASP	2.6
1	C	491	GLU	2.5
1	C	720	GLU	2.5
1	B	356	VAL	2.5
1	D	585	TYR	2.4
1	A	314	PHE	2.4
1	C	721	GLN	2.4
1	D	536	SER	2.4
1	A	557	LYS	2.3
1	C	435	GLU	2.3
1	A	723	GLU	2.3
1	A	348	ASP	2.3
1	D	494	ARG	2.3
1	D	462	GLY	2.3
1	A	523	GLU	2.2
1	D	457	VAL	2.2
1	D	463	THR	2.2
1	D	461	PRO	2.2
1	D	458	VAL	2.2
1	A	558	TYR	2.2
1	D	561	GLU	2.2
1	B	462	GLY	2.2
1	B	647	TYR	2.2
1	C	645	ASP	2.2
1	B	457	VAL	2.2
1	A	554	SER	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	561	GLU	2.1
1	D	591	LEU	2.1
1	D	587	ILE	2.1
1	B	648	ASP	2.1

## 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 ⓘ

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, 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	CL	A	733	1/1	0.69	25.45	130,130,130,130	0
5	EDO	D	734	4/4	0.39	12.57	90,90,90,90	0
3	CL	B	735	1/1	0.67	9.52	127,127,127,127	0
3	CL	B	732	1/1	0.47	6.75	100,100,100,100	0
3	CL	C	734	1/1	0.49	5.64	106,106,106,106	0
3	CL	A	734	1/1	0.23	2.82	96,96,96,96	0
3	CL	A	736	1/1	0.28	2.58	106,106,106,106	0
3	CL	B	733	1/1	0.38	2.53	121,121,121,121	0
3	CL	A	739	1/1	0.21	1.87	88,88,88,88	0
3	CL	C	735	1/1	0.27	1.68	118,118,118,118	0
2	SO4	A	731	5/5	0.25	1.64	166,166,166,166	0
3	CL	C	736	1/1	0.22	1.60	106,106,106,106	0
2	SO4	B	731	5/5	0.26	1.43	171,171,171,171	0
3	CL	A	735	1/1	0.28	0.70	94,94,94,94	0
4	C2G	D	730	30/30	0.25	0.69	122,122,123,123	0
3	CL	A	732	1/1	0.23	0.23	114,114,114,114	0
3	CL	D	733	1/1	0.19	0.21	92,92,92,92	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	C	733	1/1	0.20	0.15	102,102,102,102	0
3	CL	B	734	1/1	0.21	0.01	111,111,111,111	0
2	SO4	A	730	5/5	0.17	-0.02	103,104,105,107	0
3	CL	A	738	1/1	0.20	-0.06	97,97,97,97	0
3	CL	C	732	1/1	0.20	-0.15	107,107,107,107	0
2	SO4	C	731	5/5	0.15	-0.47	124,125,127,128	0
3	CL	A	737	1/1	0.18	-0.53	108,108,108,108	0
3	CL	D	732	1/1	0.13	-1.21	101,101,101,101	0
2	SO4	C	730	5/5	0.16	-1.46	129,129,131,132	0
2	SO4	B	730	5/5	0.10	-1.52	90,90,91,94	0
3	CL	B	736	1/1	0.09	-2.52	83,83,83,83	0
3	CL	D	731	1/1	0.06	-3.24	68,68,68,68	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.