



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 27, 2014 – 11:26 PM GMT

PDB ID : 2VKD  
Title : CRYSTAL STRUCTURE OF THE CATALYTIC DOMAIN OF LETHAL TOXIN FROM CLOSTRIDIUM SORDELLII IN COMPLEX WITH UDP-GLC AND MANGANESE ION  
Authors : Ziegler, M.O.P.; Jank, T.; Aktories, K.; Schulz, G.E.  
Deposited on : 2007-12-18  
Resolution : 2.53 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>

---

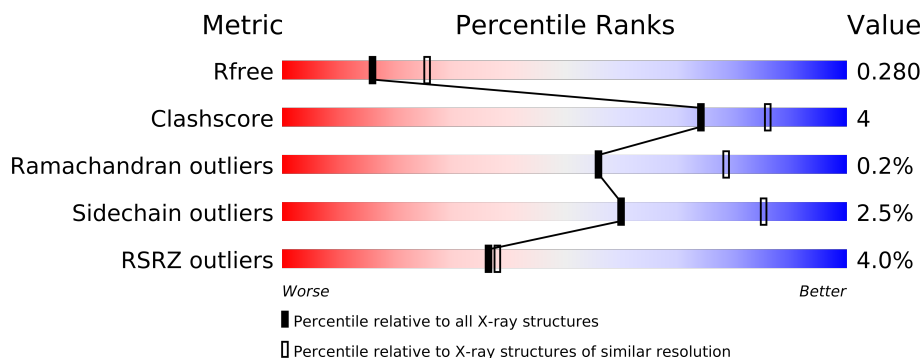
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.53 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	3240 (2.54-2.50)
Clashscore	79885	4080 (2.54-2.50)
Ramachandran outliers	78287	3990 (2.54-2.50)
Sidechain outliers	78261	3992 (2.54-2.50)
RSRZ outliers	66119	3241 (2.54-2.50)

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	546	
1	B	546	
1	C	546	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 13572 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 CYTOTOXIN L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	542	Total	C	N	O	S	0	0	0
			4447	2838	721	872	16			
1	B	538	Total	C	N	O	S	0	0	0
			4422	2824	716	867	15			
1	C	541	Total	C	N	O	S	0	0	0
			4442	2835	720	871	16			

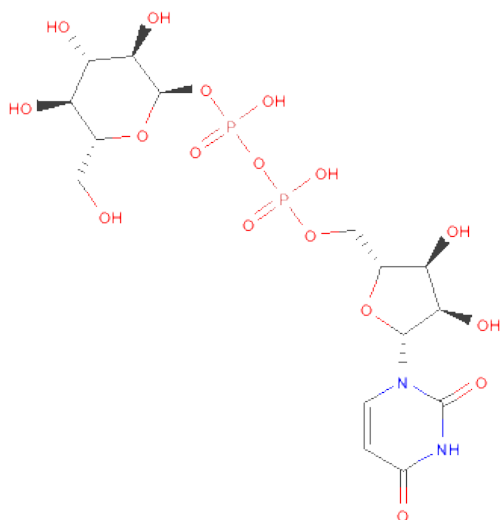
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	13	ALA	VAL	ENGINEERED MUTATION	UNP Q46342
A	289	MET	ILE	ENGINEERED MUTATION	UNP Q46342
B	13	ALA	VAL	ENGINEERED MUTATION	UNP Q46342
B	289	MET	ILE	ENGINEERED MUTATION	UNP Q46342
C	13	ALA	VAL	ENGINEERED MUTATION	UNP Q46342
C	289	MET	ILE	ENGINEERED MUTATION	UNP Q46342

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mn	0	0
			1	1		
2	A	3	Total	Mn	0	0
			3	3		
2	C	1	Total	Mn	0	0
			1	1		

- Molecule 3 is URIDINE-5'-DIPHOSPHATE-GLUCOSE (three-letter code: UPG) (formula: C<sub>15</sub>H<sub>24</sub>N<sub>2</sub>O<sub>17</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	B	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	C	1	Total	C	N	O	P	0	0
			36	15	2	17	2		

- Molecule 4 is water.

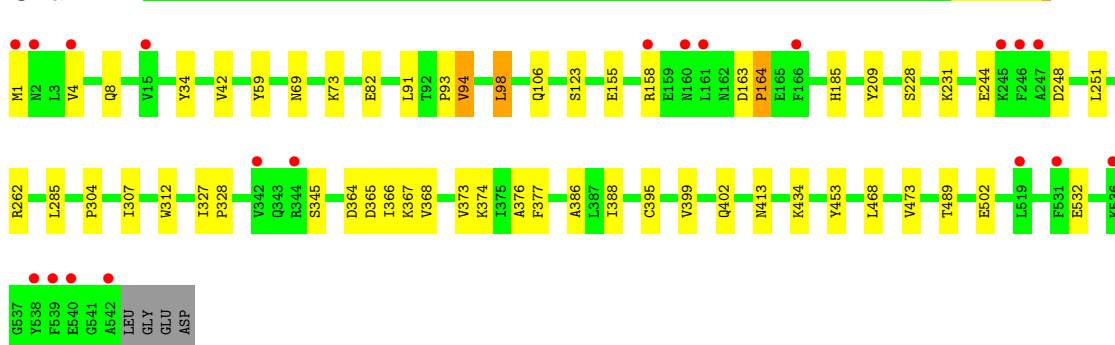
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	47	Total	O	0	0
			47	47		
4	B	57	Total	O	0	0
			57	57		
4	C	44	Total	O	0	0
			44	44		

### 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.

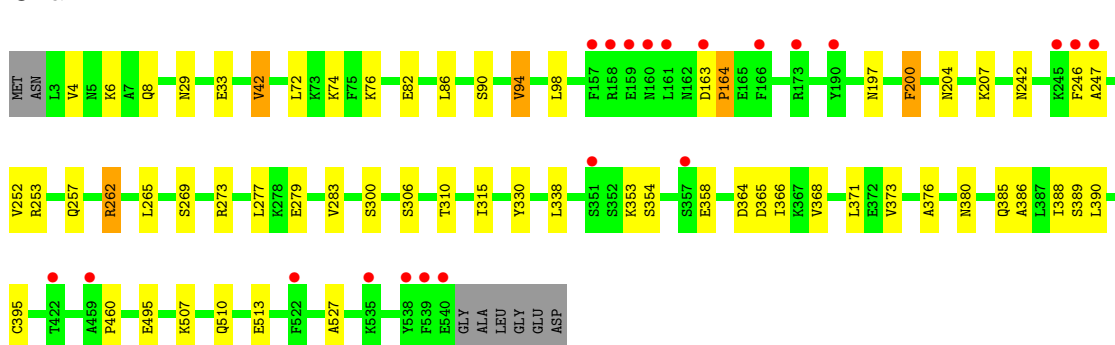
#### • Molecule 1: CYTOTOXIN L

Chain A:



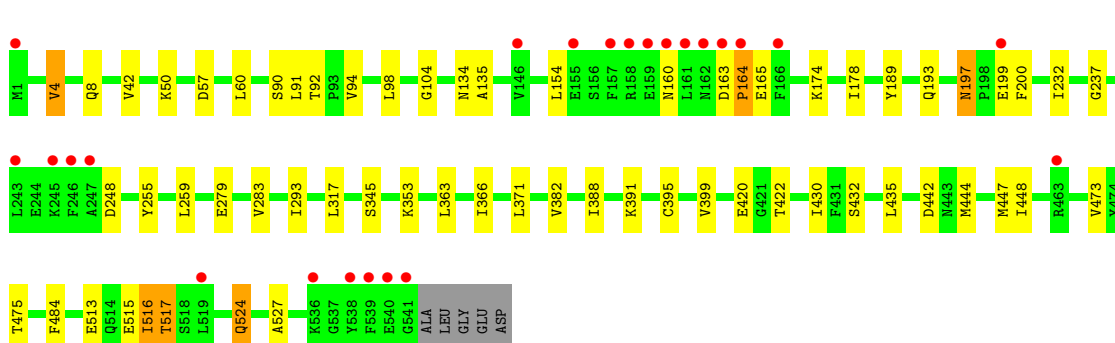
#### • Molecule 1: CYTOTOXIN L

Chain B:



#### • Molecule 1: CYTOTOXIN L

Chain C:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.29Å 190.32Å 204.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	90.17 – 2.53 90.05 – 2.53	Depositor EDS
% Data completeness (in resolution range)	99.7 (90.17-2.53) 99.7 (90.05-2.53)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.46 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.244 , 0.288 0.240 , 0.280	Depositor DCC
$R_{free}$ test set	3780 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.2	Xtriage
Anisotropy	0.705	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 15.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.22$	Xtriage
Outliers	2 of 75590 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	13572	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	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 35.55 % 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 5.7375e-04.*

<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: UPG, MN

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	0.42	0/4530	0.54	0/6114
1	B	0.44	0/4505	0.55	0/6081
1	C	0.43	0/4525	0.54	0/6107
All	All	0.43	0/13560	0.54	0/18302

There are no bond length outliers.

There are no bond angle outliers.

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	4447	0	4406	34	0
1	B	4422	0	4380	37	0
1	C	4442	0	4401	37	0
2	A	3	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	36	0	22	0	0
3	B	36	0	22	0	0
3	C	36	0	22	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	47	0	0	0	0
4	B	57	0	0	1	0
4	C	44	0	0	1	0
All	All	13572	0	13253	108	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 4.

All (108) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:163:ASP:HB2	1:A:164:PRO:HA	1.39	1.03
1:B:262:ARG:HH11	1:B:262:ARG:HG3	1.18	1.02
1:A:106:GLN:HE21	1:A:231:LYS:HD2	1.35	0.91
1:B:262:ARG:NH1	1:B:262:ARG:HG3	1.85	0.89
1:A:4:VAL:HG22	1:A:8:GLN:HB2	1.62	0.81
1:C:366:ILE:HG21	1:C:388:ILE:HD13	1.65	0.78
1:C:163:ASP:HB3	1:C:164:PRO:HA	1.67	0.74
1:C:4:VAL:HG22	1:C:8:GLN:HB3	1.68	0.73
1:A:98:LEU:HD21	1:A:285:LEU:HD12	1.70	0.72
1:A:1:MET:HB2	1:C:422:THR:HG21	1.71	0.71
1:B:262:ARG:HH11	1:B:262:ARG:CG	2.02	0.69
1:C:513:GLU:O	1:C:516:ILE:HG12	1.95	0.67
1:C:42:VAL:HG21	1:C:90:SER:CB	2.27	0.65
1:C:160:ASN:HD22	1:C:164:PRO:HD3	1.60	0.65
1:B:204:ASN:HA	1:B:207:LYS:HE2	1.77	0.65
1:A:4:VAL:HG22	1:A:8:GLN:CB	2.25	0.65
1:A:163:ASP:HB2	1:A:164:PRO:CA	2.22	0.64
1:C:4:VAL:HG22	1:C:8:GLN:CB	2.28	0.64
1:C:371:LEU:O	1:C:395:CYS:HB2	2.00	0.62
1:A:304:PRO:HG2	1:A:307:ILE:HD12	1.81	0.61
1:B:253:ARG:NH2	4:B:2027:HOH:O	2.34	0.60
1:A:373:VAL:HG23	1:A:395:CYS:HB3	1.83	0.60
1:C:524:GLN:HA	1:C:527:ALA:HB3	1.85	0.59
1:C:363:LEU:HD22	1:C:366:ILE:HD11	1.85	0.58
1:C:165:GLU:HA	1:C:165:GLU:OE1	2.04	0.57
1:A:304:PRO:HD2	1:A:312:TRP:CZ3	2.39	0.57
1:B:371:LEU:O	1:B:395:CYS:HB2	2.05	0.56
1:C:42:VAL:HG21	1:C:90:SER:HB2	1.87	0.56
1:A:93:PRO:HA	1:A:366:ILE:O	2.06	0.55
1:B:385:GLN:NE2	1:B:510:GLN:HE21	2.03	0.55
3:C:1544:UPG:H3'	3:C:1544:UPG:H5C2	1.88	0.55
1:B:98:LEU:HD23	1:B:283:VAL:HB	1.87	0.55

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:155:GLU:HB3	1:A:158:ARG:HH21	1.71	0.54
1:C:399:VAL:HG13	1:C:473:VAL:HG12	1.90	0.54
1:B:29:ASN:O	1:B:33:GLU:HG2	2.08	0.54
1:A:185:HIS:HD2	1:A:209:TYR:OH	1.91	0.54
1:A:413:ASN:HD22	1:A:434:LYS:HB3	1.72	0.53
1:B:163:ASP:HB2	1:B:164:PRO:HA	1.91	0.53
1:C:42:VAL:HG21	1:C:90:SER:HB3	1.90	0.53
1:A:366:ILE:HG21	1:A:388:ILE:HD13	1.93	0.50
1:A:399:VAL:HG13	1:A:473:VAL:HG12	1.94	0.49
1:B:380:ASN:ND2	1:B:495:GLU:OE2	2.45	0.49
1:A:91:LEU:HB3	1:A:367:LYS:HB3	1.94	0.49
1:A:262:ARG:HD2	1:A:453:TYR:CE2	2.46	0.49
1:B:373:VAL:HG23	1:B:395:CYS:HB3	1.95	0.49
1:C:420:GLU:HG3	1:C:430:ILE:HD13	1.94	0.48
1:C:255:TYR:CE2	1:C:259:LEU:HD22	2.47	0.48
1:B:273:ARG:O	1:B:277:LEU:HB2	2.13	0.48
1:B:94:VAL:HG23	1:B:366:ILE:HG13	1.95	0.48
1:B:385:GLN:HE21	1:B:510:GLN:HE21	1.62	0.47
1:B:72:LEU:O	1:B:76:LYS:HG3	2.15	0.47
1:A:374:LYS:HB2	1:A:388:ILE:HB	1.96	0.47
1:B:376:ALA:HB2	1:B:386:ALA:HB3	1.96	0.47
1:B:353:LYS:HD2	1:B:358:GLU:HB3	1.97	0.47
1:B:4:VAL:CG2	1:B:8:GLN:HB3	2.45	0.47
1:B:330:TYR:HA	1:B:513:GLU:OE1	2.14	0.47
1:A:248:ASP:HB3	1:A:251:LEU:HB2	1.97	0.46
1:A:69:ASN:O	1:A:73:LYS:HG2	2.14	0.46
1:C:279:GLU:HB3	4:C:2024:HOH:O	2.16	0.46
1:B:373:VAL:HG22	1:B:389:SER:HB2	1.98	0.46
1:C:197:ASN:HB3	1:C:200:PHE:HD2	1.80	0.46
1:C:524:GLN:H	1:C:524:GLN:HG2	1.56	0.46
1:A:163:ASP:CB	1:A:164:PRO:HA	2.26	0.46
1:C:104:GLY:HA2	1:C:135:ALA:O	2.16	0.45
1:A:106:GLN:HG3	1:A:228:SER:HA	1.99	0.45
1:C:399:VAL:HG13	1:C:473:VAL:CG1	2.46	0.45
1:C:197:ASN:HD22	1:C:199:GLU:H	1.64	0.45
1:C:435:LEU:HD22	1:C:447:MET:SD	2.57	0.45
1:B:246:PHE:HA	1:B:279:GLU:HG3	1.98	0.45
1:B:247:ALA:HB1	1:B:252:VAL:HG21	1.99	0.44
1:A:94:VAL:HG13	1:A:368:VAL:CG1	2.47	0.44
1:B:247:ALA:HB1	1:B:252:VAL:CG2	2.47	0.44
1:A:376:ALA:HB2	1:A:386:ALA:HB3	1.98	0.44
1:C:382:VAL:HG21	1:C:475:THR:HG23	1.99	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:265:LEU:O	1:B:269:SER:HB2	2.18	0.44
1:C:524:GLN:HA	1:C:527:ALA:CB	2.48	0.43
1:B:94:VAL:HA	1:B:390:LEU:HD23	2.00	0.43
1:A:395:CYS:O	1:A:399:VAL:HG23	2.18	0.43
1:A:59:TYR:CD2	1:A:59:TYR:C	2.91	0.43
1:A:402:GLN:HB2	1:A:473:VAL:HG13	2.01	0.43
1:B:42:VAL:HG21	1:B:90:SER:HB2	2.00	0.43
1:B:197:ASN:HB3	1:B:200:PHE:CD2	2.54	0.43
1:B:246:PHE:CD1	1:B:279:GLU:HB3	2.54	0.43
1:A:364:ASP:HB3	1:A:365:ASP:H	1.67	0.42
1:A:106:GLN:HE21	1:A:231:LYS:CD	2.17	0.42
1:A:327:ILE:HA	1:A:328:PRO:HD3	1.88	0.42
1:C:232:ILE:HG13	1:C:237:GLY:HA3	2.01	0.42
1:C:444:MET:O	1:C:448:ILE:HG12	2.19	0.42
1:A:34:TYR:OH	1:A:82:GLU:OE2	2.29	0.42
1:B:82:GLU:O	1:B:82:GLU:HG3	2.20	0.42
1:B:366:ILE:HD12	1:B:388:ILE:HD13	2.00	0.42
1:C:164:PRO:HB2	1:C:165:GLU:H	1.61	0.41
1:C:174:LYS:O	1:C:178:ILE:HG12	2.19	0.41
1:B:315:ILE:HD13	1:B:338:LEU:HD11	2.02	0.41
1:B:283:VAL:CG2	1:B:366:ILE:HD11	2.50	0.41
1:C:57:ASP:HA	1:C:60:LEU:HD12	2.02	0.41
1:C:279:GLU:O	1:C:391:LYS:HE2	2.21	0.41
1:A:377:PHE:CE1	1:A:502:GLU:HB3	2.56	0.41
1:C:293:ILE:HD11	1:C:317:LEU:HD23	2.02	0.41
1:B:460:PRO:HB3	1:B:527:ALA:HB2	2.03	0.41
1:B:94:VAL:CG2	1:B:366:ILE:HG13	2.51	0.40
1:A:106:GLN:NE2	1:A:231:LYS:HD2	2.18	0.40
1:C:50:LYS:HA	1:C:484:PHE:CD1	2.56	0.40
1:C:189:TYR:O	1:C:193:GLN:HG2	2.21	0.40
1:B:74:LYS:HA	1:B:74:LYS:HD2	1.72	0.40
1:B:364:ASP:HB3	1:B:365:ASP:H	1.71	0.40
1:C:515:GLU:C	1:C:517:THR:H	2.25	0.40
1:C:98:LEU:HD12	1:C:283:VAL:HB	2.02	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	540/546 (99%)	513 (95%)	26 (5%)	1 (0%)	56	78
1	B	536/546 (98%)	507 (95%)	28 (5%)	1 (0%)	56	78
1	C	539/546 (99%)	505 (94%)	32 (6%)	2 (0%)	43	66
All	All	1615/1638 (99%)	1525 (94%)	86 (5%)	4 (0%)	56	78

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	164	PRO
1	B	164	PRO
1	C	164	PRO
1	C	516	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	502/505 (99%)	493 (98%)	9 (2%)	71	91
1	B	500/505 (99%)	486 (97%)	14 (3%)	56	82
1	C	502/505 (99%)	488 (97%)	14 (3%)	56	82
All	All	1504/1515 (99%)	1467 (98%)	37 (2%)	60	85

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	42	VAL
1	A	94	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	98	LEU
1	A	123	SER
1	A	244	GLU
1	A	345	SER
1	A	468	LEU
1	A	489	THR
1	A	532	GLU
1	B	6	LYS
1	B	42	VAL
1	B	86	LEU
1	B	94	VAL
1	B	200	PHE
1	B	242	ASN
1	B	257	GLN
1	B	262	ARG
1	B	300	SER
1	B	306	SER
1	B	310	THR
1	B	354	SER
1	B	368	VAL
1	B	507	LYS
1	C	4	VAL
1	C	91	LEU
1	C	92	THR
1	C	94	VAL
1	C	134	ASN
1	C	154	LEU
1	C	197	ASN
1	C	248	ASP
1	C	345	SER
1	C	353	LYS
1	C	432	SER
1	C	442	ASP
1	C	517	THR
1	C	524	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (34) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	106	GLN
1	A	113	ASN
1	A	160	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	185	HIS
1	A	238	ASN
1	A	343	GLN
1	A	413	ASN
1	A	415	ASN
1	B	53	ASN
1	B	88	ASN
1	B	89	ASN
1	B	106	GLN
1	B	139	ASN
1	B	257	GLN
1	B	309	ASN
1	B	343	GLN
1	B	385	GLN
1	B	415	ASN
1	B	492	HIS
1	B	523	ASN
1	C	10	GLN
1	C	69	ASN
1	C	134	ASN
1	C	151	ASN
1	C	160	ASN
1	C	197	ASN
1	C	238	ASN
1	C	256	ASN
1	C	302	ASN
1	C	385	GLN
1	C	415	ASN
1	C	467	ASN
1	C	510	GLN
1	C	524	GLN

### 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 8 ligands modelled in this entry, 5 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
3	UPG	A	1544	2	38,38,38	0.99	3 (7%)	54,58,58	1.75	9 (16%)
3	UPG	B	1544	2	38,38,38	0.93	3 (7%)	54,58,58	1.42	7 (12%)
3	UPG	C	1544	2	38,38,38	0.97	2 (5%)	54,58,58	1.61	8 (14%)

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
3	UPG	A	1544	2	-	0/21/59/59	0/3/3/3
3	UPG	B	1544	2	-	0/21/59/59	0/3/3/3
3	UPG	C	1544	2	-	0/21/59/59	0/3/3/3

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1544	UPG	C2-N1	4.06	1.42	1.38
3	C	1544	UPG	C2-N1	3.81	1.42	1.38
3	B	1544	UPG	C2-N1	2.57	1.41	1.38
3	B	1544	UPG	C6-C5	2.39	1.39	1.36
3	A	1544	UPG	C6-C5	2.36	1.39	1.36
3	C	1544	UPG	C6-C5	2.33	1.39	1.36
3	A	1544	UPG	O4C-C1C	2.15	1.44	1.41
3	B	1544	UPG	O4C-C1C	2.09	1.44	1.41

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1544	UPG	C2-N1-C1C	6.51	122.29	118.21
3	A	1544	UPG	N3-C2-N1	5.93	120.92	115.97
3	B	1544	UPG	N3-C2-N1	5.69	120.72	115.97
3	C	1544	UPG	N3-C2-N1	5.42	120.50	115.97
3	C	1544	UPG	C2-N1-C1C	4.35	120.93	118.21
3	A	1544	UPG	C3'-C4'-C5'	3.52	116.49	110.20
3	C	1544	UPG	O5'-C5'-C4'	3.47	116.18	109.76
3	C	1544	UPG	O5'-C1'-O3B	-3.45	106.85	111.36
3	B	1544	UPG	O5'-C5'-C4'	3.42	116.09	109.76
3	B	1544	UPG	C3'-C4'-C5'	3.26	116.03	110.20
3	A	1544	UPG	O5'-C5'-C4'	3.21	115.70	109.76
3	C	1544	UPG	C3'-C4'-C5'	3.10	115.74	110.20
3	A	1544	UPG	C6'-C5'-C4'	-3.09	105.52	113.00
3	C	1544	UPG	C6'-C5'-C4'	-2.97	105.81	113.00
3	C	1544	UPG	PB-O3A-PA	-2.92	123.11	131.68
3	A	1544	UPG	PB-O3A-PA	-2.83	123.39	131.68
3	B	1544	UPG	C2-N1-C1C	2.73	119.92	118.21
3	C	1544	UPG	C1'-O5'-C5'	2.44	118.46	113.73
3	B	1544	UPG	O5'-C1'-C2'	-2.40	105.39	110.31
3	B	1544	UPG	O4C-C1C-N1	2.28	112.87	108.06
3	A	1544	UPG	C1'-O5'-C5'	2.27	118.14	113.73
3	A	1544	UPG	C3C-C2C-C1C	2.18	104.32	100.91
3	B	1544	UPG	PB-O3A-PA	-2.18	125.29	131.68
3	A	1544	UPG	O4C-C1C-N1	2.02	112.31	108.06

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	542/546 (99%)	0.31	20 (3%) 39 41	12, 29, 59, 87	0
1	B	538/546 (98%)	0.25	21 (3%) 37 39	9, 28, 72, 99	0
1	C	541/546 (99%)	0.34	24 (4%) 33 35	12, 30, 71, 89	0
All	All	1621/1638 (98%)	0.30	65 (4%) 36 38	9, 29, 65, 99	0

All (65) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	MET	8.7
1	C	247	ALA	7.3
1	B	539	PHE	7.3
1	C	166	PHE	6.9
1	B	538	TYR	6.7
1	C	161	LEU	6.3
1	C	157	PHE	6.1
1	A	540	GLU	5.6
1	A	539	PHE	5.5
1	B	540	GLU	5.1
1	C	1	MET	4.8
1	C	538	TYR	4.7
1	C	243	LEU	4.7
1	B	246	PHE	4.6
1	C	541	GLY	4.6
1	C	539	PHE	4.6
1	A	542	ALA	4.2
1	B	158	ARG	4.2
1	C	246	PHE	3.9
1	B	161	LEU	3.8
1	C	540	GLU	3.6
1	A	161	LEU	3.6
1	A	246	PHE	3.4

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	159	GLU	3.4
1	B	245	LYS	3.4
1	A	538	TYR	3.3
1	A	158	ARG	3.1
1	C	160	ASN	3.1
1	B	422	THR	3.1
1	A	342	VAL	3.0
1	C	536	LYS	3.0
1	A	344	ARG	3.0
1	C	155	GLU	2.9
1	C	519	LEU	2.8
1	A	245	LYS	2.8
1	A	166	PHE	2.8
1	C	199	GLU	2.8
1	C	159	GLU	2.8
1	B	160	ASN	2.7
1	C	158	ARG	2.7
1	C	463	ARG	2.7
1	C	163	ASP	2.6
1	C	162	ASN	2.4
1	B	163	ASP	2.4
1	C	245	LYS	2.3
1	B	173	ARG	2.3
1	B	522	PHE	2.3
1	A	247	ALA	2.3
1	B	459	ALA	2.3
1	B	535	LYS	2.3
1	B	247	ALA	2.3
1	A	519	LEU	2.2
1	B	351	SER	2.2
1	B	166	PHE	2.2
1	C	164	PRO	2.2
1	A	2	ASN	2.2
1	A	160	ASN	2.2
1	B	190	TYR	2.2
1	B	357	SER	2.1
1	A	536	LYS	2.1
1	A	531	PHE	2.1
1	A	15	VAL	2.1
1	C	146	VAL	2.0
1	A	4	VAL	2.0
1	B	157	PHE	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

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
2	MN	C	1542	1/1	0.19	0.78	31,31,31,31	0
2	MN	B	1541	1/1	0.20	0.64	35,35,35,35	0
2	MN	A	1551	1/1	0.16	0.03	56,56,56,56	0
3	UPG	C	1544	36/36	0.15	-0.30	26,31,35,38	0
3	UPG	A	1544	36/36	0.16	-0.30	30,33,37,38	0
2	MN	A	1543	1/1	0.16	-0.31	35,35,35,35	0
3	UPG	B	1544	36/36	0.16	-0.33	31,38,42,43	0
2	MN	A	1552	1/1	0.14	-0.89	63,63,63,63	0

## 6.5 Other polymers

There are no such residues in this entry.