Full wwPDB X-ray Structure Validation Report ⓘ

Feb 22, 2024 – 06:16 pm GMT

|		
PDB ID	:	6QCV
Title	:	Crystal structure of influenza B polymerase initiation state with capped 14-mer RNA primer and CTP
Authors	:	Cusack, S.; Drncova, P.
Deposited on	:	2018-12-31
Resolution	:	3.24 Å(reported)

1   Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.24 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	Percentile Ranks	Value
Rfree		0.230
Clashscore		2
Ramachandran outliers		0.7%
Sidechain outliers		3.3%
RSRZ outliers		6.5%
RNA backbone		0.39

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
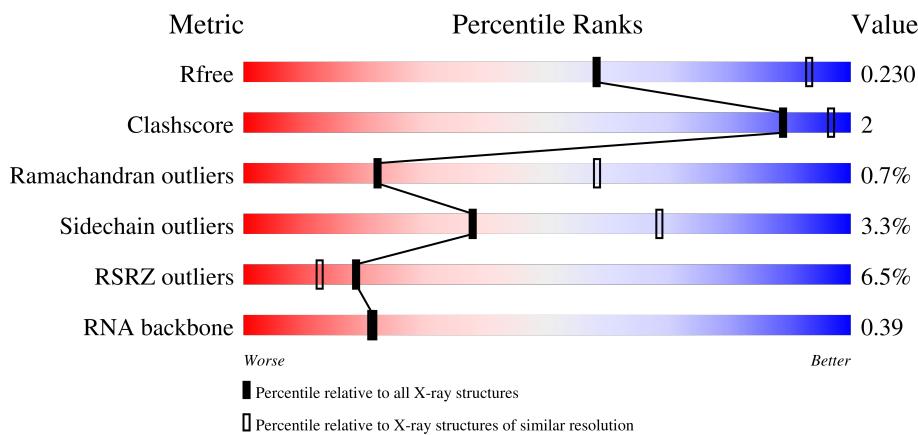
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1619 (3.28-3.20)
Clashscore	141614	1755 (3.28-3.20)
Ramachandran outliers	138981	1728 (3.28-3.20)
Sidechain outliers	138945	1727 (3.28-3.20)
RSRZ outliers	127900	1567 (3.28-3.20)
RNA backbone	3102	1034 (3.58-2.90)

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	A	751	5% / 88% / 8% / 5%
2	B	772	9% / 89% / 7% / · ·
3	C	798	4% / 83% / 8% / · / 8%
4	M	15	47% / 60% / 27% / 13%
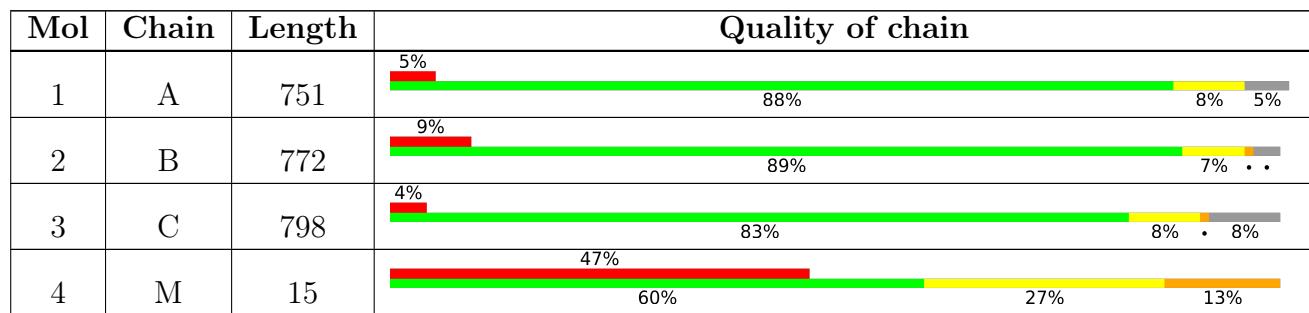
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	R	21	5% 67% 33%
6	V	14	36% 64%

2 Entry composition ⓘ

There are 9 unique types of molecules in this entry. The entry contains 18569 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Polymerase acidic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	717	5755	3657	963	1095	40	0	0	0

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	GLY	-	expression tag	UNP Q5V8Z9
A	-12	SER	-	expression tag	UNP Q5V8Z9
A	-11	HIS	-	expression tag	UNP Q5V8Z9
A	-10	HIS	-	expression tag	UNP Q5V8Z9
A	-9	HIS	-	expression tag	UNP Q5V8Z9
A	-8	HIS	-	expression tag	UNP Q5V8Z9
A	-7	HIS	-	expression tag	UNP Q5V8Z9
A	-6	HIS	-	expression tag	UNP Q5V8Z9
A	-5	HIS	-	expression tag	UNP Q5V8Z9
A	-4	HIS	-	expression tag	UNP Q5V8Z9
A	-3	GLY	-	expression tag	UNP Q5V8Z9
A	-2	SER	-	expression tag	UNP Q5V8Z9
A	-1	GLY	-	expression tag	UNP Q5V8Z9
A	0	SER	-	expression tag	UNP Q5V8Z9
A	727	GLY	-	expression tag	UNP Q5V8Z9
A	728	SER	-	expression tag	UNP Q5V8Z9
A	729	GLY	-	expression tag	UNP Q5V8Z9
A	730	SER	-	expression tag	UNP Q5V8Z9
A	731	GLY	-	expression tag	UNP Q5V8Z9
A	732	GLU	-	expression tag	UNP Q5V8Z9
A	733	ASN	-	expression tag	UNP Q5V8Z9
A	734	LEU	-	expression tag	UNP Q5V8Z9
A	735	TYR	-	expression tag	UNP Q5V8Z9
A	736	PHE	-	expression tag	UNP Q5V8Z9
A	737	GLN	-	expression tag	UNP Q5V8Z9

- Molecule 2 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	746	5852	3692	1016	1092	52	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	GLY	-	expression tag	UNP Q5V8Y6
B	-7	SER	-	expression tag	UNP Q5V8Y6
B	-6	GLY	-	expression tag	UNP Q5V8Y6
B	-5	SER	-	expression tag	UNP Q5V8Y6
B	-4	GLY	-	expression tag	UNP Q5V8Y6
B	-3	SER	-	expression tag	UNP Q5V8Y6
B	-2	GLY	-	expression tag	UNP Q5V8Y6
B	-1	SER	-	expression tag	UNP Q5V8Y6
B	0	GLY	-	expression tag	UNP Q5V8Y6
B	753	GLY	-	expression tag	UNP Q5V8Y6
B	754	SER	-	expression tag	UNP Q5V8Y6
B	755	GLY	-	expression tag	UNP Q5V8Y6
B	756	SER	-	expression tag	UNP Q5V8Y6
B	757	GLY	-	expression tag	UNP Q5V8Y6
B	758	GLU	-	expression tag	UNP Q5V8Y6
B	759	ASN	-	expression tag	UNP Q5V8Y6
B	760	LEU	-	expression tag	UNP Q5V8Y6
B	761	TYR	-	expression tag	UNP Q5V8Y6
B	762	PHE	-	expression tag	UNP Q5V8Y6
B	763	GLN	-	expression tag	UNP Q5V8Y6

- Molecule 3 is a protein called Polymerase basic protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	733	5863	3726	1026	1071	40	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-8	GLY	-	expression tag	UNP Q5V8X3
C	-7	SER	-	expression tag	UNP Q5V8X3
C	-6	GLY	-	expression tag	UNP Q5V8X3
C	-5	SER	-	expression tag	UNP Q5V8X3
C	-4	GLY	-	expression tag	UNP Q5V8X3
C	-3	SER	-	expression tag	UNP Q5V8X3
C	-2	GLY	-	expression tag	UNP Q5V8X3

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	SER	-	expression tag	UNP Q5V8X3
C	0	GLY	-	expression tag	UNP Q5V8X3
C	771	GLY	-	expression tag	UNP Q5V8X3
C	772	TRP	-	expression tag	UNP Q5V8X3
C	773	SER	-	expression tag	UNP Q5V8X3
C	774	HIS	-	expression tag	UNP Q5V8X3
C	775	PRO	-	expression tag	UNP Q5V8X3
C	776	GLN	-	expression tag	UNP Q5V8X3
C	777	PHE	-	expression tag	UNP Q5V8X3
C	778	GLU	-	expression tag	UNP Q5V8X3
C	779	LYS	-	expression tag	UNP Q5V8X3
C	780	GLY	-	expression tag	UNP Q5V8X3
C	781	SER	-	expression tag	UNP Q5V8X3
C	782	GLY	-	expression tag	UNP Q5V8X3
C	783	SER	-	expression tag	UNP Q5V8X3
C	784	GLU	-	expression tag	UNP Q5V8X3
C	785	ASN	-	expression tag	UNP Q5V8X3
C	786	LEU	-	expression tag	UNP Q5V8X3
C	787	TYR	-	expression tag	UNP Q5V8X3
C	788	PHE	-	expression tag	UNP Q5V8X3
C	789	GLN	-	expression tag	UNP Q5V8X3

- Molecule 4 is a RNA chain called RNA (5'-D(*(GDM))-R(P*GP*AP*AP*UP*GP*CP*UP *AP*UP*AP*AP*UP*AP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	M	15	330	146	61	107	16	0	0	0

- Molecule 5 is a RNA chain called RNA (5'-R(*UP*AP*UP*AP*CP*CP*UP*CP*UP*GP* CP*UP*UP*CP*UP*GP*CP*UP*AP*UP*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
5	R	21	427	193	63	151	20	0	0	0

- Molecule 6 is a RNA chain called RNA (5'-R(P*AP*GP*UP*AP*GP*UP*AP*AP*CP*AP *AP*GP*AP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
6	V	14	307	137	62	94	14	0	0	0

- Molecule 7 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).
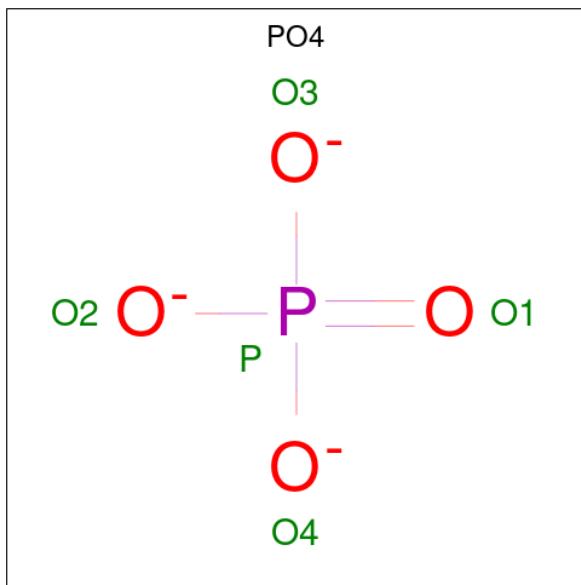
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 5	O 4	P 1	0	0

- Molecule 8 is CYTIDINE-5'-TRIPHOSPHATE (three-letter code: CTP) (formula: $C_9H_{16}N_3O_{14}P_3$) (labeled as "Ligand of Interest" by depositor).
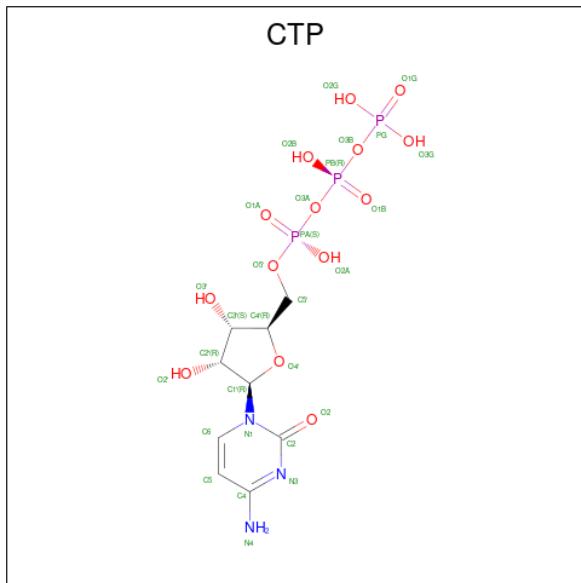
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	B	1	Total 29	C 9	N 3	O 14	P 3	0	0

• Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
9	B	1	1	1	0	0

3 Residue-property plots ⓘ

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: Polymerase acidic protein

Chain A:
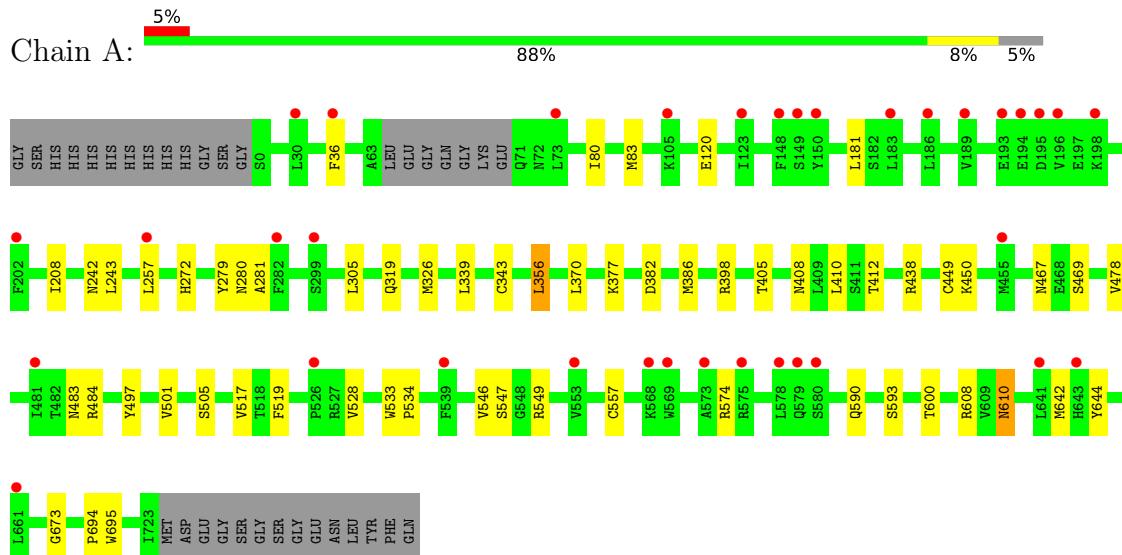
• Molecule 2: RNA-directed RNA polymerase catalytic subunit

Chain B:
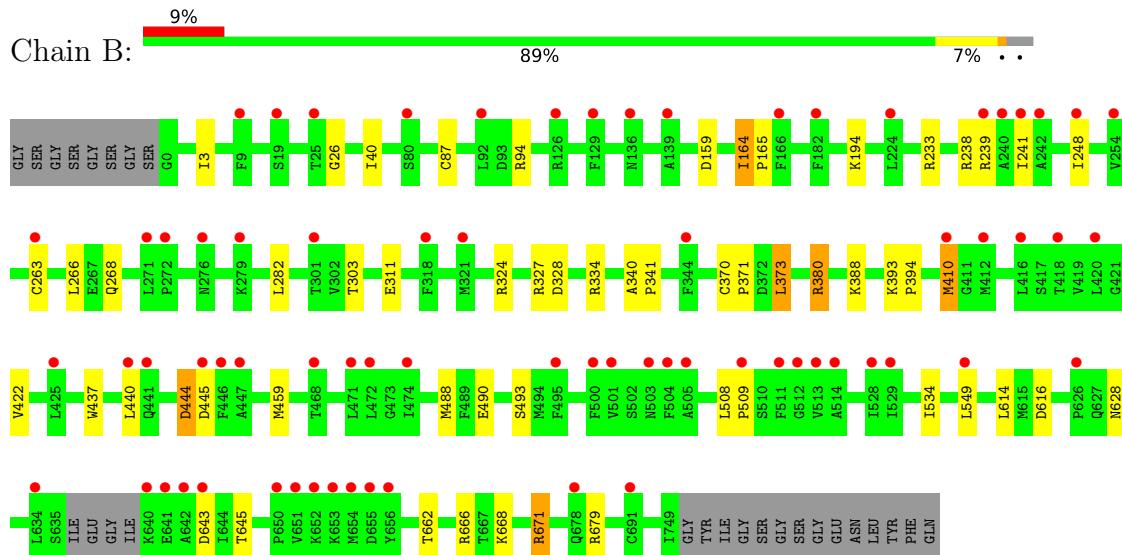
• Molecule 3: Polymerase basic protein 2

Chain C:

4%

83%

8% · 8%

GLY SER GLY SER GLY SER G0 L12 R34 K35 S46 L47 K50 M53 C54 P58 L59 A60 L61 T62 K80 I86 K89 C93 S94 I95 F120 R123 T136 R144 L148 L149 N163 L169 E173 R178 Y207

R211 E212 L213 L220 P221 L234 Q238 N241 S258 Q259 S260 M261 A264 C265 R266 R271 E280 D292 C299 D307 V308 N342 L376 K377 K378 A391 F422 L423 N424 R425 A426 G427 Q428 L429 L430 S431 Y434 Q437 S460 A471

K477 V485 ILE ASP ASP PHE SER SER THR GLU T494 L503 I506 V512 I513 A516 S520 Q527 I530 W538 L546 K569 E570 D571 M572 F573 Q574 W575 D576 A577 F578 F581 I584 Q603 M610 L619 P620 P625 P626 E647 F659

L668 T669 I670 G691 N692 G697 L717 I728 Y731 K734 V738 V739 K740 ARG LYS ARG TYR SER ALA LEU SER ASN ILE SER GLN GLY ILE LYS ARG ARG MET THR VAL GLU SER MET GLY TRP ALA LEU SER GLY TRP SER HIS PRO GLN PHE GLU LYS

GLY SER GLY SER GLU ASN LEU TYR PHE GLN

• Molecule 4: RNA (5'-D(*(GDM))-R(P*GP*AP*AP*UP*GP*CP*UP*AP*UP*AP*AP*UP*AP*G)-3')

Chain M:

47%

60%

27%

13%

M7G0 G1 A2 G5 C6 U7 A8 U9 A10 A11 G14
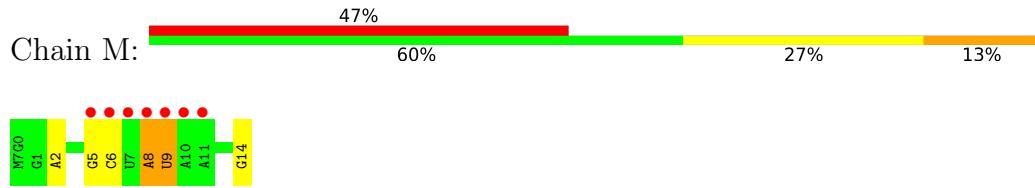
• Molecule 5: RNA (5'-R(*UP*AP*UP*AP*CP*CP*UP*CP*UP*GP*CP*UP*UP*CP*UP*GP*CP*UP*AP*UP*U)-3')

Chain R:

5%

67%

33%

U1 C5 C6 C11 U12 U13 C14 U15 U21

• Molecule 6: RNA (5'-R(P*AP*GP*UP*AP*GP*UP*AP*AP*CP*AP*AP*GP*AP*G)-3')

Chain V:

36%

64%

A1 G5 U6 A7 A8 C9 A10 A11 G12 A13 G14

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	200.19Å 200.19Å 256.21Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 3.24 173.37 – 3.24	Depositor EDS
% Data completeness (in resolution range)	67.7 (50.00-3.24) 67.7 (173.37-3.24)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.69 (at 3.26Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.209 , 0.234 0.207 , 0.230	Depositor DCC
R_{free} test set	3069 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å^2)	97.9	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 74.9	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.035 for -h,-k,l	Xtriage
F_o,F_c correlation	0.83	EDS
Total number of atoms	18569	wwPDB-VP
Average B, all atoms (Å^2)	126.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.52% 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: PO4, CTP, M7G, MG

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	0.67	0/5871	0.70	0/7918
2	B	0.66	0/5967	0.70	0/8043
3	C	0.67	0/5963	0.71	0/8014
4	M	0.20	0/337	0.65	0/523
5	R	0.27	0/473	0.69	0/732
6	V	0.60	1/345 (0.3%)	0.64	0/535
All	All	0.65	1/18956 (0.0%)	0.70	0/25765

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	V	1	A	OP3-P	-9.91	1.49	1.61

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	A	5755	0	5730	25	0
2	B	5852	0	5872	24	0
3	C	5863	0	6031	26	0
4	M	330	0	167	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	R	427	0	219	2	0
6	V	307	0	153	1	0
7	A	5	0	0	0	0
8	B	29	0	12	0	0
9	B	1	0	0	0	0
All	All	18569	0	18184	71	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 2.

All (71) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:467:ASN:HA	3:C:54:CYS:SG	2.39	0.63
2:B:340:ALA:HB3	2:B:341:PRO:HD3	1.84	0.59
1:A:356:LEU:HD13	1:A:478:VAL:HG23	1.85	0.58
1:A:574:ARG:HB3	2:B:549:LEU:HD22	1.86	0.57
1:A:386:MET:O	2:B:380:ARG:NH1	2.40	0.54
1:A:593:SER:HB3	3:C:241:ASN:HB3	1.93	0.50
2:B:370:CYS:HA	2:B:373:LEU:HD13	1.94	0.50
1:A:546:VAL:HG12	1:A:549:ARG:HB2	1.93	0.50
3:C:80:LYS:HA	3:C:93:CYS:HA	1.94	0.49
3:C:728:ILE:HD13	3:C:738:VAL:HB	1.95	0.49
2:B:26:GLY:O	2:B:233:ARG:NH2	2.45	0.49
3:C:731:TYR:CE2	3:C:734:LYS:HG3	2.48	0.49
5:R:21:U:O2	5:R:21:U:O4'	2.30	0.49
1:A:642:MET:HG3	1:A:695:TRP:CZ3	2.48	0.48
3:C:264:ALA:HB2	3:C:530:ILE:HD11	1.96	0.48
2:B:241:ILE:HG22	2:B:410:MET:HG2	1.95	0.48
3:C:717:LEU:HD21	3:C:728:ILE:HD11	1.96	0.48
2:B:282:LEU:HD22	2:B:440:LEU:HD13	1.96	0.47
1:A:80:ILE:HA	1:A:83:MET:HE2	1.96	0.47
1:A:610:ASN:HD22	1:A:610:ASN:N	2.13	0.47
3:C:581:PHE:O	3:C:584:ILE:HG13	2.15	0.47
3:C:659:PHE:CZ	3:C:668:LEU:HD11	2.50	0.46
3:C:603:GLN:OE1	3:C:697:GLY:N	2.47	0.46
1:A:36:PHE:CZ	1:A:181:LEU:HD11	2.51	0.46
3:C:603:GLN:NE2	3:C:692:ASN:O	2.49	0.46
1:A:483:ASN:HB3	1:A:497:TYR:HE2	1.80	0.45
4:M:8:A:O2'	4:M:9:U:OP1	2.27	0.45
3:C:258:SER:O	3:C:261:MET:HB3	2.17	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:410:LEU:HA	1:A:450:LYS:HB3	1.98	0.45
2:B:327:ARG:HA	2:B:334:ARG:NH1	2.32	0.45
3:C:578:PHE:CD1	3:C:620:PRO:HA	2.52	0.45
1:A:280:ASN:HD21	1:A:694:PRO:HD3	1.82	0.44
1:A:644:TYR:HA	2:B:26:GLY:HA2	2.00	0.44
2:B:508:LEU:HB3	2:B:509:PRO:HD3	2.00	0.44
1:A:242:ASN:ND2	2:B:87:CYS:SG	2.91	0.44
3:C:58:PRO:HD2	3:C:94:SER:HA	1.99	0.44
2:B:238:ARG:HA	2:B:238:ARG:NE	2.32	0.43
3:C:506:ILE:HG22	3:C:512:VAL:HG12	2.00	0.43
1:A:408:ASN:HA	2:B:3:ILE:HD12	2.00	0.43
1:A:519:PHE:HD1	1:A:557:CYS:HG	1.67	0.43
3:C:89:LYS:HA	3:C:89:LYS:HE2	2.00	0.43
3:C:431:SER:HB3	3:C:434:TYR:HB2	2.01	0.43
1:A:501:VAL:CG2	1:A:517:VAL:HB	2.49	0.43
2:B:263:CYS:HA	2:B:266:LEU:HD12	2.01	0.42
1:A:438:ARG:HA	1:A:600:THR:HG21	2.01	0.42
3:C:538:TRP:CZ3	3:C:546:LEU:HD23	2.54	0.42
3:C:53:MET:SD	3:C:95:ILE:HG21	2.60	0.42
3:C:266:ARG:CZ	3:C:520:SER:HB3	2.49	0.42
1:A:533:TRP:N	1:A:534:PRO:CD	2.82	0.42
2:B:628:ASN:HD22	2:B:662:THR:HG22	1.85	0.41
1:A:343:CYS:SG	1:A:501:VAL:HG13	2.60	0.41
6:V:8:A:H2'	6:V:9:C:C6	2.55	0.41
3:C:163:ASN:HD22	3:C:178:ARG:HG3	1.85	0.41
2:B:303:THR:HG23	2:B:490:GLU:O	2.21	0.41
1:A:343:CYS:HA	1:A:478:VAL:HG21	2.03	0.41
2:B:164:ILE:N	2:B:165:PRO:HD2	2.36	0.41
5:R:5:C:H4'	5:R:6:C:OP2	2.21	0.41
2:B:266:LEU:HD13	2:B:422:VAL:HG11	2.03	0.41
2:B:444:ASP:OD1	4:M:14:G:O3'	2.37	0.41
1:A:272:HIS:HA	1:A:398:ARG:HD3	2.03	0.41
3:C:619:LEU:HB3	3:C:620:PRO:HD3	2.02	0.41
1:A:279:TYR:CE1	1:A:281:ALA:HB3	2.55	0.41
1:A:469:SER:O	1:A:505:SER:OG	2.35	0.41
2:B:40:ILE:HG21	2:B:388:LYS:HB3	2.03	0.41
2:B:370:CYS:N	2:B:371:PRO:CD	2.84	0.41
2:B:508:LEU:N	2:B:509:PRO:CD	2.84	0.41
3:C:625:PRO:HA	3:C:626:PRO:HD3	1.98	0.41
2:B:393:LYS:HB3	2:B:394:PRO:HD3	2.02	0.40
2:B:671:ARG:HB3	3:C:86:ILE:HD11	2.04	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:61:LEU:HD21	3:C:93:CYS:HB3	2.02	0.40
3:C:120:PHE:O	3:C:123:ARG:HB3	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	713/751 (95%)	653 (92%)	57 (8%)	3 (0%)	34	68
2	B	742/772 (96%)	690 (93%)	47 (6%)	5 (1%)	22	58
3	C	729/798 (91%)	671 (92%)	51 (7%)	7 (1%)	15	50
All	All	2184/2321 (94%)	2014 (92%)	155 (7%)	15 (1%)	22	58

All (15) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	547	SER
1	A	673	GLY
2	B	194	LYS
3	C	471	ALA
2	B	643	ASP
2	B	679	ARG
3	C	391	ALA
3	C	477	LYS
3	C	691	GLY
3	C	292	ASP
1	A	528	VAL
2	B	410	MET
3	C	460	SER
2	B	534	ILE
3	C	308	VAL

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	638/664 (96%)	619 (97%)	19 (3%)	41	70
2	B	640/657 (97%)	617 (96%)	23 (4%)	35	66
3	C	640/694 (92%)	618 (97%)	22 (3%)	37	68
All	All	1918/2015 (95%)	1854 (97%)	64 (3%)	38	68

All (64) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	120	GLU
1	A	208	ILE
1	A	243	LEU
1	A	257	LEU
1	A	305	LEU
1	A	319	GLN
1	A	326	MET
1	A	339	LEU
1	A	356	LEU
1	A	370	LEU
1	A	377	LYS
1	A	382	ASP
1	A	405	THR
1	A	412	THR
1	A	449	CYS
1	A	484	ARG
1	A	590	GLN
1	A	608	ARG
1	A	610	ASN
2	B	94	ARG
2	B	159	ASP
2	B	164	ILE
2	B	239	ARG
2	B	248	ILE
2	B	268	GLN
2	B	311	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	324	ARG
2	B	328	ASP
2	B	373	LEU
2	B	380	ARG
2	B	437	TRP
2	B	444	ASP
2	B	445	ASP
2	B	459	MET
2	B	488	MET
2	B	493	SER
2	B	614	LEU
2	B	616	ASP
2	B	645	THR
2	B	666	ARG
2	B	668	LYS
2	B	671	ARG
3	C	34	ARG
3	C	61	LEU
3	C	62	THR
3	C	89	LYS
3	C	136	THR
3	C	173	GLU
3	C	178	ARG
3	C	238	GLN
3	C	259	GLN
3	C	271	ARG
3	C	280	GLU
3	C	299	CYS
3	C	307	ASP
3	C	378	LYS
3	C	437	GLN
3	C	513	ILE
3	C	527	GLN
3	C	530	ILE
3	C	572	MET
3	C	576	ASP
3	C	610	MET
3	C	647	GLU

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

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	M	13/15 (86%)	5 (38%)	1 (7%)
5	R	19/21 (90%)	3 (15%)	1 (5%)
6	V	13/14 (92%)	6 (46%)	1 (7%)
All	All	45/50 (90%)	14 (31%)	3 (6%)

All (14) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	M	2	A
4	M	5	G
4	M	6	C
4	M	8	A
4	M	9	U
5	R	11	C
5	R	14	C
5	R	15	U
6	V	5	G
6	V	6	U
6	V	7	A
6	V	11	A
6	V	13	A
6	V	14	G

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	M	8	A
5	R	13	U
6	V	5	G

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	#$|Z| > 2$
7	PO4	A	801	-	4,4,4	0.64	0	6,6,6	0.45	0
8	CTP	B	801	9	26,30,30	0.50	0	39,47,47	0.54	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
8	CTP	B	801	9	-	4/22/38/38	0/2/2/2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	B	801	CTP	PB-O3B-PG-O2G
8	B	801	CTP	PB-O3B-PG-O3G
8	B	801	CTP	C5'-O5'-PA-O3A
8	B	801	CTP	C5'-O5'-PA-O1A

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier.

Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand CTP B 801

Bond lengths   Bond angles

Torsions   Rings
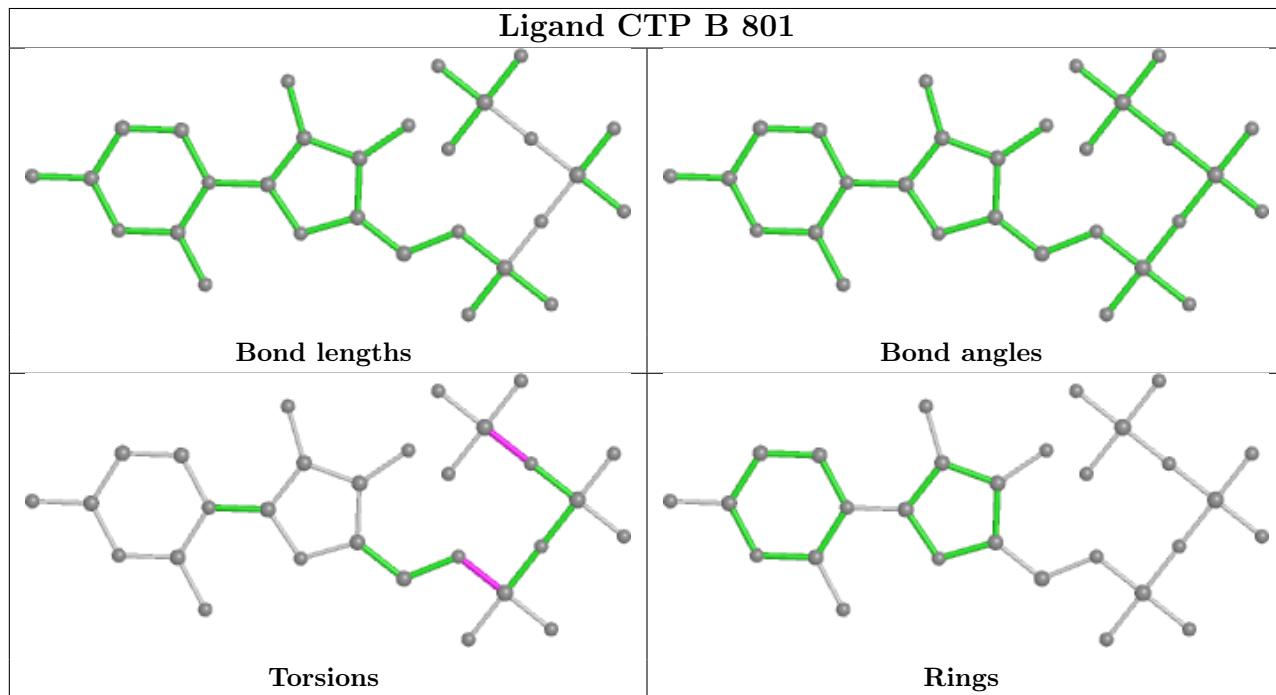
5.7 Other polymers ⓘ

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(Å²)	Q<0.9
1	A	717/751 (95%)	0.56	35 (4%)	29	19	83, 117, 212, 228	0
2	B	746/772 (96%)	0.92	71 (9%)	8	6	85, 113, 146, 180	0
3	C	733/798 (91%)	0.54	33 (4%)	33	23	89, 123, 192, 238	0
4	M	14/15 (93%)	2.24	7 (50%)	0	0	124, 159, 197, 200	0
5	R	21/21 (100%)	1.01	1 (4%)	30	20	88, 102, 127, 139	0
6	V	14/14 (100%)	0.69	0	100	100	88, 94, 100, 101	0
All	All	2245/2371 (94%)	0.69	147 (6%)	18	12	83, 117, 199, 238	0

All (147) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	653	LYS	10.0
1	A	150	TYR	6.5
4	M	9	U	6.1
2	B	654	MET	5.3
4	M	6	C	4.0
3	C	426	ALA	3.9
4	M	8	A	3.8
2	B	652	LYS	3.8
2	B	271	LEU	3.8
3	C	425	ARG	3.7
2	B	634	LEU	3.5
1	A	195	ASP	3.5
3	C	207	TYR	3.5
1	A	194	GLU	3.3
2	B	224	LEU	3.1
3	C	569	LYS	3.1
4	M	5	G	3.1
1	A	186	LEU	3.1
2	B	440	LEU	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	C	570	GLU	3.1
3	C	430	LEU	3.1
3	C	221	PRO	3.0
1	A	202	PHE	3.0
2	B	511	PHE	3.0
2	B	642	ALA	3.0
2	B	441	GLN	3.0
3	C	213	LEU	3.0
1	A	189	VAL	3.0
4	M	10	A	3.0
1	A	148	PHE	2.9
2	B	468	THR	2.9
2	B	242	ALA	2.9
2	B	129	PHE	2.8
3	C	220	LEU	2.8
4	M	7	U	2.8
1	A	481	ILE	2.8
1	A	193	GLU	2.8
2	B	529	ILE	2.8
2	B	656	TYR	2.7
3	C	423	LEU	2.7
2	B	650	PRO	2.7
1	A	569	TRP	2.6
3	C	144	ARG	2.6
2	B	418	THR	2.6
2	B	279	LYS	2.6
1	A	30	LEU	2.6
1	A	123	ILE	2.6
1	A	573	ALA	2.6
2	B	512	GLY	2.6
2	B	241	ILE	2.6
2	B	651	VAL	2.6
1	A	455	MET	2.6
2	B	509	PRO	2.6
2	B	420	LEU	2.6
2	B	495	PHE	2.5
2	B	640	LYS	2.5
2	B	691	CYS	2.5
1	A	526	PRO	2.5
2	B	80	SER	2.5
2	B	643	ASP	2.5
2	B	321	MET	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	503	ASN	2.4
2	B	501	VAL	2.4
4	M	11	A	2.4
2	B	25	THR	2.4
1	A	105	LYS	2.4
2	B	513	VAL	2.4
2	B	240	ALA	2.4
3	C	427	GLY	2.3
3	C	516	ALA	2.3
2	B	514	ALA	2.3
2	B	500	PHE	2.3
1	A	36	PHE	2.3
2	B	272	PRO	2.3
1	A	73	LEU	2.3
2	B	139	ALA	2.3
2	B	276	ASN	2.3
2	B	126	ARG	2.3
2	B	447	ALA	2.3
1	A	196	VAL	2.3
1	A	661	LEU	2.2
5	R	21	U	2.2
1	A	282	PHE	2.2
1	A	149	SER	2.2
1	A	641	LEU	2.2
3	C	47	LEU	2.2
1	A	257	LEU	2.2
1	A	580	SER	2.2
3	C	169	LEU	2.2
1	A	198	LYS	2.2
1	A	299	SER	2.2
2	B	19	SER	2.2
1	A	643	HIS	2.2
3	C	12	LEU	2.2
2	B	136	ASN	2.2
3	C	428	GLN	2.2
2	B	678	GLN	2.2
2	B	248	ILE	2.2
2	B	528	ILE	2.2
2	B	425	LEU	2.1
3	C	503	LEU	2.1
2	B	505	ALA	2.1
2	B	239	ARG	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	C	148	LEU	2.1
2	B	472	LEU	2.1
3	C	149	LEU	2.1
3	C	234	LEU	2.1
1	A	568	LYS	2.1
3	C	35	LYS	2.1
3	C	211	ARG	2.1
3	C	50	LYS	2.1
3	C	342	ASN	2.1
2	B	641	GLU	2.1
3	C	572	MET	2.1
2	B	412	MET	2.1
3	C	59	LEU	2.1
2	B	445	ASP	2.1
1	A	578	LEU	2.1
2	B	410	MET	2.1
2	B	416	LEU	2.1
1	A	539	PHE	2.1
2	B	166	PHE	2.1
2	B	318	PHE	2.1
2	B	626	PRO	2.1
3	C	670	ILE	2.1
1	A	183	LEU	2.1
3	C	376	LEU	2.1
3	C	95	ILE	2.1
2	B	92	LEU	2.1
2	B	344	PHE	2.1
2	B	504	PHE	2.1
2	B	301	THR	2.1
1	A	579	GLN	2.1
3	C	574	GLN	2.1
2	B	471	LEU	2.1
3	C	46	SER	2.1
2	B	446	PHE	2.0
2	B	549	LEU	2.0
2	B	9	PHE	2.0
2	B	254	VAL	2.0
2	B	263	CYS	2.0
1	A	575	ARG	2.0
2	B	474	ILE	2.0
3	C	422	PHE	2.0
1	A	553	VAL	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	655	ASP	2.0
2	B	182	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 monosaccharides 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. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	PO4	A	801	5/5	0.90	0.21	139,140,141,141	0
8	CTP	B	801	29/29	0.91	0.26	121,135,146,152	0
9	MG	B	802	1/1	0.99	0.10	129,129,129,129	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around CTP B 801:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers ⓘ

There are no such residues in this entry.