W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Sep 6, 2023 – 11:44 PM EDT

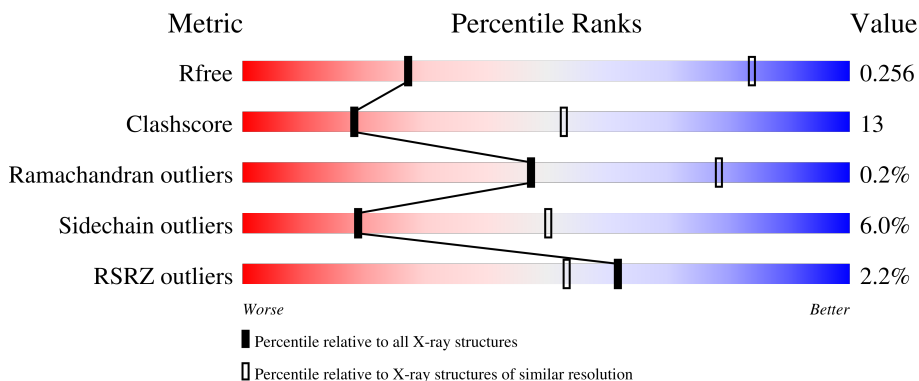|		
PDB ID	:	4G7Z
Title	:	Crystal structure of Thermus thermophilus transcription initiation complex containing 5-BrU at template-strand position +1
Authors	:	Zhang, Y.; Ebright, R.H.
Deposited on	:	2012-07-20
Resolution	:	3.81 Å(reported)

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 3.81 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		0.256
Clashscore		13
Ramachandran outliers		0.2%
Sidechain outliers		6.0%
RSRZ outliers		2.2%

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	1231 (4.04-3.60)
Clashscore	141614	1031 (4.02-3.62)
Ramachandran outliers	138981	1261 (4.04-3.60)
Sidechain outliers	138945	1255 (4.04-3.60)
RSRZ outliers	127900	1139 (4.04-3.60)

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	315	42% 27% • 28%
1	B	315	43% 26% • 30%
1	K	315	% 55% 16% • 28%
1	L	315	54% 17% • 29%
2	C	1119	% 60% 36% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	M	1119	5% / 70% / 27%
3	D	1524	% / 58% / 35%
3	N	1524	2% / 72% / 24%
4	E	99	% / 60% / 31% / 5%
4	O	99	4% / 74% / 19% / 5%
5	F	443	5% / 56% / 21% / 22%
5	P	443	5% / 58% / 17% / 22%
6	H	27	19% / 48% / 22% / 11%
6	R	27	7% / 67% / 15% / 11%
7	G	21	5% / 29% / 38% / 10% / 24%
7	Q	21	29% / 33% / 14% / 24%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	BRU	Q	15	-	-	X	-

2 Entry composition ⓘ

There are 9 unique types of molecules in this entry. The entry contains 56872 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 DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	226	Total 1782	C 1138	N 310	O 332	S 2	0	0	0
1	B	222	Total 1750	C 1118	N 304	O 326	S 2	0	0	0
1	K	226	Total 1782	C 1138	N 310	O 332	S 2	0	0	0
1	L	225	Total 1773	C 1133	N 308	O 330	S 2	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1111	Total 8770	C 5548	N 1564	O 1634	S 24	0	0	0
2	M	1111	Total 8770	C 5548	N 1564	O 1634	S 24	0	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1482	Total 11704	C 7421	N 2059	O 2189	S 35	0	0	0
3	N	1489	Total 11746	C 7446	N 2066	O 2198	S 36	0	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	94	Total 761	C 486	N 132	O 139	S 4	0	0	0
4	O	94	Total 761	C 486	N 132	O 139	S 4	0	0	0

- Molecule 5 is a protein called RNA polymerase sigma factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	346	Total 2807	C 1770	N 509	O 524	S 4	0	0	0
5	P	347	Total 2814	C 1774	N 510	O 526	S 4	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-19	MET	-	expression tag	UNP Q5SKW1
F	-18	GLY	-	expression tag	UNP Q5SKW1
F	-17	SER	-	expression tag	UNP Q5SKW1
F	-16	SER	-	expression tag	UNP Q5SKW1
F	-15	HIS	-	expression tag	UNP Q5SKW1
F	-14	HIS	-	expression tag	UNP Q5SKW1
F	-13	HIS	-	expression tag	UNP Q5SKW1
F	-12	HIS	-	expression tag	UNP Q5SKW1
F	-11	HIS	-	expression tag	UNP Q5SKW1
F	-10	HIS	-	expression tag	UNP Q5SKW1
F	-9	SER	-	expression tag	UNP Q5SKW1
F	-8	SER	-	expression tag	UNP Q5SKW1
F	-7	GLY	-	expression tag	UNP Q5SKW1
F	-6	LEU	-	expression tag	UNP Q5SKW1
F	-5	VAL	-	expression tag	UNP Q5SKW1
F	-4	PRO	-	expression tag	UNP Q5SKW1
F	-3	ARG	-	expression tag	UNP Q5SKW1
F	-2	GLY	-	expression tag	UNP Q5SKW1
F	-1	SER	-	expression tag	UNP Q5SKW1
F	0	HIS	-	expression tag	UNP Q5SKW1
P	-19	MET	-	expression tag	UNP Q5SKW1
P	-18	GLY	-	expression tag	UNP Q5SKW1
P	-17	SER	-	expression tag	UNP Q5SKW1
P	-16	SER	-	expression tag	UNP Q5SKW1
P	-15	HIS	-	expression tag	UNP Q5SKW1
P	-14	HIS	-	expression tag	UNP Q5SKW1
P	-13	HIS	-	expression tag	UNP Q5SKW1
P	-12	HIS	-	expression tag	UNP Q5SKW1
P	-11	HIS	-	expression tag	UNP Q5SKW1
P	-10	HIS	-	expression tag	UNP Q5SKW1
P	-9	SER	-	expression tag	UNP Q5SKW1
P	-8	SER	-	expression tag	UNP Q5SKW1
P	-7	GLY	-	expression tag	UNP Q5SKW1
P	-6	LEU	-	expression tag	UNP Q5SKW1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
P	-5	VAL	-	expression tag	UNP Q5SKW1
P	-4	PRO	-	expression tag	UNP Q5SKW1
P	-3	ARG	-	expression tag	UNP Q5SKW1
P	-2	GLY	-	expression tag	UNP Q5SKW1
P	-1	SER	-	expression tag	UNP Q5SKW1
P	0	HIS	-	expression tag	UNP Q5SKW1

- Molecule 6 is a DNA chain called 5'-D(*TP*AP*TP*AP*AP*TP*GP*GP*GP*AP*GP*C P*TP*GP*TP*CP*AP*CP*GP*GP*AP*TP*GP*CP*AP*GP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	H	24	Total 495	C 236	N 94	O 142	P 23	0	0	0
6	R	24	Total 495	C 236	N 94	O 142	P 23	0	0	0

- Molecule 7 is a DNA chain called 5'-D(*CP*CP*TP*GP*CP*AP*TP*CP*CP*GP*TP*G P*AP*GP*(BRU)P*CP*GP*AP*GP*GP*G)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
7	G	16	Total 328	Br 1	C 155	N 63	O 94	P 15	0	0	0
7	Q	16	Total 328	Br 1	C 155	N 63	O 94	P 15	0	0	0

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	D	2	Total 2	Zn 2	0	0
8	N	2	Total 2	Zn 2	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	1	Total 1	Mg 1	0	0
9	N	1	Total 1	Mg 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.
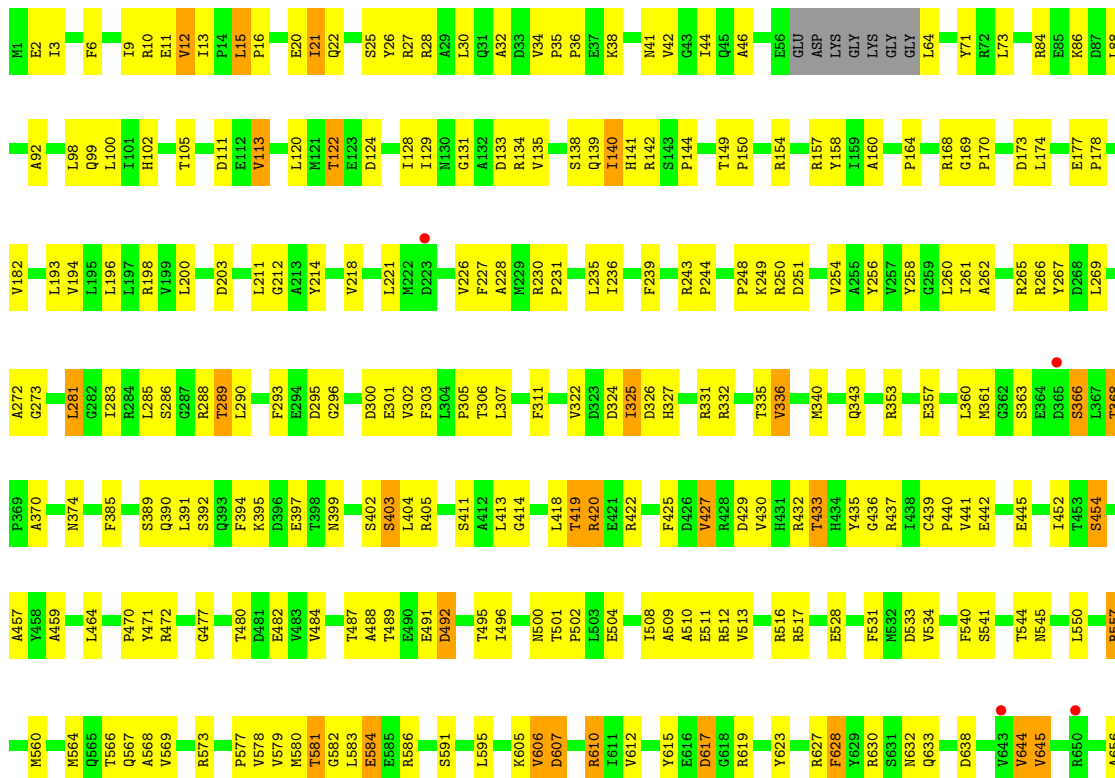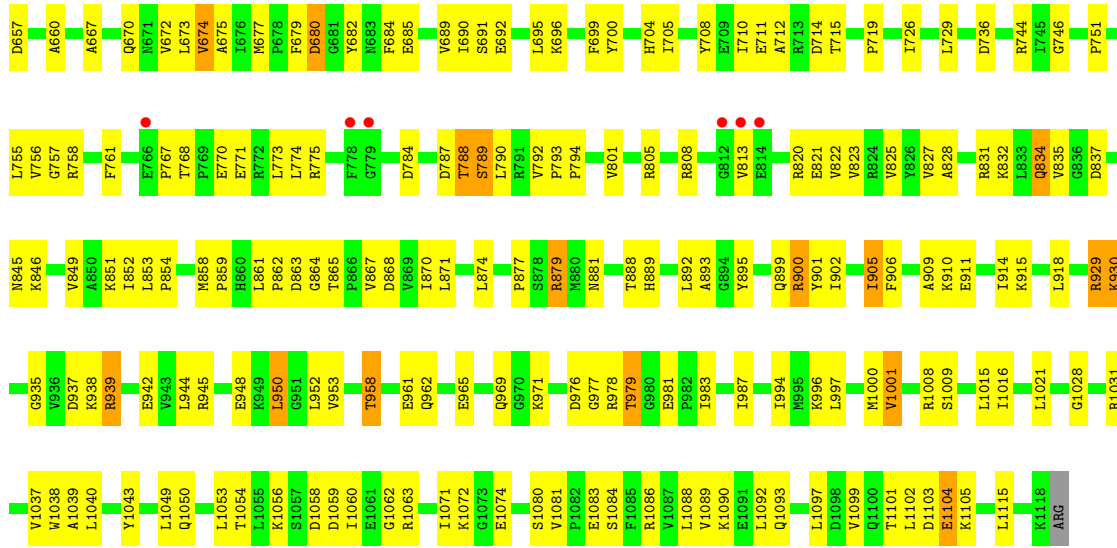
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:

• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain B:

• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain K:

P152 D168 F171 A178 F179 D183 T184 R185 L186 G187 Q188 R189 T190 D191 L192 D193 T196 L197 T201 D202 G203 T206 P207 L208 E209 V215 E216 R219 Q229 ALA ALA ALA VAL ALA ALA PRO GLU GLU ALA LYS GLU PRO GLU ALA PRO PRO GLU GLN GLU GLU

GLU LEU ASP PRO LEU GLU GLU LEU LEU GLY LEU SER THR ARG VAL LEU LEU HIS SER LEU LYS GLU GLY GLY ILE GLU SER VAL ARG ALA LEU LEU ALA LEU ASN LEU ASP LEU LYS ASN ILE PRO GLY ILE GLY ARG SER LEU GLU GLU ILE LYS GLU ALA LEU GLU LYS LYS GLY

PHE THR LEU LYS GLU

• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain L:　　　　54%　　　　17%　　　•　　　　29%

MET LEU ASP S4 K5 L6 R18 E19 Y20 G21 E22 F23 V24 F32 G33 V34 T35 P39 L40 L44 I48 V56 Y57 I58 V61 L62 H63 E64 F65 V71 D74 V75 I78 I79 L82 L90 N91 L94 L115 P116 V117 A118 D119 N124

L132 V142 V151 H156 G157 I158 N163 P173 R176 R185 T190 D191 K194 L197 R198 T201 V205 T206 P207 L208 E209 A210 L211 N212 Q213 I217 L222 P228 GLN ALA ALA ALA VAL ALA ALA PRO GLU GLU ALA LYS GLU GLU PRO GLU ALA PRO

PRO GLU GLN GLU GLU GLU LEU ASP LEU PRO LEU GLU GLU LEU GLY LEU SER THR ARG VAL LEU HIS SER LEU LYS GLU LYS GLU ILE GLU SER VAL ARG ALA LEU LEU ALA LEU ASN LEU LYS ASP LEU LYS ASN ILE PRO GLY ILE GLY GLU ARG SER LEU GLU GLU ILE LYS GLU ALA

LEU GLU LYS LYS GLY PHE THR LEU LYS GLU

• Molecule 2: DNA-directed RNA polymerase subunit beta

%

Chain C:　　　　60%　　　　36%　　　••　　••

M1 E2 I3 F6 I9 R10 E11 V12 I13 P14 L15 P16 E20 I21 Q22 S25 Y26 R27 R28 A29 L30 Q31 A32 D33 V34 P35 P36 E37 K38 N41 V42 G43 I44 Q45 A46 E56 GLU ASP LYS GLY LYS GLY L64 Y71 R72 L73 R84 E85 K86 D87 L88

A92 L98 Q99 L100 I101 H102 T105 D111 E112 V113 L120 M121 T122 E123 D124 I128 I129 N130 G131 A132 D133 R134 V135 S138 Q139 I140 H141 R142 S143 P144 T149 P150 R154 R157 Y158 I159 A160 P164 R168 G169 P170 D173 L174 E177 P178

V182 L193 V194 L195 L196 L197 R198 V199 L200 D203 L211 G212 Y214 V218 L221 M222 D223 V226 F227 A228 M229 R230 P231 L235 I236 F239 R243 P244 P248 K249 R250 D251 V254 A255 Y256 V257 Y258 G259 L260 I261 A262 R265 R266 Y267 D268 L269

A272 G273 L281 G282 I283 R284 L285 S286 G287 R288 T289 L290 F293 E294 D295 G296 D300 E301 V302 F303 L304 P305 T306 L307 F311 V322 D323 D324 I325 D326 H327 R331 R332 T335 V336 M340 Q343 R353 E357 L360 M361 G362 S363 E364 D365 S366 L367 T368

P369 A370 N374 F385 S389 Q390 L391 S392 Q393 F394 K395 D396 E397 T398 N399 S402 S403 L404 R405 S411 A412 L413 G414 L418 T419 R420 E421 R422 F425 D426 V427 R428 D429 V430 H431 R432 T433 H434 Y435 G436 R437 I438 C439 P440 V441 E442 E445 I452 T453 S454

A457 Y458 A459 L464 P470 Y471 R472 G477 T480 D481 E482 V483 V484 T487 A488 T489 E490 E491 D492 T495 I496 N500 T501 P502 L503 E504 I508 A509 A510 E511 R512 V513 R516 R517 E528 F531 M532 D533 V534 F540 S541 T544 N545 L550 R557

M560 M564 Q565 T566 Q567 A568 V569 R573 P577 V578 V579 M580 T581 G582 L583 E584 E585 R586 S591 L595 K605 V606 D607 R610 I611 V612 Y615 E616 D617 G618 R619 Y623 R627 F628 Y629 R630 S631 N632 Q633 D638 V643 V644 V645 R650 A656

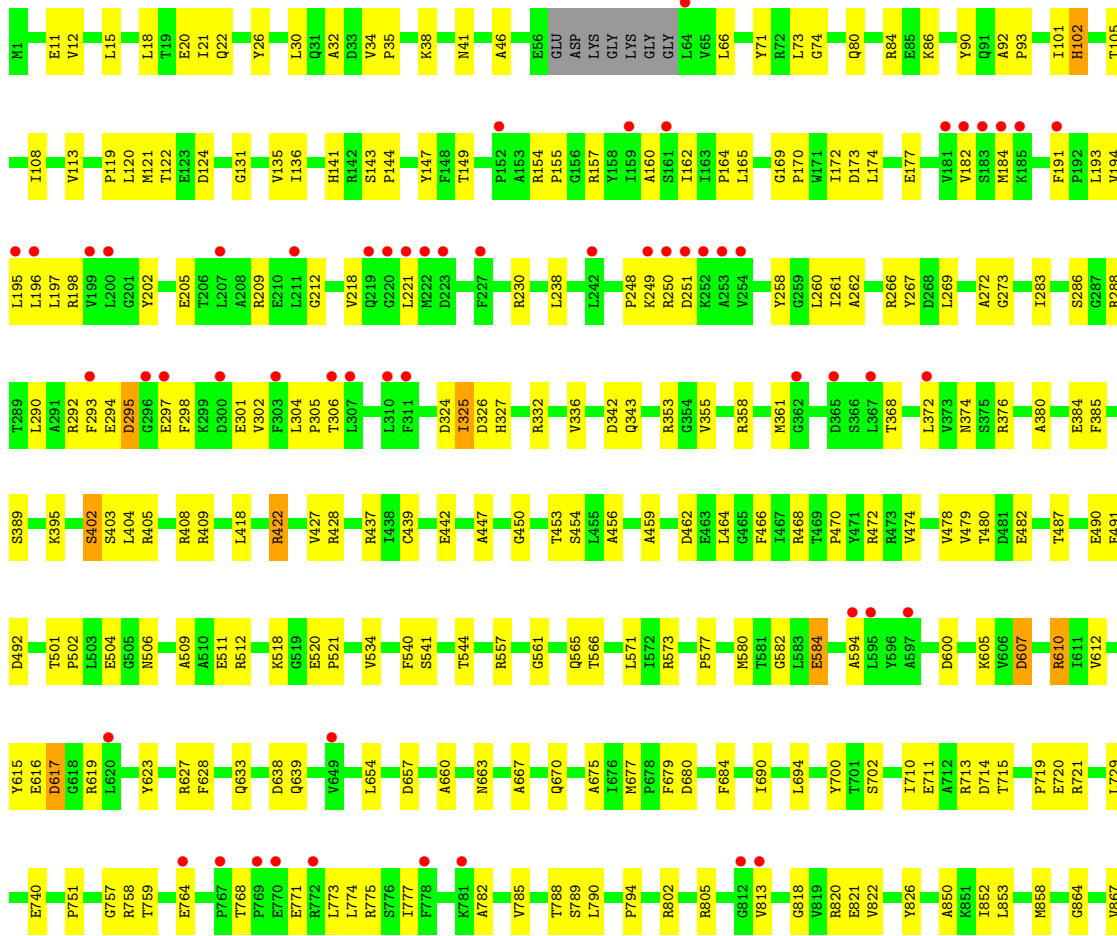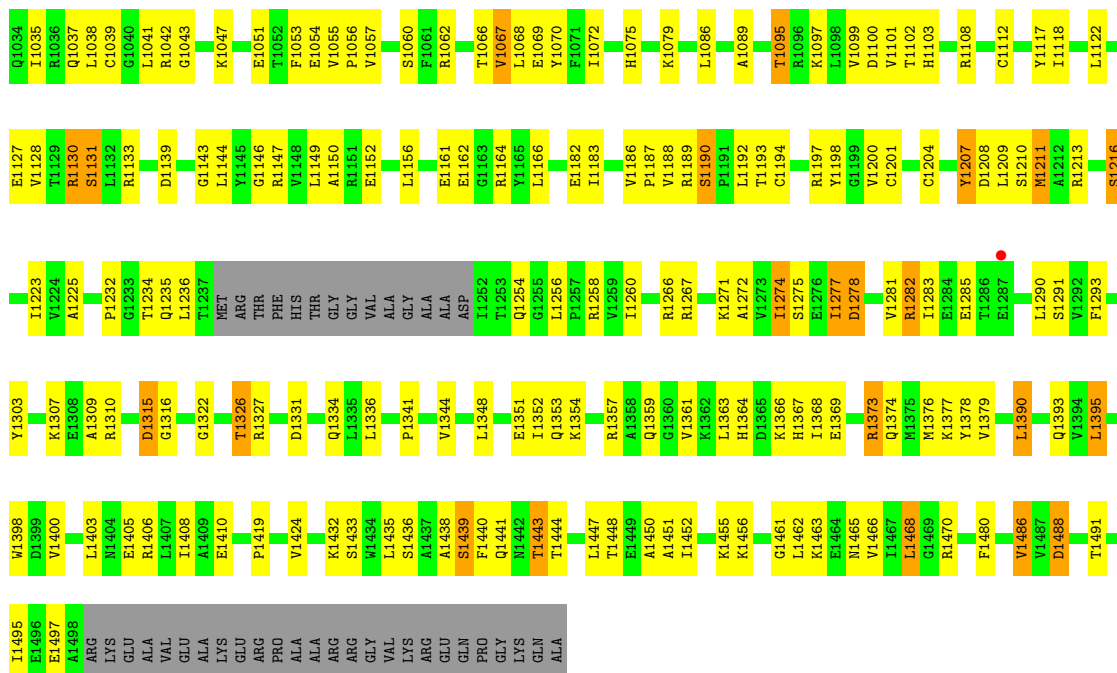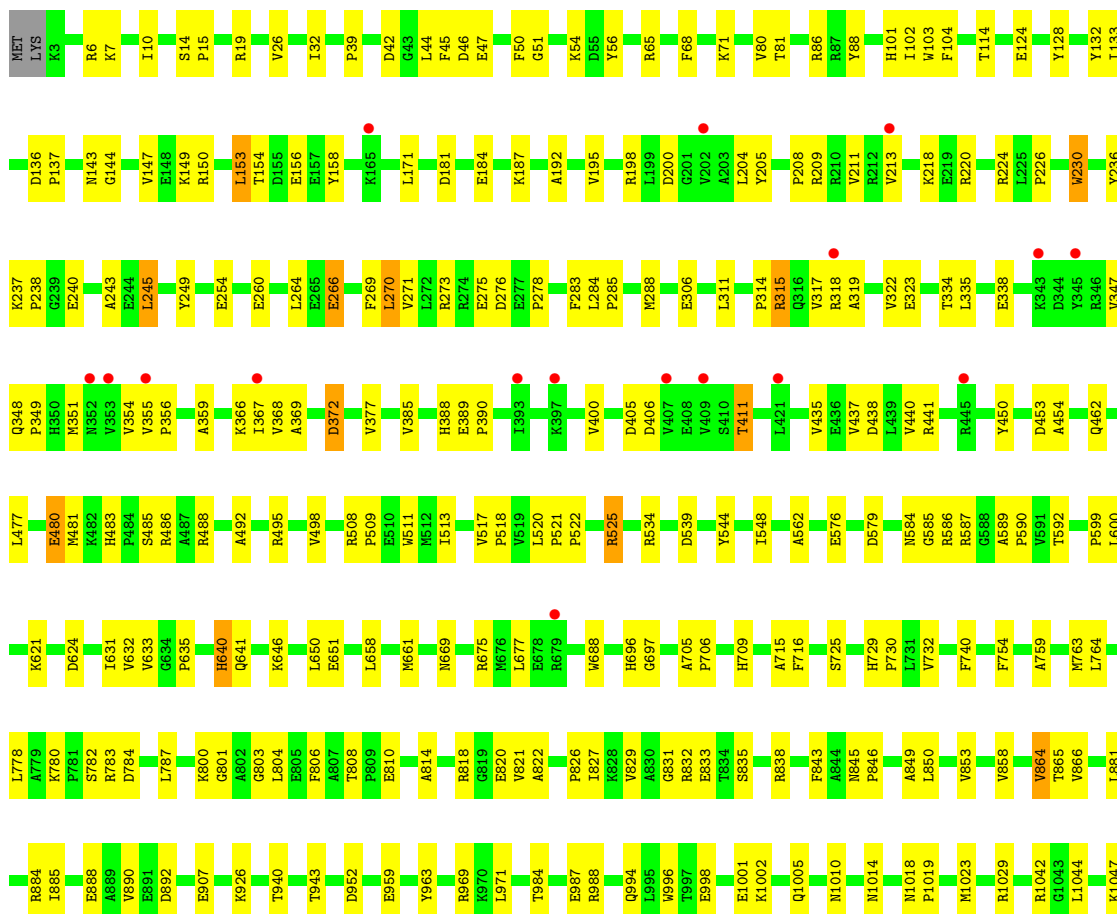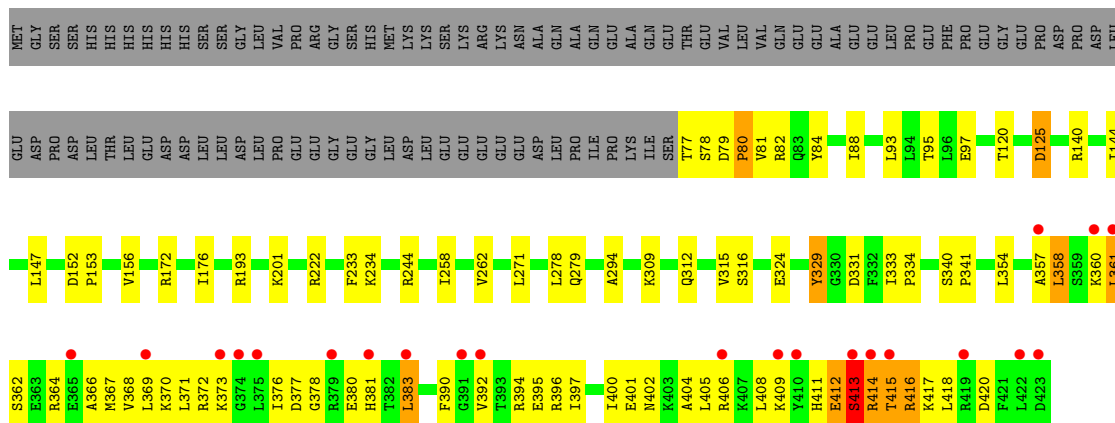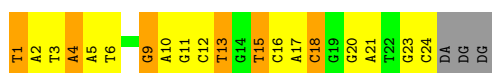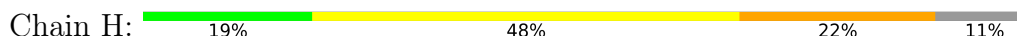• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain M:

• Molecule 3: DNA-directed RNA polymerase subunit beta'
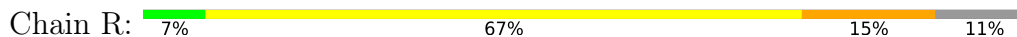
● Molecule 3: DNA-directed RNA polymerase subunit beta'

Chain N:

• Molecule 4: DNA-directed RNA polymerase subunit omega

• Molecule 4: DNA-directed RNA polymerase subunit omega

• Molecule 5: RNA polymerase sigma factor

• Molecule 5: RNA polymerase sigma factor

Chain P:

• Molecule 6: 5'-D(*TP*AP*TP*AP*AP*TP*GP*GP*GP*AP*GP*CP*TP*GP*TP*CP*AP*C
P*GP*GP*AP*TP*GP*CP*AP*GP*G)-3'

Chain H:

• Molecule 6: 5'-D(*TP*AP*TP*AP*AP*TP*GP*GP*GP*AP*GP*CP*TP*GP*TP*CP*AP*C
P*GP*GP*AP*TP*GP*CP*AP*GP*G)-3'

Chain R:

• Molecule 7: 5'-D(*CP*CP*TP*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*(BRU)P*CP*
GP*AP*GP*GP*G)-3'

Chain G:

• Molecule 7: 5'-D(*CP*CP*TP*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*(BRU)P*CP*
GP*AP*GP*GP*G)-3'

Chain Q:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	185.47Å 103.59Å 296.51Å 90.00° 98.53° 90.00°	Depositor
Resolution (Å)	49.85 – 3.81 49.85 – 3.81	Depositor EDS
% Data completeness (in resolution range)	51.6 (49.85-3.81) 82.2 (49.85-3.81)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	0.17	Depositor
$< I/\sigma(I) >$ [1]	3.44 (at 3.77Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.176 , 0.232 0.192 , 0.256	Depositor DCC
R_{free} test set	1776 reflections (1.82%)	wwPDB-VP
Wilson B-factor (Å^2)	120.0	Xtriage
Anisotropy	0.137	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 99.0	EDS
L-test for twinning[2]	$< |L| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	56872	wwPDB-VP
Average B, all atoms (Å^2)	151.0	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 40.98 % 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 2.5249e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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: BRU, ZN, 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	0.46	0/1814	0.66	0/2466
1	B	0.45	0/1782	0.67	0/2424
1	K	0.26	0/1814	0.45	0/2466
1	L	0.26	0/1805	0.47	0/2454
2	C	0.47	0/8937	0.70	0/12087
2	M	0.26	0/8937	0.48	1/12087 (0.0%)
3	D	0.48	0/11910	0.69	1/16105 (0.0%)
3	N	0.25	0/11952	0.47	0/16162
4	E	0.45	0/775	0.66	0/1045
4	O	0.24	0/775	0.43	0/1045
5	F	0.25	0/2852	0.44	0/3837
5	P	0.26	0/2859	0.45	0/3847
6	H	0.89	0/556	1.73	20/858 (2.3%)
6	R	1.00	2/556 (0.4%)	1.66	15/858 (1.7%)
7	G	0.82	0/345	1.34	4/529 (0.8%)
7	Q	0.79	0/345	1.37	6/529 (1.1%)
All	All	0.39	2/58014 (0.0%)	0.64	47/78799 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	M	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	R	5	DA	N9-C4	5.04	1.40	1.37
6	R	12	DC	C1'-N1	5.01	1.55	1.49

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	H	12	DC	O4'-C1'-N1	10.39	115.27	108.00
6	R	17	DA	O4'-C1'-N9	9.91	114.94	108.00
6	H	18	DC	O4'-C4'-C3'	-9.12	100.53	106.00
2	M	422	ARG	NE-CZ-NH2	-8.82	115.89	120.30
6	R	23	DG	O4'-C1'-N9	8.35	113.84	108.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	M	294	GLU	Peptide

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	1782	0	1834	69	0
1	B	1750	0	1797	63	0
1	K	1782	0	1834	32	0
1	L	1773	0	1826	35	0
2	C	8770	0	8874	308	0
2	M	8770	0	8874	214	0
3	D	11704	0	11934	397	0
3	N	11746	0	11974	270	0
4	E	761	0	778	31	0
4	O	761	0	778	18	0
5	F	2807	0	2882	74	0
5	P	2814	0	2889	120	0
6	H	495	0	272	18	0
6	R	495	0	272	15	0
7	G	328	0	178	7	0
7	Q	328	0	178	13	0
8	D	2	0	0	0	0
8	N	2	0	0	0	0
9	D	1	0	0	0	0
9	N	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	56872	0	57174	1519	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 13.

The worst 5 of 1519 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)
5:P:408:LEU:O	5:P:412:GLU:CB	1.72	1.37
5:P:408:LEU:O	5:P:412:GLU:HB2	1.07	1.23
5:P:408:LEU:HD23	5:P:412:GLU:CG	1.69	1.21
7:Q:14:DG:H2"	7:Q:15:BRU:O5'	1.43	1.15
3:N:1238:MET:SD	3:N:1359:GLN:OE1	2.07	1.12

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	224/315 (71%)	219 (98%)	3 (1%)	2 (1%)	17	54
1	B	220/315 (70%)	209 (95%)	11 (5%)	0	100	100
1	K	224/315 (71%)	220 (98%)	4 (2%)	0	100	100
1	L	223/315 (71%)	214 (96%)	9 (4%)	0	100	100
2	C	1107/1119 (99%)	1066 (96%)	40 (4%)	1 (0%)	51	83
2	M	1107/1119 (99%)	1065 (96%)	40 (4%)	2 (0%)	47	78
3	D	1478/1524 (97%)	1418 (96%)	59 (4%)	1 (0%)	51	83
3	N	1485/1524 (97%)	1423 (96%)	59 (4%)	3 (0%)	47	78
4	E	92/99 (93%)	88 (96%)	4 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	O	92/99 (93%)	88 (96%)	4 (4%)	0	100	100
5	F	344/443 (78%)	337 (98%)	7 (2%)	0	100	100
5	P	345/443 (78%)	331 (96%)	12 (4%)	2 (1%)	25	62
All	All	6941/7630 (91%)	6678 (96%)	252 (4%)	11 (0%)	47	78

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	325	ILE
2	M	295	ASP
2	M	325	ILE
5	P	413	SER
3	N	1234	THR

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	199/273 (73%)	185 (93%)	14 (7%)	15	45
1	B	195/273 (71%)	182 (93%)	13 (7%)	16	47
1	K	199/273 (73%)	191 (96%)	8 (4%)	31	59
1	L	198/273 (72%)	192 (97%)	6 (3%)	41	66
2	C	936/941 (100%)	854 (91%)	82 (9%)	10	38
2	M	936/941 (100%)	903 (96%)	33 (4%)	36	63
3	D	1250/1279 (98%)	1139 (91%)	111 (9%)	9	37
3	N	1253/1279 (98%)	1197 (96%)	56 (4%)	27	57
4	E	83/88 (94%)	76 (92%)	7 (8%)	11	40
4	O	83/88 (94%)	81 (98%)	2 (2%)	49	71
5	F	301/388 (78%)	291 (97%)	10 (3%)	38	64
5	P	302/388 (78%)	285 (94%)	17 (6%)	21	52
All	All	5935/6484 (92%)	5576 (94%)	359 (6%)	19	50

5 of 359 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	F	138	SER
3	N	153	LEU
1	K	34	VAL
2	M	402	SER
3	N	450	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	M	704	HIS
3	N	483	HIS
3	N	1404	ASN
3	D	824	ASN
3	D	1172	HIS

5.3.3 RNA ⓘ

There are no RNA molecules 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	#$|Z| > 2$
7	BRU	G	15	7	18,21,22	2.76	8 (44%)	26,30,33	2.14	5 (19%)
7	BRU	Q	15	7	18,21,22	2.75	8 (44%)	26,30,33	2.11	5 (19%)

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
7	BRU	G	15	7	-	<mark>2/7/21/22</mark>	0/2/2/2
7	BRU	Q	15	7	-	0/7/21/22	0/2/2/2

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	G	15	BRU	O4-C4	6.78	1.36	1.23
7	Q	15	BRU	O4-C4	6.76	1.36	1.23
7	G	15	BRU	C6-N1	5.34	1.47	1.38
7	Q	15	BRU	C6-N1	5.32	1.47	1.38
7	Q	15	BRU	C2-N1	3.45	1.44	1.38

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	G	15	BRU	C5-C4-N3	5.83	120.05	113.34
7	Q	15	BRU	C5-C4-N3	5.75	119.95	113.34
7	G	15	BRU	C4-N3-C2	-5.72	119.94	127.35
7	Q	15	BRU	C4-N3-C2	-5.66	120.02	127.35
7	G	15	BRU	O4-C4-C5	-4.69	119.95	125.84

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	G	15	BRU	C3'-C4'-C5'-O5'
7	G	15	BRU	O4'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	G	15	BRU	1	0
7	Q	15	BRU	10	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

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	226/315 (71%)	-0.28	1 (0%)	92	88	112, 138, 163, 174	0
1	B	222/315 (70%)	-0.34	0	100	100	110, 145, 190, 215	0
1	K	226/315 (71%)	-0.23	3 (1%)	77	70	114, 145, 171, 193	0
1	L	225/315 (71%)	-0.23	1 (0%)	92	88	114, 150, 198, 222	0
2	C	1111/1119 (99%)	-0.18	10 (0%)	84	78	89, 136, 212, 240	0
2	M	1111/1119 (99%)	0.04	56 (5%)	28	24	98, 154, 239, 269	0
3	D	1482/1524 (97%)	-0.21	12 (0%)	86	81	86, 136, 196, 258	1 (0%)
3	N	1489/1524 (97%)	-0.12	26 (1%)	70	62	84, 139, 199, 263	1 (0%)
4	E	94/99 (94%)	-0.08	1 (1%)	80	74	107, 138, 175, 185	0
4	O	94/99 (94%)	0.03	4 (4%)	35	29	109, 141, 191, 201	0
5	F	346/443 (78%)	-0.03	20 (5%)	23	18	112, 152, 238, 251	0
5	P	347/443 (78%)	-0.02	22 (6%)	20	15	117, 161, 253, 290	0
6	H	24/27 (88%)	-0.36	0	100	100	144, 191, 241, 247	0
6	R	24/27 (88%)	-0.38	0	100	100	142, 204, 241, 254	0
7	G	15/21 (71%)	-0.29	1 (6%)	17	13	146, 181, 229, 233	0
7	Q	15/21 (71%)	-0.36	0	100	100	160, 197, 234, 235	0
All	All	7051/7726 (91%)	-0.13	157 (2%)	62	53	84, 144, 219, 290	2 (0%)

The worst 5 of 157 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	N	1247	ALA	5.2
2	M	207	LEU	5.0
3	N	1238	MET	5.0
2	M	221	LEU	4.9
2	M	181	VAL	4.9

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. 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	BRU	Q	15	20/21	0.83	0.24	177,196,247,303	1
7	BRU	G	15	20/21	0.89	0.25	138,145,192,257	1

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
9	MG	D	2003	1/1	0.94	0.23	111,111,111,111	0
9	MG	N	2003	1/1	0.98	0.22	127,127,127,127	0
8	ZN	N	2002	1/1	0.99	0.06	138,138,138,138	0
8	ZN	D	2002	1/1	0.99	0.09	164,164,164,164	0
8	ZN	N	2001	1/1	0.99	0.22	166,166,166,166	0
8	ZN	D	2001	1/1	1.00	0.17	168,168,168,168	0

6.5 Other polymers ⓘ

There are no such residues in this entry.