



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 05:11 am BST

PDB ID : 4G7O
Title : Crystal structure of Thermus thermophilus transcription initiation complex containing 2 nt of RNA
Authors : Zhang, Y.; Ebright, R.H.
Deposited on : 2012-07-20
Resolution : 2.99 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

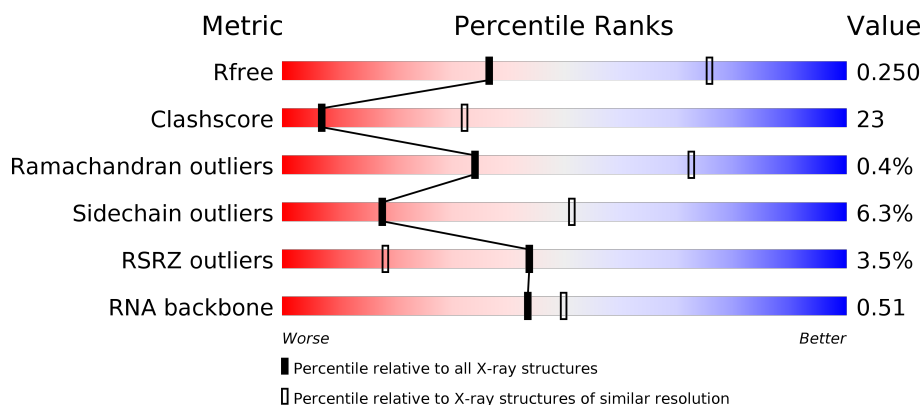
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.99 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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)
RNA backbone	3102	1173 (3.30-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria 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 $\leq 5\%$. 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	
1	B	315	
1	K	315	
1	L	315	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	1119	
2	M	1119	
3	D	1524	
3	N	1524	
4	E	99	
4	O	99	
5	F	443	
5	P	443	
6	G	21	
6	Q	21	
7	H	27	
7	R	27	
8	I	2	
8	S	2	

2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 57231 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	C	N	O	S	0	0	0
			1782	1138	310	332	2			
1	B	222	Total	C	N	O	S	0	0	0
			1750	1118	304	326	2			
1	K	226	Total	C	N	O	S	0	0	0
			1782	1138	310	332	2			
1	L	225	Total	C	N	O	S	0	0	0
			1773	1133	308	330	2			

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

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

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

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

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

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

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

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

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(*CP*CP*T*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*CP*GP*AP*GP*GP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	16	Total	C	N	O	P	0	0	0
			328	156	63	94	15			
6	Q	16	Total	C	N	O	P	0	0	0
			328	156	63	94	15			

- Molecule 7 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
7	H	24	Total	C	N	O	P	0	0	0
			495	236	94	142	23			
7	R	24	Total	C	N	O	P	0	0	0
			495	236	94	142	23			

- Molecule 8 is a RNA chain called 5'-R(*GP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	I	2	Total	C	N	O	P	0	0	0
			42	20	10	11	1			
8	S	2	Total	C	N	O	P	0	0	0
			42	20	10	11	1			

- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	D	2	Total Mg 2 2	0	0
10	K	1	Total Mg 1 1	0	0
10	N	2	Total Mg 2 2	0	0

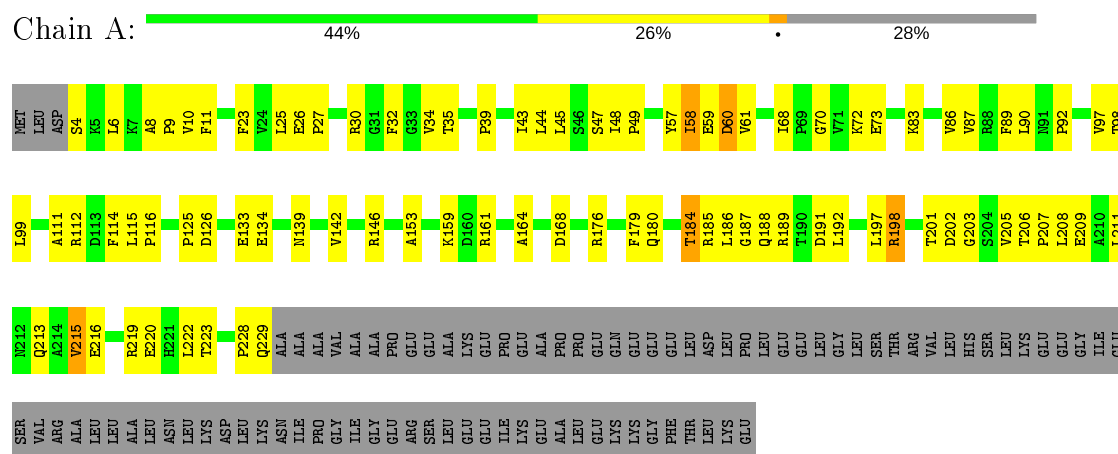
- Molecule 11 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	14	Total O 14 14	0	0
11	B	3	Total O 3 3	0	0
11	C	53	Total O 53 53	0	0
11	D	58	Total O 58 58	0	0
11	E	8	Total O 8 8	0	0
11	F	10	Total O 10 10	0	0
11	G	5	Total O 5 5	0	0
11	H	2	Total O 2 2	0	0
11	K	4	Total O 4 4	0	0
11	L	3	Total O 3 3	0	0
11	M	33	Total O 33 33	0	0
11	N	52	Total O 52 52	0	0
11	O	5	Total O 5 5	0	0
11	P	16	Total O 16 16	0	0
11	Q	2	Total O 2 2	0	0
11	R	3	Total O 3 3	0	0
11	S	1	Total O 1 1	0	0

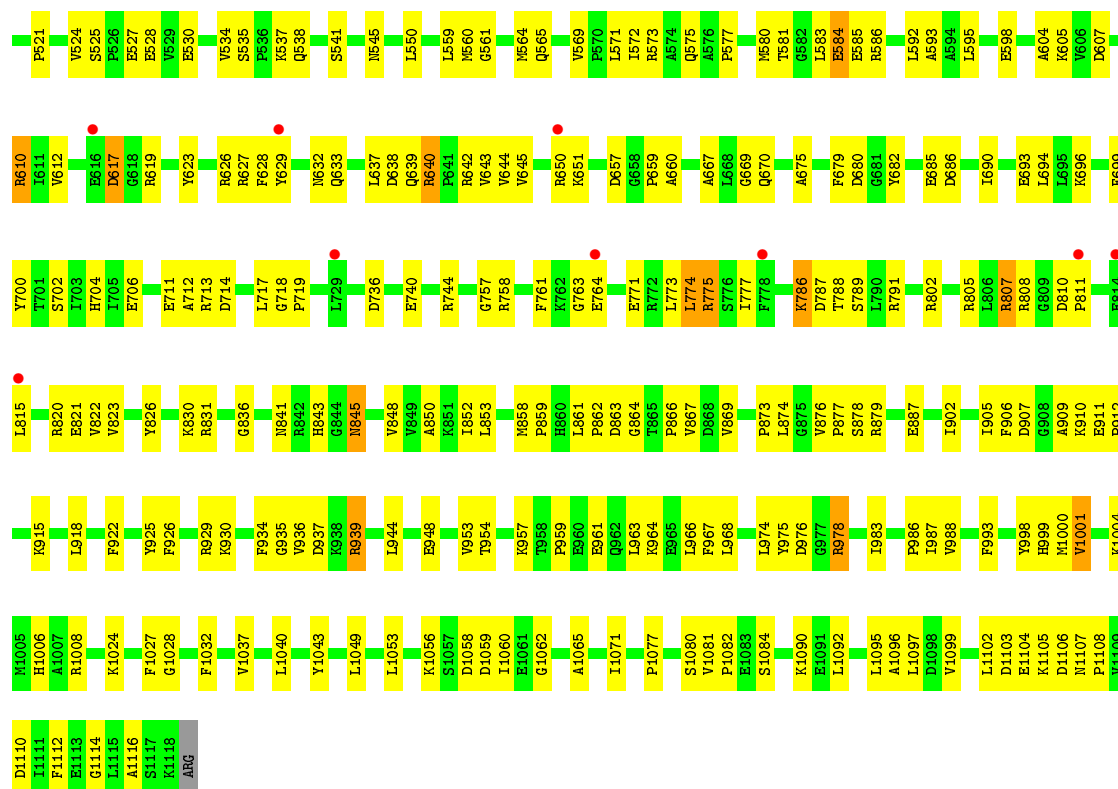
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

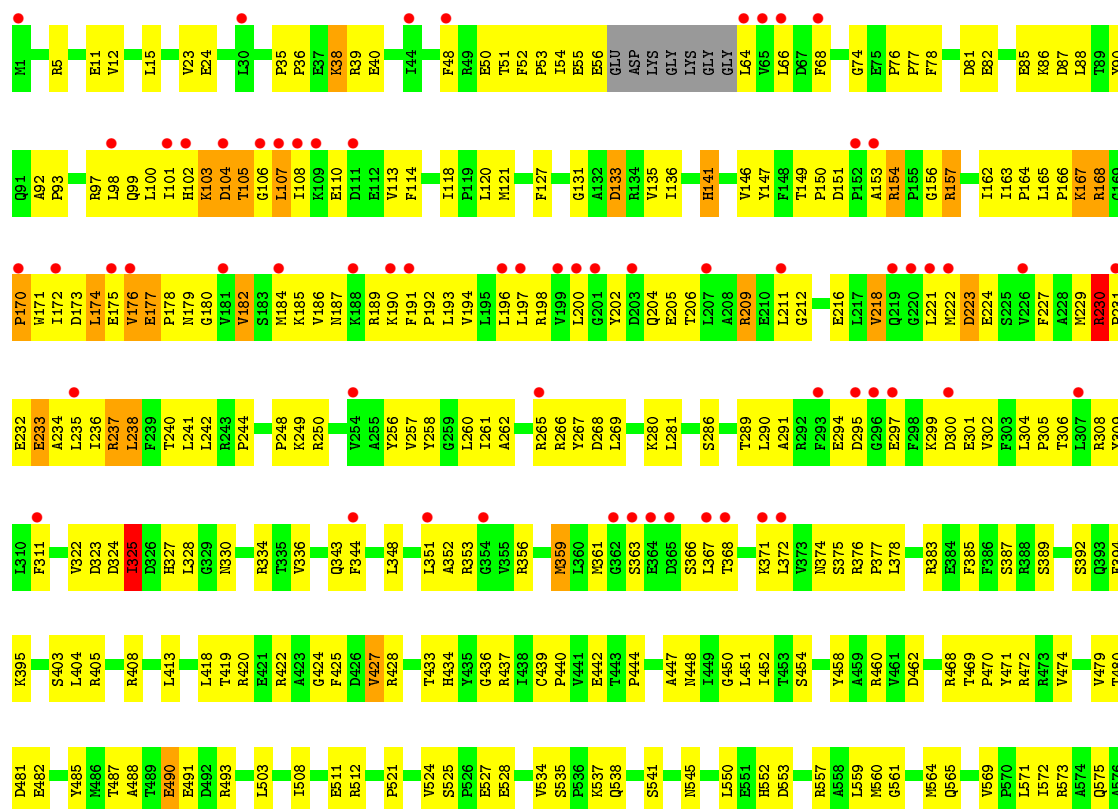
• Molecule 1: DNA-directed RNA polymerase subunit alpha



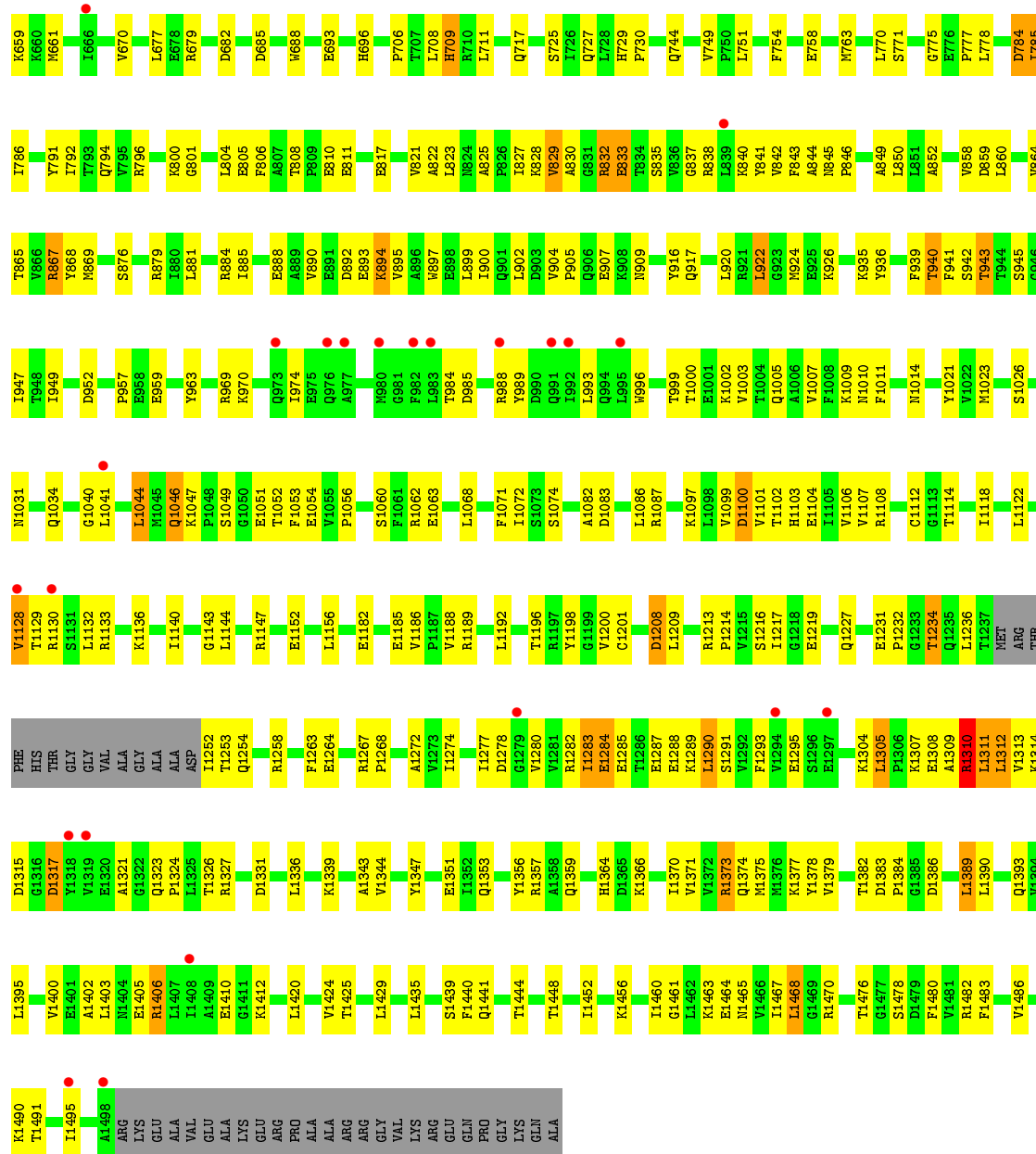




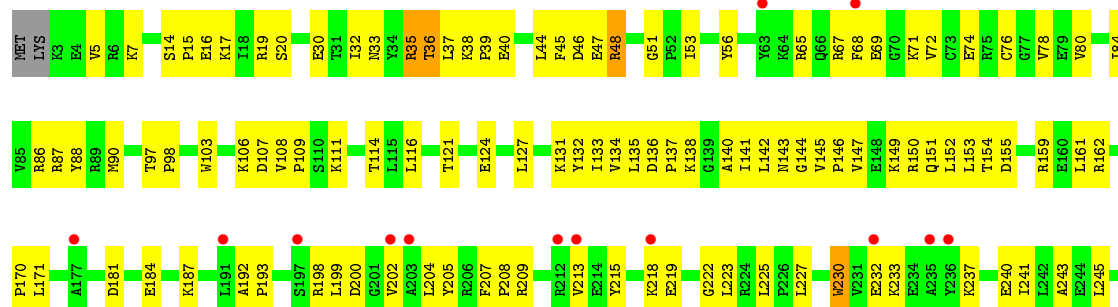
• Molecule 2: DNA-directed RNA polymerase subunit beta



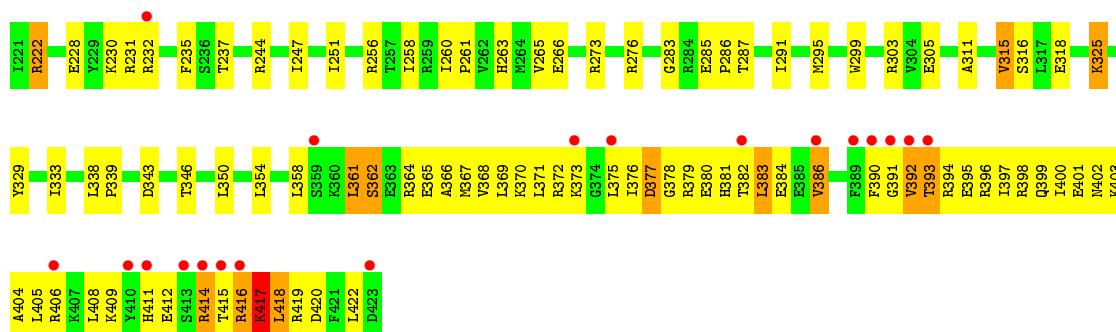




• Molecule 3: DNA-directed RNA polymerase subunit beta'



V1424	V1439	F1440	Q1441	T1444	T1448	I1452	L1459	L1462	K1463	L1464	H1465	V1466	I1467	L1468	G1469	R1470	L1471	I1472	P1473	G1477	S1478	P1479	D1479	D1480	D1481	D1482	F1483	V1486	V1487	T1491	K1492	K1493	A1494	I1495	E1496	E1497	A1498	ARG	LYS	GLU	ALA	VAL	ALA	LYS	GLU	ARG	PRO	ALA	ALA				
V1344	Y1347	E1351	I1352	Q1353	Y1356	R1357	A1358	Q1359	K1362	K1366	H1367	I1368	V1371	V1372	R1373	Q1374	M1375	E1380	V1381	T1382	P1383	P1384	G1385	D1386	E1393	L1394	L1395	V1400	L1403	I1404	E1405	L1406	I1407	I1408	A1409	E1410	G1411	K1412	T1413	P1414	V1415	H1416	K1417	K1418	E1419	L1420							
I1274	S1275	E1276	I1277	D1278	G1279	V1280	V1281	R1282	I1283	E1284	E1285	T1286	E1287	E1288	K1289	L1290	S1291	V1292	V1293	V1294	E1295	F1299	S1300	K1301	E1302	Y1303	L1304	L1305	P1306	K1307	E1308	A1309	R1310	L1311	L1312	V1313	K1314	D1315	G1316	D1317	Y1318	V1319	E1320	A1321	P1324	L1325	T1326	R1327	D1331	L1336	K1339	E1342	A1343
V1188	R1189	C1194	Y1198	Y1199	V1200	C1204	Y1205	D1208	R1213	P1214	V1215	E1219	Q1227	E1231	P1232	G1233	T1234	T1237	M1238	ARG	THR	PHE	HIS	THR	GLY	V1246	A1247	G1248	A1249	A1250	D1251	I1252	T1253	Q1254	G1255	L1256	P1257	R1258	V1259	I1260	E1261	L1262	F1263	E1264	R1267	P1268							
S1060	F1061	R1062	L1068	F1071	I1072	S1073	S1074	A1082	D1083	L1086	Y1093	K1097	L1098	V1099	D1100	A1106	V1101	F1108	K1109	M1101	F1101	M1104	L11020	Y11021	V11022	S11026	G11027	A11028	M11031	Q11034	G11040	L11041	L11044	M1045	Q1046	K1047	P1048	S1049	G1050	E1051	T1052	F1053	E1054	Y1055	P1056								
R884	I885	E888	A889	V890	E891	D892	K894	M897	E898	L899	I900	Q901	L902	L903	V904	P905	Q906	E907	K908	N909	Y916	Q917	L920	R921	L922	K926	L931	K935	Y936	F939	T940	F941	S942	T943	T944	S945	G946	I947	T948	I949	D952	P957	E958	E959	Y963	K970							
G901	L904	E905	F906	T908	F909	E910	E911	E917	E920	V921	A922	L923	N924	A925	P926	I927	K928	V929	A930	G931	R932	E933	T934	S935	V936	R937	L939	K940	Y941	F943	A944	N945	P946	L950	V958	D959	L960	V964	T965	V966	R967	Y968	M969	S976	R979	L981	K900						
Q880	R881	D882	D885	M888	E893	P706	L707	L708	H709	L710	L711	Q717	P718	V719	L720	V721	S725	L726	L727	L728	H729	L639	H640	L731	Q641	C642	G643	K646	R647	M648	A649	L650	E651	F652	G654	P655	F656	L657	L658	K659	K660	M661	N669	V670	K671	L792	L793	Q794	L795	R796	K900		
P599	L600	R601	S602	L603	I606	R613	F614	L618	G620	K621	R622	V623	D624	Y625	V630	L631	V632	V633	G634	L637	K638	L639	H640	L731	Q641	C642	G643	K646	R647	M648	A649	L650	E651	F652	G654	P655	F656	L657	L658	K659	K660	M661	N669	V670	K671	L792	L793	Q794	L795	R796	K900		
R495	L496	E497	V498	N511	N512	L513	P518	V519	L520	P521	R525	V530	R534	F535	A536	N541	R545	L545	D453	A454	R455	N456	Q457	A460	Q463	L465	L468	A472	L473	E474	K475	E476	L477	L478	E479	N480	N584	G585	R586	R587	V591	L592	N593	P594	G595	S596	D597	R598					
L333	T334	L337	E338	K339	V347	Q348	H349	K350	K351	K352	V353	V354	P356	A359	I367	V368	I371	D372	P373	E376	I378	A379	E380	V385	E388	E389	P390	A391	S392	L393	L394	V395	V396	R399	V400	Y401	D405	D406	V407	S408	R486	T411	R488	G412	P413	R414	L421						
Y249	L250	F251	R252	E255	E256	V259	E260	L264	E265	E266	F269	L270	V271	L272	R273	R274	E275	D276	E277	P278	V279	A280	T281	F282	Y283	L284	P285	V286	V298	E299	K300	G301	Q302	P303	E306	L310	L311	R312	N313	P314	R315	Q316	V317	A319	E323	E326	V331	Y332					



- Molecule 6: 5'-D(*CP*CP*T*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*CP*GP*AP*GP*GP*G)-3'

Chain G: 48% 19% 10% 24%



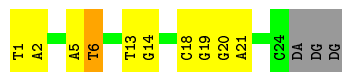
- Molecule 6: 5'-D(*CP*CP*T*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*CP*GP*AP*GP*GP*G)-3'

Chain Q: 33% 43% 24%



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

Chain H: 52% 33% 11%



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

Chain R: 52% 37% 11%



- Molecule 8: 5'-R(*GP*A)-3'

Chain I: 100%

There are no outlier residues recorded for this chain.

- Molecule 8: 5'-R(*GP*A)-3'

Chain S:  100%

There are no outlier residues recorded for this chain.

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	185.48 Å 103.88 Å 297.34 Å 90.00° 98.23° 90.00°	Depositor
Resolution (Å)	46.61 – 2.99 49.59 – 2.99	Depositor EDS
% Data completeness (in resolution range)	95.6 (46.61-2.99) 95.5 (49.59-2.99)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	0.17	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.99 (at 3.01 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.205 , 0.247 0.207 , 0.250	Depositor DCC
R_{free} test set	1946 reflections (0.88%)	wwPDB-VP
Wilson B-factor (Å ²)	65.4	Xtriage
Anisotropy	0.459	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 42.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	57231	wwPDB-VP
Average B, all atoms (Å ²)	77.0	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 36.33 % 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 5.0787e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.29	0/1814	0.50	0/2466
1	B	0.27	0/1782	0.49	0/2424
1	K	0.27	0/1814	0.49	0/2466
1	L	0.28	0/1805	0.51	0/2454
2	C	0.30	0/8937	0.49	0/12087
2	M	0.31	0/8937	0.50	0/12087
3	D	0.31	1/11910 (0.0%)	0.49	0/16105
3	N	0.31	1/11952 (0.0%)	0.50	0/16162
4	E	0.28	0/775	0.44	0/1045
4	O	0.26	0/775	0.44	0/1045
5	F	0.28	0/2852	0.46	0/3837
5	P	0.28	0/2859	0.45	0/3847
6	G	0.55	0/368	1.09	2/567 (0.4%)
6	Q	0.47	0/368	1.05	0/567
7	H	0.51	0/556	1.26	3/858 (0.3%)
7	R	0.52	0/556	1.23	2/858 (0.2%)
8	I	0.35	0/47	0.68	0/72
8	S	0.25	0/47	0.55	0/72
All	All	0.31	2/58154 (0.0%)	0.53	7/79019 (0.0%)

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
5	P	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	N	599	PRO	C-N	5.34	1.46	1.34
3	D	105	VAL	C-N	5.13	1.45	1.34

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	R	17	DA	O4'-C1'-N9	6.77	112.74	108.00
7	H	6	DT	O4'-C1'-N1	-6.45	103.49	108.00
7	H	20	DG	C1'-O4'-C4'	-6.17	103.93	110.10
6	G	16	DC	O4'-C4'-C3'	-5.77	102.19	104.50
7	R	20	DG	C1'-O4'-C4'	-5.40	104.70	110.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	M	104	ASP	Peptide
5	P	82	ARG	Sidechain

5.2 Too-close contacts [i](#)

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	81	0
1	B	1750	0	1797	92	0
1	K	1782	0	1834	60	0
1	L	1773	0	1826	76	0
2	C	8770	0	8872	436	0
2	M	8770	0	8871	518	2
3	D	11704	0	11934	544	1
3	N	11746	0	11974	664	0
4	E	761	0	778	26	0
4	O	761	0	778	25	0
5	F	2807	0	2882	111	0
5	P	2814	0	2888	168	1
6	G	328	0	181	4	0
6	Q	328	0	181	6	0
7	H	495	0	272	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	R	495	0	272	14	0
8	I	42	0	23	0	0
8	S	42	0	23	0	0
9	D	2	0	0	0	0
9	N	2	0	0	0	0
10	D	2	0	0	0	0
10	K	1	0	0	0	0
10	N	2	0	0	0	0
11	A	14	0	0	0	0
11	B	3	0	0	0	0
11	C	53	0	0	0	0
11	D	58	0	0	0	0
11	E	8	0	0	0	0
11	F	10	0	0	0	0
11	G	5	0	0	0	0
11	H	2	0	0	0	0
11	K	4	0	0	0	0
11	L	3	0	0	0	0
11	M	33	0	0	1	0
11	N	52	0	0	1	0
11	O	5	0	0	0	0
11	P	16	0	0	0	0
11	Q	2	0	0	0	0
11	R	3	0	0	0	0
11	S	1	0	0	0	0
All	All	57231	0	57220	2617	2

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 23.

The worst 5 of 2617 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)
2:M:157:ARG:NH1	2:M:176:VAL:HG11	1.23	1.47
2:C:64:LEU:N	2:C:103:LYS:NZ	1.61	1.47
2:M:64:LEU:N	2:M:103:LYS:NZ	1.64	1.45
2:M:107:LEU:HD12	2:M:108:ILE:N	1.11	1.40
2:M:740:GLU:OE1	2:M:805:ARG:NH1	1.57	1.36

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:1314:LYS:NZ	2:M:784:ASP:OD2[1_445]	1.83	0.37
2:M:40:GLU:OE1	5:P:364:ARG:NH1[1_545]	2.03	0.17

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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%)	218 (97%)	5 (2%)	1 (0%)	34	72
1	B	220/315 (70%)	212 (96%)	6 (3%)	2 (1%)	17	55
1	K	224/315 (71%)	218 (97%)	6 (3%)	0	100	100
1	L	223/315 (71%)	213 (96%)	8 (4%)	2 (1%)	17	55
2	C	1107/1119 (99%)	1066 (96%)	37 (3%)	4 (0%)	34	72
2	M	1107/1119 (99%)	1063 (96%)	39 (4%)	5 (0%)	29	68
3	D	1478/1524 (97%)	1423 (96%)	53 (4%)	2 (0%)	51	85
3	N	1485/1524 (97%)	1423 (96%)	57 (4%)	5 (0%)	41	76
4	E	92/99 (93%)	89 (97%)	3 (3%)	0	100	100
4	O	92/99 (93%)	88 (96%)	4 (4%)	0	100	100
5	F	344/443 (78%)	335 (97%)	9 (3%)	0	100	100
5	P	345/443 (78%)	329 (95%)	12 (4%)	4 (1%)	13	48
All	All	6941/7630 (91%)	6677 (96%)	239 (3%)	25 (0%)	34	72

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	325	ILE
2	M	325	ILE
5	P	78	SER
5	P	417	LYS
3	D	1310	ARG

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%)	189 (95%)	10 (5%)	24	60
1	B	195/273 (71%)	182 (93%)	13 (7%)	16	49
1	K	199/273 (73%)	189 (95%)	10 (5%)	24	60
1	L	198/273 (72%)	188 (95%)	10 (5%)	24	60
2	C	936/941 (100%)	880 (94%)	56 (6%)	19	53
2	M	936/941 (100%)	873 (93%)	63 (7%)	16	49
3	D	1250/1279 (98%)	1165 (93%)	85 (7%)	16	48
3	N	1253/1279 (98%)	1169 (93%)	84 (7%)	16	49
4	E	83/88 (94%)	80 (96%)	3 (4%)	35	70
4	O	83/88 (94%)	81 (98%)	2 (2%)	49	79
5	F	301/388 (78%)	287 (95%)	14 (5%)	26	63
5	P	302/388 (78%)	280 (93%)	22 (7%)	14	44
All	All	5935/6484 (92%)	5563 (94%)	372 (6%)	18	51

5 of 372 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	F	117	SER
2	M	107	LEU
3	N	1389	LEU
5	F	218	GLN
1	K	185	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
2	M	187	ASN
2	M	330	ASN
3	N	1195	GLN
3	D	1172	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	221	HIS

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	I	1/2 (50%)	0	0
8	S	1/2 (50%)	0	0
All	All	2/4 (50%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 9 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 [i](#)

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	0 100 100	47, 67, 98, 118	0
1	B	222/315 (70%)	-0.17	2 (0%) 84 63	50, 79, 119, 142	0
1	K	226/315 (71%)	-0.21	0 100 100	49, 73, 103, 122	0
1	L	225/315 (71%)	-0.14	3 (1%) 77 51	51, 81, 126, 148	0
2	C	1111/1119 (99%)	-0.03	29 (2%) 56 27	34, 62, 122, 168	0
2	M	1111/1119 (99%)	0.24	82 (7%) 14 4	38, 76, 142, 179	0
3	D	1482/1524 (97%)	0.02	47 (3%) 47 20	32, 66, 130, 178	1 (0%)
3	N	1489/1524 (97%)	0.06	55 (3%) 41 17	34, 69, 129, 201	1 (0%)
4	E	94/99 (94%)	-0.08	3 (3%) 47 20	44, 66, 113, 121	0
4	O	94/99 (94%)	-0.10	3 (3%) 47 20	49, 72, 116, 130	0
5	F	346/443 (78%)	-0.14	6 (1%) 70 41	42, 70, 124, 147	0
5	P	347/443 (78%)	0.04	19 (5%) 25 9	46, 83, 156, 214	0
6	G	16/21 (76%)	-0.49	0 100 100	42, 78, 172, 179	0
6	Q	16/21 (76%)	-0.63	0 100 100	53, 80, 182, 183	0
7	H	24/27 (88%)	-0.39	0 100 100	67, 95, 146, 201	0
7	R	24/27 (88%)	-0.39	0 100 100	65, 106, 160, 210	0
8	I	2/2 (100%)	-0.43	0 100 100	50, 50, 50, 55	0
8	S	2/2 (100%)	-0.41	0 100 100	62, 62, 62, 63	0
All	All	7057/7730 (91%)	0.01	249 (3%) 44 18	32, 71, 132, 214	2 (0%)

The worst 5 of 249 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	M	363	SER	6.5
5	P	392	VAL	6.1
5	P	411	HIS	6.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	M	191	PHE	5.9
5	P	375	LEU	5.6

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

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(\AA^2)	Q<0.9
10	MG	N	2003	1/1	0.78	0.20	46,46,46,46	0
10	MG	N	2004	1/1	0.82	0.37	65,65,65,65	0
10	MG	K	1001	1/1	0.89	0.66	69,69,69,69	0
10	MG	D	2003	1/1	0.94	0.24	40,40,40,40	0
10	MG	D	2004	1/1	0.97	0.44	63,63,63,63	0
9	ZN	D	2002	1/1	0.98	0.07	83,83,83,83	0
9	ZN	N	2001	1/1	0.98	0.20	62,62,62,62	0
9	ZN	N	2002	1/1	0.99	0.05	107,107,107,107	0
9	ZN	D	2001	1/1	0.99	0.13	62,62,62,62	0

6.5 Other polymers [i](#)

There are no such residues in this entry.