



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 10:44 AM GMT

PDB ID : 1SFO  
Title : RNA POLYMERASE II STRAND SEPARATED ELONGATION COMPLEX  
Authors : Westover, K.D.; Bushnell, D.A.; Kornberg, R.D.  
Deposited on : 2004-02-20  
Resolution : 3.61 Å(reported)

This is a wwPDB validation summary 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>

---

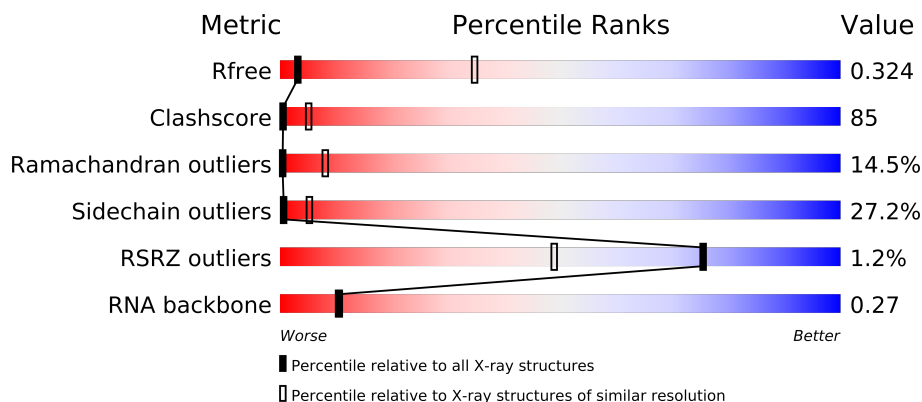
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.61 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	1006 (3.88-3.36)
Clashscore	79885	1171 (3.84-3.40)
Ramachandran outliers	78287	1125 (3.84-3.40)
Sidechain outliers	78261	1124 (3.84-3.40)
RSRZ outliers	66119	1007 (3.88-3.36)
RNA backbone	1838	1000 (4.46-2.80)

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	R	10	
2	T	14	
3	A	1733	
4	B	1224	
5	C	318	
6	E	215	
7	F	155	
8	H	146	
9	I	122	
10	J	70	
11	K	120	
12	L	70	

## 2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 28647 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 RNA chain called RNA STRAND.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	10	Total	C	N	O	P	0	0	0
			217	98	45	65	9			

- Molecule 2 is a DNA chain called DNA STRAND.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	14	Total	C	N	O	P	0	0	0
			279	135	48	83	13			

- Molecule 3 is a protein called DNA-directed RNA polymerase II largest subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	1395	Total	C	N	O	S	0	0	0
			10969	6917	1923	2068	61			

- Molecule 4 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	1106	Total	C	N	O	S	0	0	0
			8793	5568	1538	1632	55			

- Molecule 5 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	E	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	11			

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	F	84	Total	C	N	O	S	0	0	0
			679	434	115	127	3			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	133	Total	C	N	O	S	0	0	0
			1068	673	180	211	4			

- Molecule 9 is a protein called DNA-directed RNA polymerase II 14.2 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	119	Total	C	N	O	S	0	0	0
			971	596	179	186	10			

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 11 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	J	1	Total	Zn	0	0
			1	1		
13	B	1	Total	Zn	0	0
			1	1		
13	I	2	Total	Zn	0	0
			2	2		
13	C	1	Total	Zn	0	0
			1	1		
13	A	2	Total	Zn	0	0
			2	2		
13	L	1	Total	Zn	0	0
			1	1		

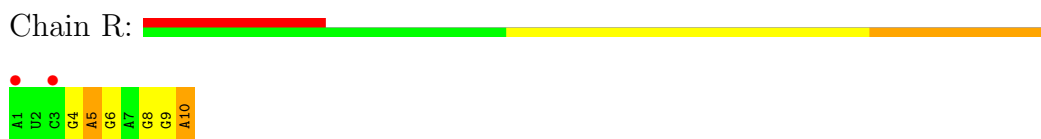
- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	1	Total	Mg	0	0
			1	1		

### 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.

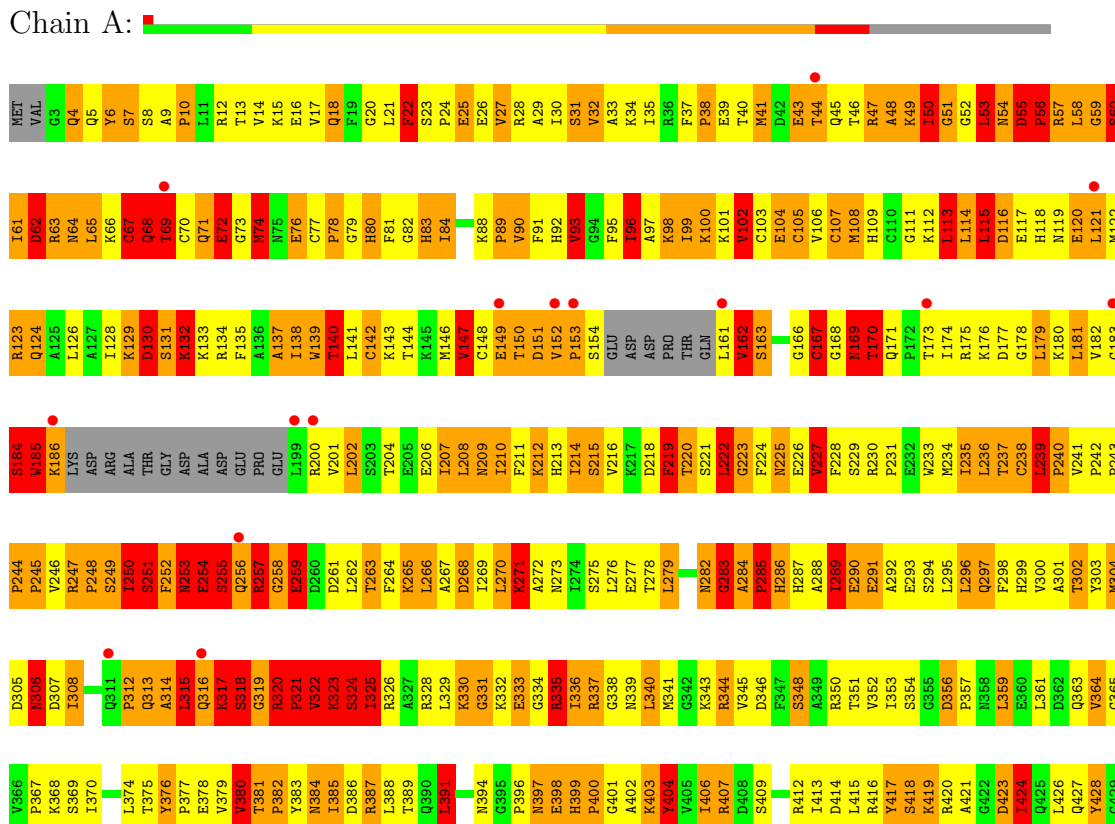
- Molecule 1: RNA STRAND



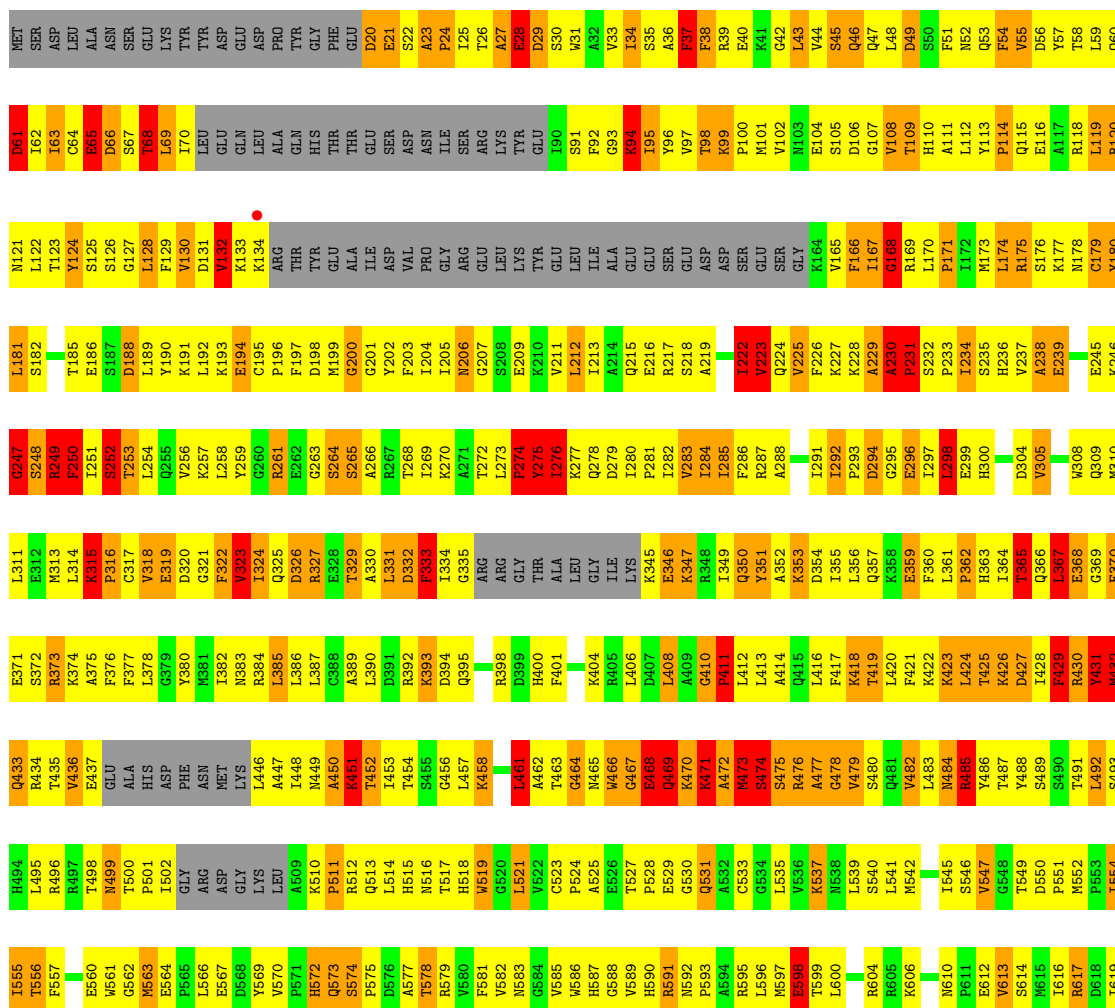
- Molecule 2: DNA STRAND



- Molecule 3: DNA-directed RNA polymerase II largest subunit



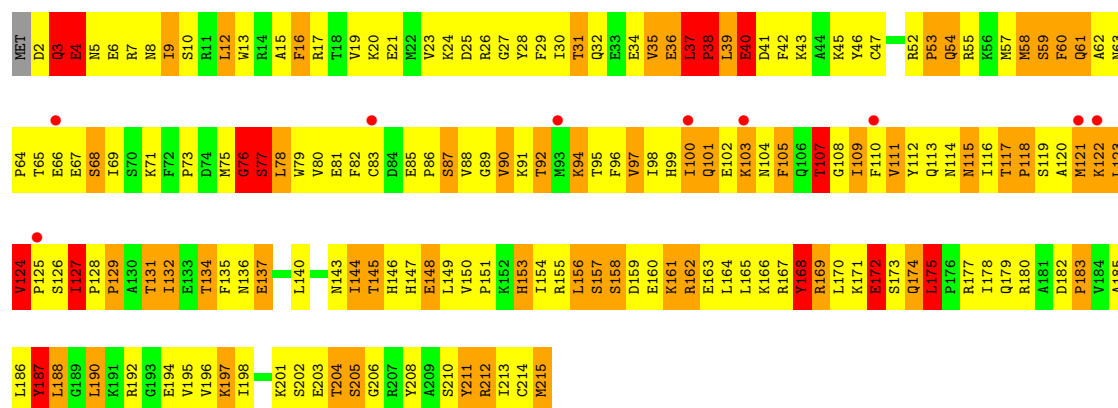
M1368	E1307	LYS	ASP	A1125	V1064	M1004	D939	E879	M818	S751	K687	B618	D557	S494	K430
A1369	T1308	SER	Q1187	A1126	G1065	E1005	R940	R890	G819	K752	K688	K619	G558	E495	K431
L1370	D1309	LEU	Q1188	D1127	V1066	E1006	R941	S881	G820	R753	K689	K620	V559	E496	V432
L1371	G1310	ASP	S1189	Q1128	A1067	T1007	F942	S882	R821	K754	V690	V622	I560	T497	E433
V1372	V1311	ALA	P1190	E1129	A1068	Q1008	L943	L883	G823	S754	L891	S625	P561	R498	R434
D1373	N1312	GLU	V1191	Q1130	A1069	M1009	R944	D884	G823	T755	D892	S625	T562	A499	R435
V1374	L1313	THR	L1192	Q1070	A1070	A1010	E945	T885	L824	T756	V693	S625	P563	E500	T436
M1375	S1314	GLU	K1132	K1073	G1073	R1012	F947	I886	I825	T757	T694	L629	A564	L501	K437
E1315	R1193	L1133	L1133	G1074	A1074	R1013		G887	D826	T759	K695	I630	I566	L504	D439
E1265	L1195	T1134	E1074	P1075	A1014	E1014	A952	G888	T827	A697	D896	I630	I567	C505	P441
E1266	L1135	R1135	A1076	P1076	V1015	V1015	N953	S889	A828	T765	Q698	V633	P568	A506	V442
D1257	S1136	S1136	T1077	Q1078	T1077	L1017	V954	A891	V829	V768	Q698	T634	P568	A506	V442
H1258	D1198	D1198	T1077	Q1078	T1077	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
M1259	R1199	R1199	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
L1260	A1200	A1200	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
K1261	A1201	A1201	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
K1262	M1202	M1202	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
K1263	M1203	M1203	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
K1264	D1204	D1204	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1265	K1205	K1205	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
N1266	D1206	D1206	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
N1267	L1207	L1207	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
M1268	T1208	T1208	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1269	M1209	M1209	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
N1270	G1210	G1210	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1271	S1150	S1150	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
T1271	E1151	E1151	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
L1272	I1152	I1152	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
K1273	V1153	V1153	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
L1274	V1154	V1154	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
K1275	D1155	D1155	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1276	P1156	P1156	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
V1276	D1157	D1157	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1277	F1158	F1158	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
N1278	L1159	L1159	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1280	S1160	S1160	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
R1281	T1161	T1161	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1282	V1162	V1162	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
V1283	L1163	L1163	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
H1284	P1164	P1164	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1285	E1165	E1165	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
K1286	D1166	D1166	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
Y1287	E1167	E1167	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1288	L1168	L1168	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
R1289	S1229	S1229	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
K1290	L1170	L1170	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
V1291	Q1171	Q1171	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
P1292	L1172	L1172	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
S1293	H1173	H1173	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
P1294	F1174	F1174	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
K1295	L1175	L1175	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1296	L1176	L1176	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
Y1298	LEU	LEU	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
P1299	ASP	ASP	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
K1300	GLU	GLU	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1301	C1240	C1240	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
P1302	ALA	ALA	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1303	ALA	ALA	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1305	GLU	GLU	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1306	GLN	GLN	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1307	SER	SER	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1308	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1309	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1310	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1311	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1312	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1313	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1314	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1315	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1316	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1317	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1318	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1319	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1320	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1321	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1322	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1323	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1324	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1325	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1326	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1327	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1328	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1329	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1330	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1331	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1332	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1333	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892	T831	V765	Q698	R635	P570	V507	L443
E1334	PRO	PRO	Q1078	Q1078	Q1078	L1017	P955	A892							





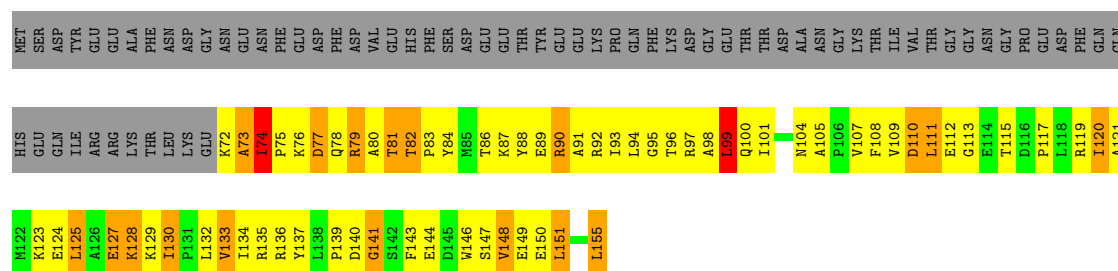


Chain E:



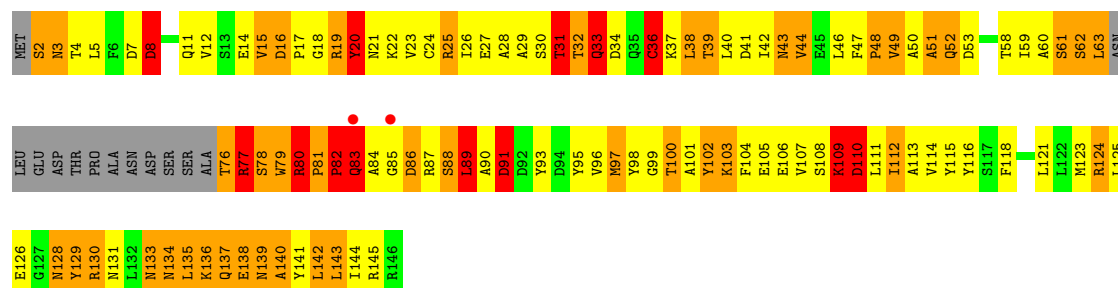
- Molecule 7: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

Chain F:



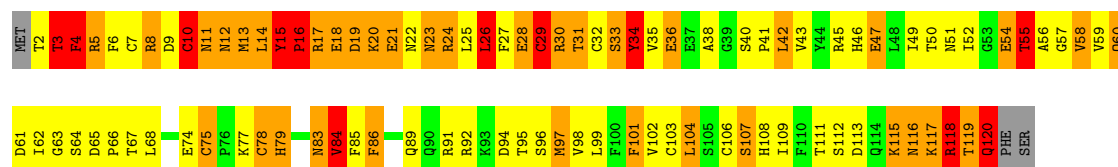
- Molecule 8: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

Chain H:



- Molecule 9: DNA-directed RNA polymerase II 14.2 kDa polypeptide

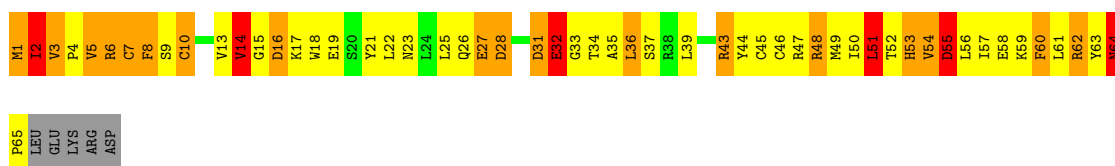
Chain I:



- Molecule 10: DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide

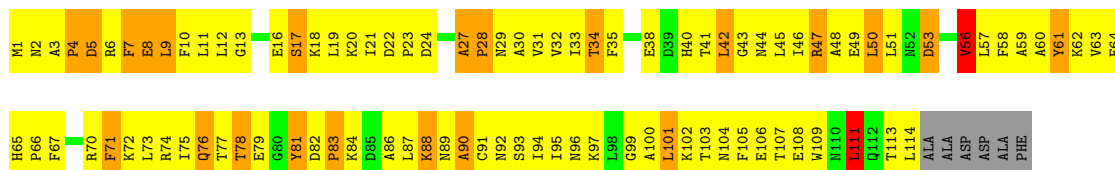
Chain J:





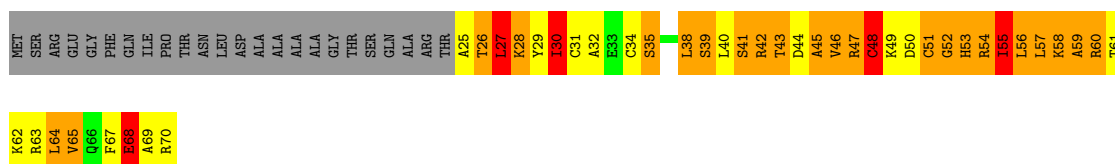
- Molecule 11: DNA-directed RNA polymerase II 13.6 kDa polypeptide

Chain K:



- Molecule 12: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide

Chain L:



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	167.08Å 221.26Å 193.69Å 90.00° 100.10° 90.00°	Depositor
Resolution (Å)	39.86 – 3.61 39.86 – 3.60	Depositor EDS
% Data completeness (in resolution range)	92.7 (39.86-3.61) 92.7 (39.86-3.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.13 (at 3.57Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.315 , 0.343 0.303 , 0.324	Depositor DCC
$R_{free}$ test set	7393 reflections (11.17%)	DCC
Wilson B-factor (Å <sup>2</sup> )	47.6	Xtriage
Anisotropy	0.357	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , -10.0	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	0 of 80020 reflections	Xtriage
$F_o, F_c$ correlation	0.76	EDS
Total number of atoms	28647	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.72% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 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	R	0.67	0/244	0.83	0/380
2	T	0.75	0/311	1.39	3/477 (0.6%)
3	A	0.88	15/11163 (0.1%)	1.45	167/15091 (1.1%)
4	B	0.84	7/8964 (0.1%)	1.38	114/12086 (0.9%)
5	C	0.78	0/2133	1.24	13/2891 (0.4%)
6	E	0.90	2/1788 (0.1%)	1.40	14/2406 (0.6%)
7	F	0.83	0/691	1.28	7/933 (0.8%)
8	H	0.85	1/1086 (0.1%)	1.59	20/1470 (1.4%)
9	I	1.03	2/989 (0.2%)	1.64	23/1331 (1.7%)
10	J	0.78	0/541	1.44	7/727 (1.0%)
11	K	0.74	0/937	1.20	5/1265 (0.4%)
12	L	0.99	1/366 (0.3%)	1.78	12/485 (2.5%)
All	All	0.86	28/29213 (0.1%)	1.41	385/39542 (1.0%)

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	T	0	5
3	A	1	6
4	B	0	7
5	C	0	2
6	E	0	1
9	I	0	1
All	All	1	22

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	319	GLY	C-O	10.97	1.41	1.23
3	A	255	SER	CA-CB	8.59	1.65	1.52
3	A	320	ARG	CA-CB	8.38	1.72	1.53
3	A	320	ARG	CG-CD	7.83	1.71	1.51
4	B	595	ARG	CG-CD	7.49	1.70	1.51

The worst 5 of 385 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	321	PRO	N-CA-C	-18.56	63.84	112.10
3	A	322	VAL	N-CA-C	14.54	150.26	111.00
10	J	10	CYS	CA-CB-SG	12.01	135.61	114.00
3	A	315	LEU	CA-CB-CG	11.82	142.50	115.30
4	B	478	GLY	N-CA-C	-11.76	83.71	113.10

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	322	VAL	CA

5 of 22 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	T	10	DT	Sidechain
2	T	11	DC	Sidechain
2	T	13	DA	Sidechain
2	T	6	DC	Sidechain
2	T	8	DT	Sidechain

## 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	R	217	0	110	18	0
2	T	279	0	160	43	0
3	A	10969	0	11070	2106	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	8793	0	8823	1592	0
5	C	2095	0	2051	337	0
6	E	1752	0	1776	303	0
7	F	679	0	701	127	0
8	H	1068	0	1040	193	0
9	I	971	0	929	162	0
10	J	532	0	542	125	0
11	K	919	0	929	175	0
12	L	364	0	387	65	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	C	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
14	A	1	0	0	0	0
All	All	28647	0	28518	4856	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 85.

The worst 5 of 4856 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:A:853:ASP:OD1	3:A:855:THR:HB	1.32	1.25
3:A:90:VAL:HG12	3:A:297:GLN:NE2	1.49	1.24
4:B:635:ARG:HB2	4:B:636:PRO:CD	1.65	1.21
3:A:321:PRO:O	3:A:322:VAL:HG22	1.41	1.18
3:A:351:THR:HG23	4:B:1103:ILE:HD12	1.23	1.18

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	
3	A	1383/1733 (80%)	851 (62%)	315 (23%)	217 (16%)	0	6
4	B	1088/1224 (89%)	730 (67%)	214 (20%)	144 (13%)	0	10
5	C	264/318 (83%)	187 (71%)	49 (19%)	28 (11%)	1	15
6	E	212/215 (99%)	142 (67%)	41 (19%)	29 (14%)	0	9
7	F	82/155 (53%)	49 (60%)	26 (32%)	7 (8%)	1	24
8	H	129/146 (88%)	87 (67%)	16 (12%)	26 (20%)	0	3
9	I	117/122 (96%)	74 (63%)	24 (20%)	19 (16%)	0	6
10	J	63/70 (90%)	42 (67%)	11 (18%)	10 (16%)	0	6
11	K	112/120 (93%)	81 (72%)	20 (18%)	11 (10%)	1	18
12	L	44/70 (63%)	20 (46%)	10 (23%)	14 (32%)	0	0
All	All	3494/4173 (84%)	2263 (65%)	726 (21%)	505 (14%)	0	8

5 of 505 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	44	THR
3	A	50	ILE
3	A	55	ASP
3	A	56	PRO
3	A	57	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	
3	A	1218/1520 (80%)	859 (70%)	359 (30%)	0	4
4	B	960/1061 (90%)	718 (75%)	242 (25%)	1	7
5	C	234/274 (85%)	175 (75%)	59 (25%)	1	7
6	E	196/197 (100%)	141 (72%)	55 (28%)	0	5
7	F	74/137 (54%)	57 (77%)	17 (23%)	1	8
8	H	117/128 (91%)	80 (68%)	37 (32%)	0	3
9	I	113/116 (97%)	84 (74%)	29 (26%)	1	7

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
10	J	60/65 (92%)	42 (70%)	18 (30%)	0	4
11	K	99/102 (97%)	81 (82%)	18 (18%)	2	16
12	L	40/57 (70%)	27 (68%)	13 (32%)	0	3
All	All	3111/3657 (85%)	2264 (73%)	847 (27%)	0	5

5 of 847 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	177	LYS
4	B	635	ARG
9	I	31	THR
4	B	234	ILE
4	B	408	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 101 such sidechains are listed below:

Mol	Chain	Res	Type
4	B	46	GLN
4	B	515	HIS
9	I	12	ASN
4	B	121	ASN
4	B	366	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	9/10 (90%)	2 (22%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	5	A
1	R	10	A

There are no RNA pucker outliers to report.

## 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 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.

## 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	R	10/10 (100%)	1.26	2 (20%) 2 2	30, 82, 200, 200	0
2	T	14/14 (100%)	1.43	4 (28%) 1 1	30, 107, 200, 200	0
3	A	1395/1733 (80%)	0.07	18 (1%) 74 47	30, 32, 134, 200	0
4	B	1106/1224 (90%)	0.01	7 (0%) 86 66	30, 30, 118, 198	0
5	C	266/318 (83%)	-0.07	0 100 100	30, 30, 84, 140	0
6	E	214/215 (99%)	0.40	9 (4%) 35 19	30, 67, 153, 200	0
7	F	84/155 (54%)	0.07	0 100 100	30, 31, 93, 141	0
8	H	133/146 (91%)	0.12	2 (1%) 70 43	30, 47, 138, 190	0
9	I	119/122 (97%)	0.09	0 100 100	30, 32, 114, 163	0
10	J	65/70 (92%)	-0.10	0 100 100	30, 30, 96, 141	0
11	K	114/120 (95%)	-0.09	0 100 100	30, 30, 73, 107	0
12	L	46/70 (65%)	0.03	0 100 100	30, 45, 132, 158	0
All	All	3566/4197 (84%)	0.06	42 (1%) 75 49	30, 31, 130, 200	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	T	1	DA	5.9
4	B	866	TYR	4.9
3	A	161	LEU	4.4
6	E	83	CYS	4.4
3	A	316	GLN	3.7

### 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( $\text{\AA}^2$ )	Q<0.9
13	ZN	I	203	1/1	0.22	-0.38	22,22,22,22	0
13	ZN	A	1734	1/1	0.26	-0.71	22,22,22,22	0
13	ZN	I	204	1/1	0.15	-0.91	22,22,22,22	0
14	MG	A	2000	1/1	0.20	-0.94	22,22,22,22	0
13	ZN	B	1307	1/1	0.11	-1.84	22,22,22,22	0
13	ZN	L	105	1/1	0.06	-2.01	22,22,22,22	0
13	ZN	A	1735	1/1	0.10	-2.08	22,22,22,22	0
13	ZN	C	319	1/1	0.07	-2.34	22,22,22,22	0
13	ZN	J	101	1/1	0.11	-4.84	22,22,22,22	0

### 6.5 Other polymers ⓘ

There are no such residues in this entry.