



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 01:17 am BST

PDB ID : 4YFK
Title : Escherichia coli RNA polymerase in complex with squaramide compound 8.
Authors : Molodtsov, V.; Fleming, P.R.; Eyermann, C.J.; Ferguson, A.D.; Foulk, M.A.;
McKinney, D.C.; Masse, C.E.; Buurman, E.T.; Murakami, K.S.
Deposited on : 2015-02-25
Resolution : 3.57 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
buster-report : 1.1.7 (2018)
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

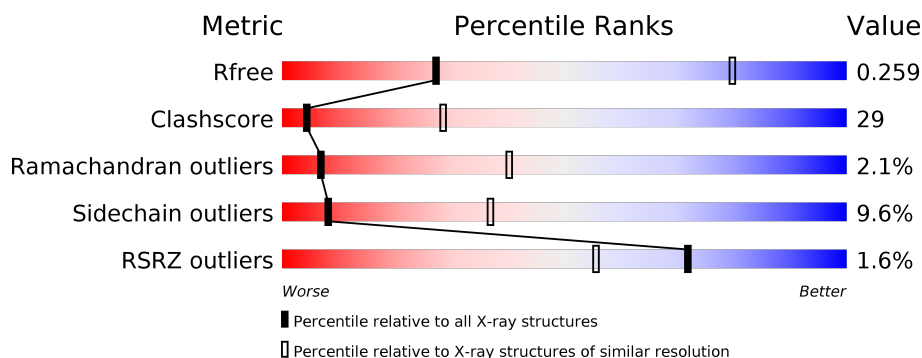
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.57 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	1094 (3.66-3.50)
Clashscore	141614	1181 (3.66-3.50)
Ramachandran outliers	138981	1143 (3.66-3.50)
Sidechain outliers	138945	1143 (3.66-3.50)
RSRZ outliers	127900	1012 (3.66-3.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 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	329	<div> <div>2%</div> <div> <div>37%</div> <div>45%</div> <div>14%</div> <div>••</div> </div> </div>
1	B	329	<div> <div>0%</div> <div> <div>30%</div> <div>33%</div> <div>•</div> <div>34%</div> </div> </div>
1	G	329	<div> <div>33%</div> <div>30%</div> <div>6%</div> <div>31%</div> </div>
1	H	329	<div> <div>2%</div> <div> <div>28%</div> <div>34%</div> <div>•</div> <div>34%</div> </div> </div>
2	C	1342	<div> <div>0%</div> <div> <div>48%</div> <div>46%</div> <div>6%</div> </div> </div>
2	I	1342	<div> <div>3%</div> <div> <div>49%</div> <div>45%</div> <div>5%</div> <div>•</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	1407	<div><div></div><div>37%40%6%17%</div></div>
3	J	1407	<div><div>%</div><div>39%37%5%18%</div></div>
4	E	91	<div><div></div><div>60%35%. .</div></div>
4	K	91	<div><div>10%</div><div>40%43%.13%</div></div>
5	F	613	<div><div>2%</div><div>38%33%5%24%</div></div>
5	L	613	<div><div>2%</div><div>33%39%5%23%</div></div>

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 55782 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	319	Total	C	N	O	S	0	0	0
			2490	1557	439	486	8			
1	B	217	Total	C	N	O	S	0	0	0
			1677	1047	295	329	6			
1	G	227	Total	C	N	O	S	0	0	0
			1755	1093	311	345	6			
1	H	216	Total	C	N	O	S	0	0	0
			1662	1038	292	326	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1340	Total	C	N	O	S	0	0	0
			10570	6631	1841	2055	43			
2	I	1340	Total	C	N	O	S	0	0	0
			10566	6629	1840	2054	43			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1163	Total	C	N	O	S	0	0	0
			9050	5690	1620	1694	46			
3	J	1152	Total	C	N	O	S	0	0	0
			8990	5654	1608	1682	46			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	89	Total	C	N	O	S	0	0	0
			691	421	129	140	1			
4	K	79	Total	C	N	O	S	0	0	0
			627	382	118	126	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	468	Total	C	N	O	S	0	0	0
			3813	2389	678	723	23			
5	L	469	Total	C	N	O	S	0	0	0
			3821	2393	679	726	23			

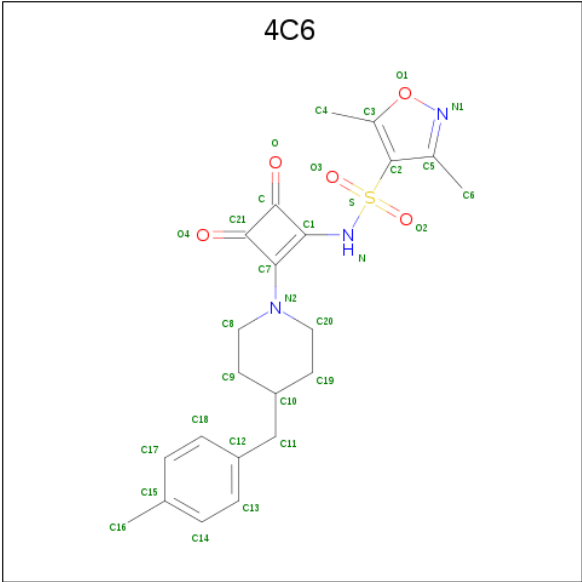
- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	1	Total	Mg	0	0
			1	1		
6	I	1	Total	Mg	0	0
			1	1		
6	D	1	Total	Mg	0	0
			1	1		
6	C	1	Total	Mg	0	0
			1	1		

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	J	2	Total	Zn	0	0
			2	2		
7	D	2	Total	Zn	0	0
			2	2		

- Molecule 8 is 3,5-dimethyl-N-{2-[4-(4-methylbenzyl)piperidin-1-yl]-3,4-dioxocyclobut-1-en-1-yl}-1,2-oxazole-4-sulfonamide (three-letter code: 4C6) (formula: C₂₂H₂₅N₃O₅S).

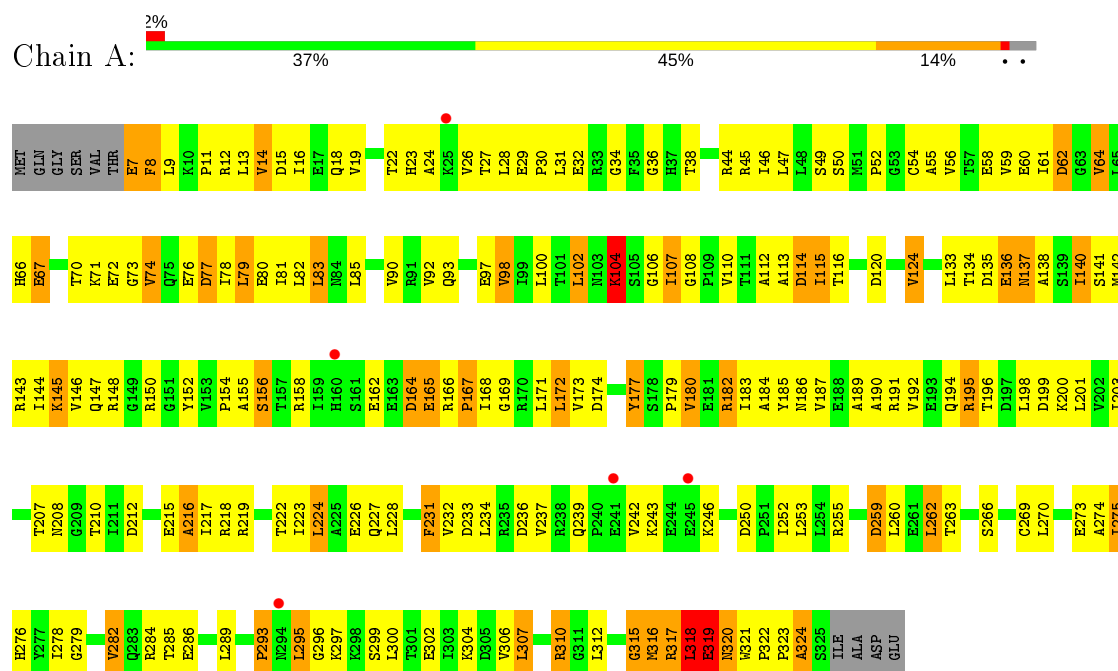


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	D	1	Total	C	N	O	S	0	0
			31	22	3	5	1		
8	J	1	Total	C	N	O	S	0	0
			31	22	3	5	1		

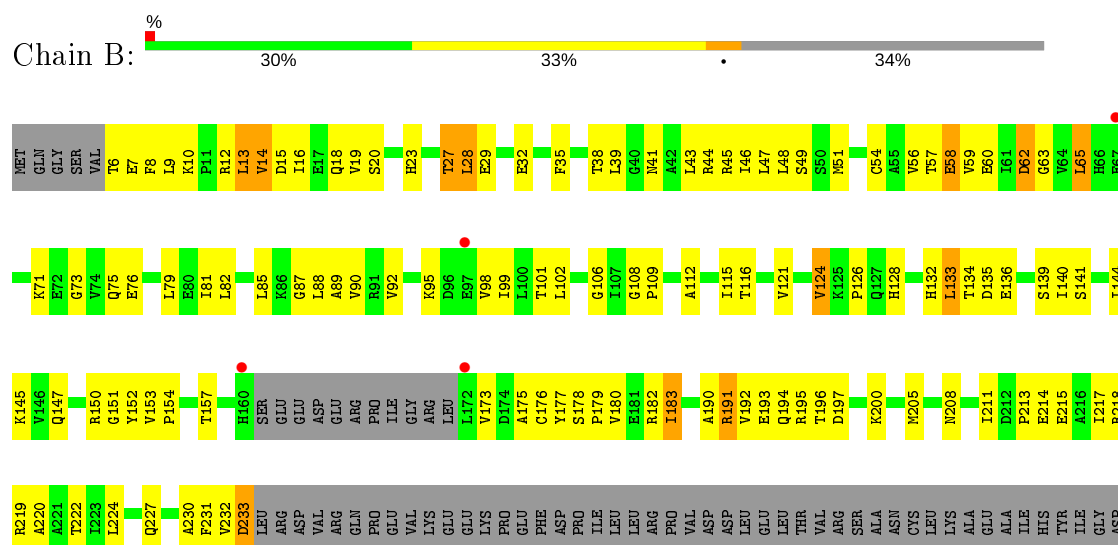
3 Residue-property plots [i](#)

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

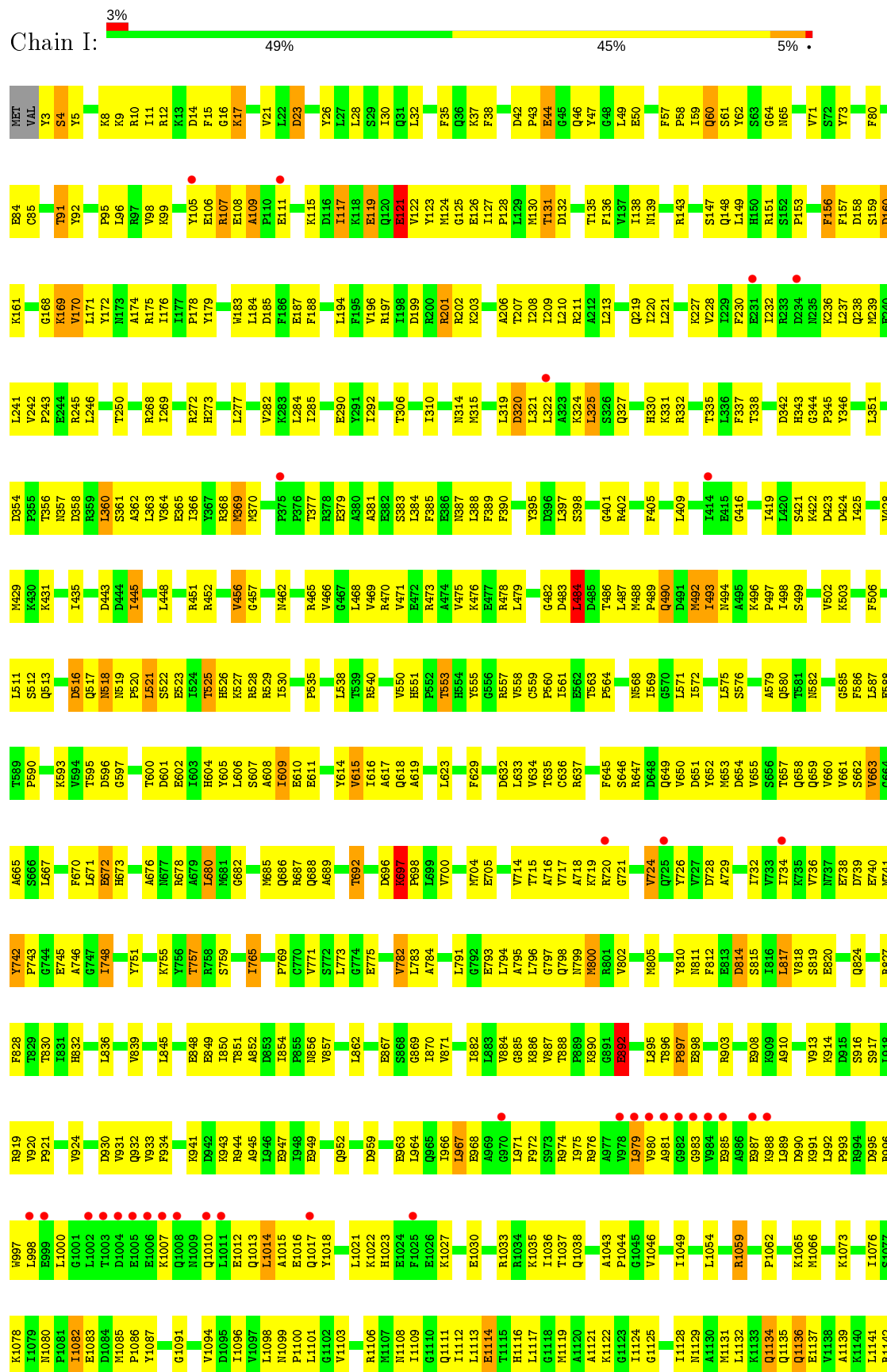


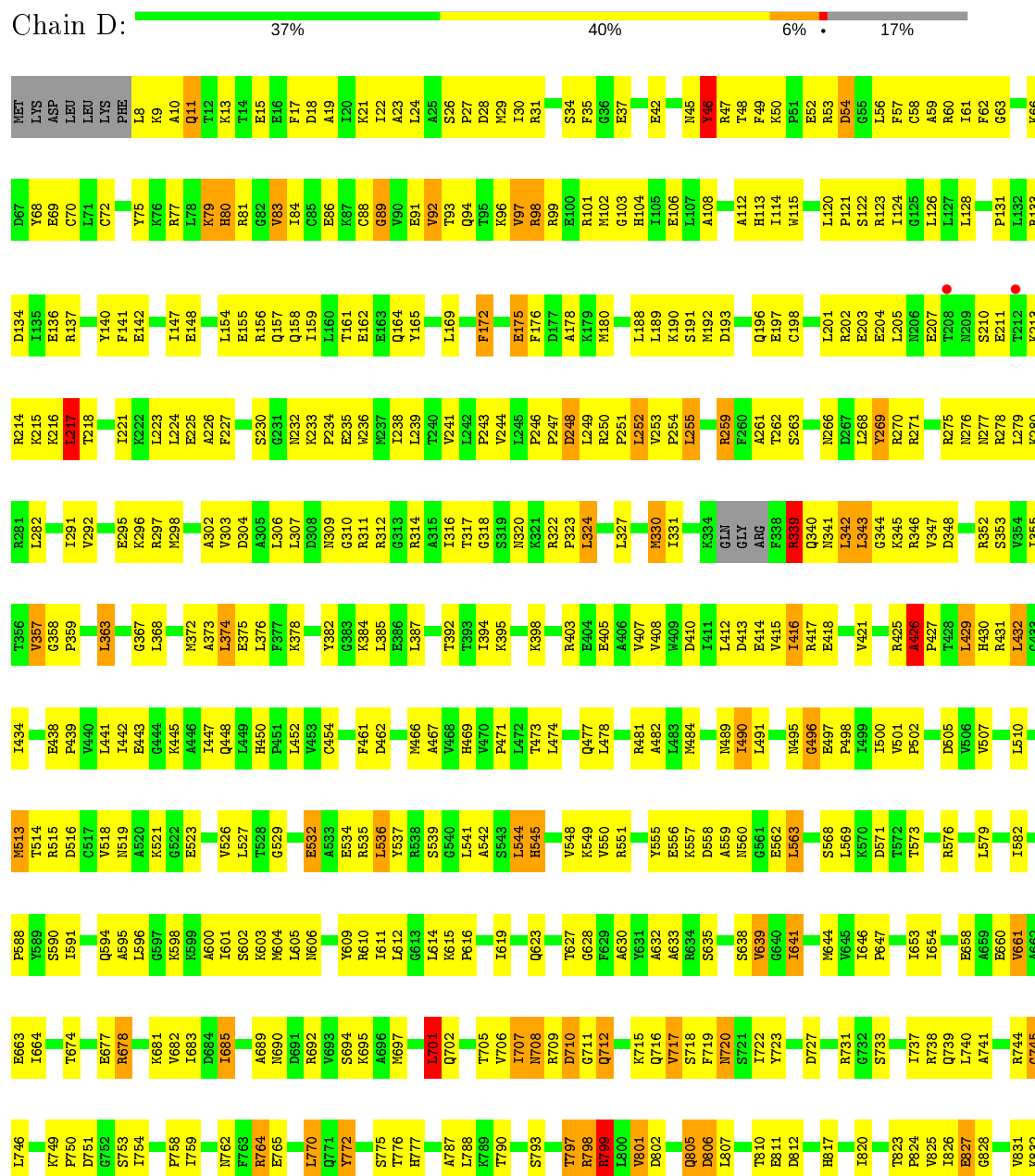
• Molecule 1: DNA-directed RNA polymerase subunit alpha

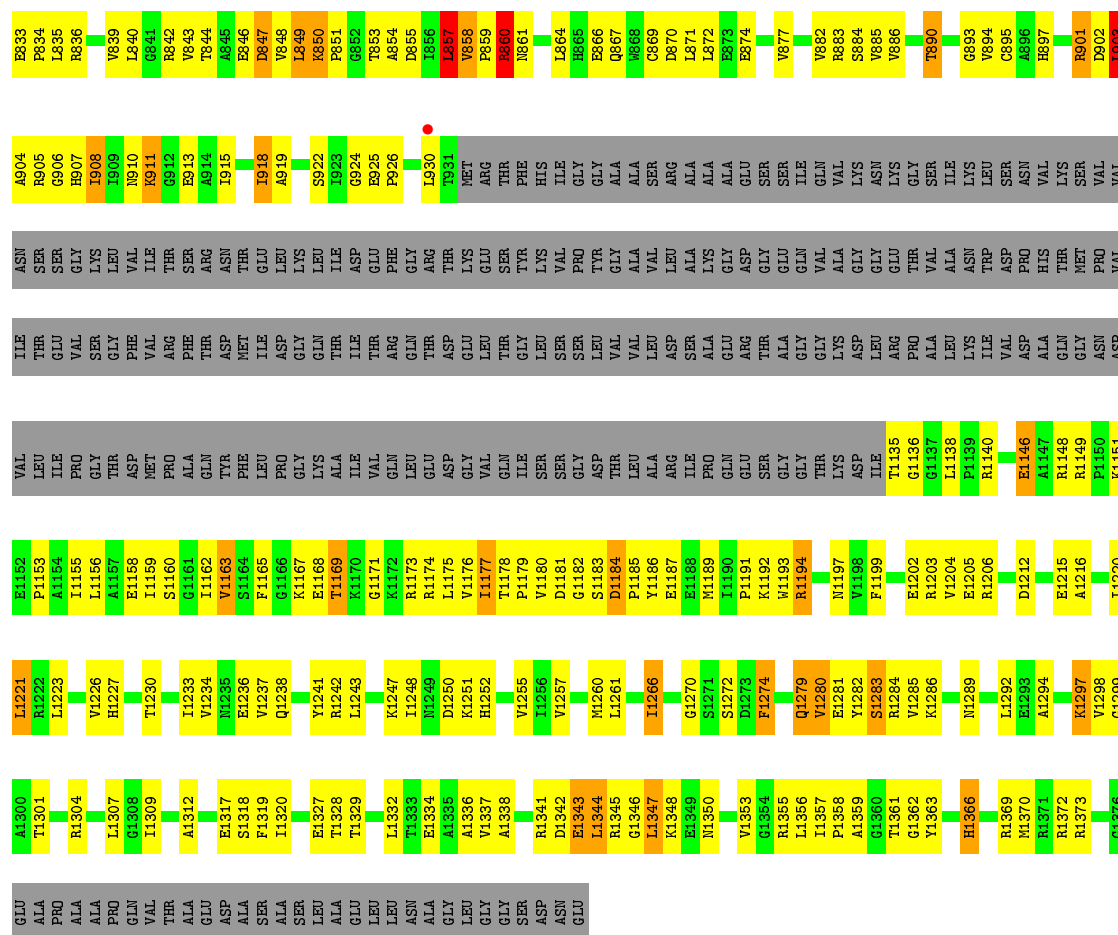


Q1288	L1212	R1142	G1071	E985	Q884	N811	D728	V650	G570	M488	A380	T306	E218	F136
E1289	R1216	Q1146	M1072	A986	L985	D814	A729	D651	L571	P489	A381	T309	Q219	V137
M1290	T1217	Q987	K1073	K987	T986	S730	S730	E382	I572	Q490	E382	I310	I138	I138
L1291	G1218	D989	P1081	K988	E988	L817	I732	M653	L575	I493	L384	C311	L221	N139
T1292	E1219	Y1149	S1077	L990	E989	V818	V736	V655	N582	M494	Y395	N314	F224	G140
K1294	G1222	L1151	K1078	K991	K900	S819	T737	S656	F586	P497	K404	F316	F225	R143
S1295	R1223	A1152	I1079	L992	L901	Q824	E738	Q658	L587	I498	S403	S318	I229	V146
D1296	D1297	D1154	P1081	R993	L902	D826	D739	V660	F590	F506	L419	D320	E231	S147
V1298	G1228	V1155	I1082	R994	R903	F828	E740	Q859	T595	I591	L420	L322	E231	Q148
N1299	Y1229	R1156	E1083	Y997	F906	T829	V742	V663	H592	F506	L419	D320	K236	L149
G1300	M1230	Q1157	D1084	L998	A910	H832	V742	V663	H592	F506	L419	D320	K236	H150
R1301	Y1231	K1158	M1085	L998	A910	H832	V742	V663	H592	F506	L419	D320	K236	R151
K1303	M1232	V1159	P1086	L1002	V913	H832	V742	V663	H592	F506	L419	D320	K236	F157
N1307	K1234	L1161	M1090	L1002	V913	H832	V742	V663	H592	F506	L419	D320	K236	D158
I1308	L1235	S1162	M1090	L1002	V913	H832	V742	V663	H592	F506	L419	D320	K236	S159
G1311	M1236	T1163	V1094	E1006	S916	A837	G747	V670	V699	Q517	V428	H330	L246	D160
N1312	H1237	F1164	D1095	K1007	S917	C838	I748	L671	T600	N518	V428	H330	L246	K161
H1313	S1238	S1165	I1096	Q1010	L918	R839	I748	E672	T600	N518	V428	H330	L246	H165
Q1314	D1240	D1166	V1097	Q1010	V920	S840	I748	H673	D601	N519	M429	K430	T250	G168
P1317	M1243	V1169	L1098	Q1017	V924	I850	I756	D675	H604	S522	K431	K332	A251	K169
G1318	R1246	R1171	P1100	L1021	D930	A852	I756	L693	V605	I524	K431	K332	A251	V170
M1319	S1247	N1175	L1101	L1021	V931	H851	I756	L693	V605	I524	K431	K332	A251	L171
P1320	T1248	G1179	P1104	K1028	D942	L865	E778	L623	N622	T539	G457	K353	R272	D185
S1322	S1250	M1180	S1105	K1028	K943	D866	E778	L623	N622	T539	G457	K353	R272	F188
L1326	Y1251	P1181	M1107	K1032	R944	E867	I783	L633	D624	E541	M459	D354	I274	D189
L1327	S1252	I1182	M1108	K1033	A945	G869	I783	L633	D624	E541	M459	D354	I274	P190
R1331	V1254	A1183	I1109	R1034	L946	S888	I783	L633	D624	E541	M459	D354	I274	N193
S1332	T1255	V1186	Q1111	I1036	E949	L865	I783	L633	D624	E541	M459	D354	I274	L194
L1333	Q1256	F1187	I1112	T1037	E950	D866	I783	L633	D624	E541	M459	D354	I274	F195
P1258	L1259	G1189	E1114	Q1038	N951	E867	I783	L633	D624	E541	M459	D354	I274	V196
G1260	Q1264	K1191	H1116	P1044	L953	G869	I783	L633	D624	E541	M459	D354	I274	R197
F1265	F1265	A1193	G1118	V1046	L960	I870	I783	L633	D624	E541	M459	D354	I274	I198
G1266	G1266	E1194	M1119	L1047	E963	V871	I783	L633	D624	E541	M459	D354	I274	R201
Q1267	Q1267	K1196	G1123	K1048	L967	G874	I783	L633	D624	E541	M459	D354	I274	R202
R1269	R1269	L1199	I1124	K1051	L971	G874	I783	L633	D624	E541	M459	D354	I274	A206
W1276	W1276	G1202	G1125	V1052	S973	L883	I783	L633	D624	E541	M459	D354	I274	T207
E1279	E1279	D1203	I1128	R1058	R974	V884	I783	L633	D624	E541	M459	D354	I274	I208
Y1281	Y1281	L1204	M1131	R1059	L975	G885	I783	L633	D624	E541	M459	D354	I274	I209
G1282	G1282	P1205	Q1134	Q1061	R976	K886	I783	L633	D624	E541	M459	D354	I274	L210
A1283	A1283	T1206	Q1135	P1062	L979	V888	I783	L633	D624	E541	M459	D354	I274	R211
A1284	A1284	S1207	Q1136	G1063	V980	T888	I783	L633	D624	E541	M459	D354	I274	Y215
Y1285	Y1285	G1286	Q1136	D1064	A981	K890	I783	L633	D624	E541	M459	D354	I274	
L1286	L1286	Q1209	E1137	K1065	G982	K891	I783	L633	D624	E541	M459	D354	I274	
L1287	L1287	I1210	L1141	H1070	V984	E892	I783	L633	D624	E541	M459	D354	I274	

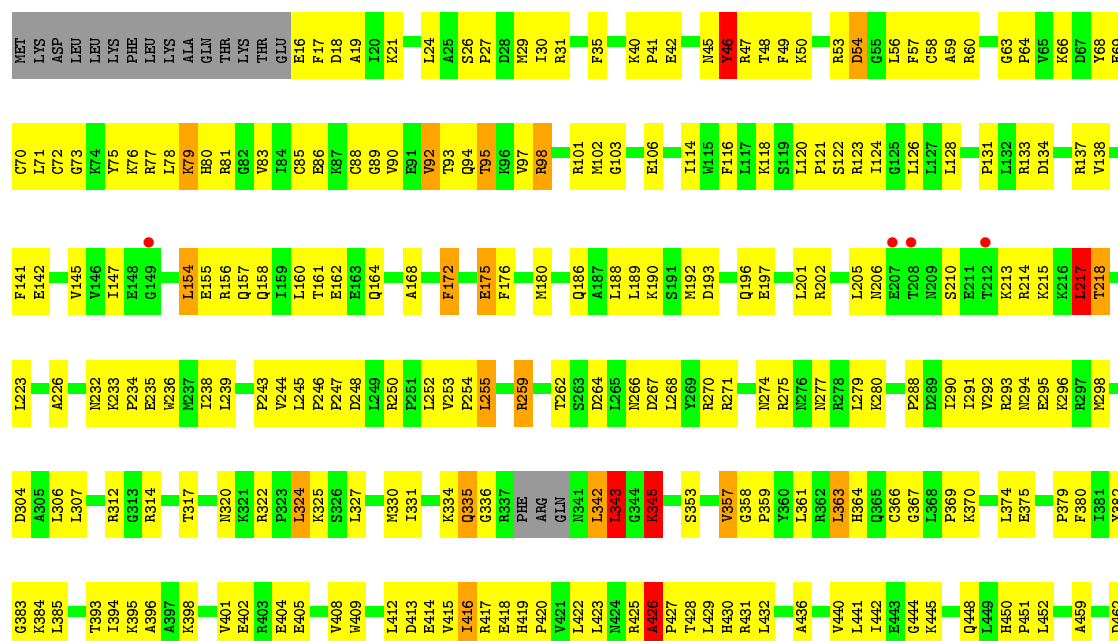
● Molecule 2: DNA-directed RNA polymerase subunit beta

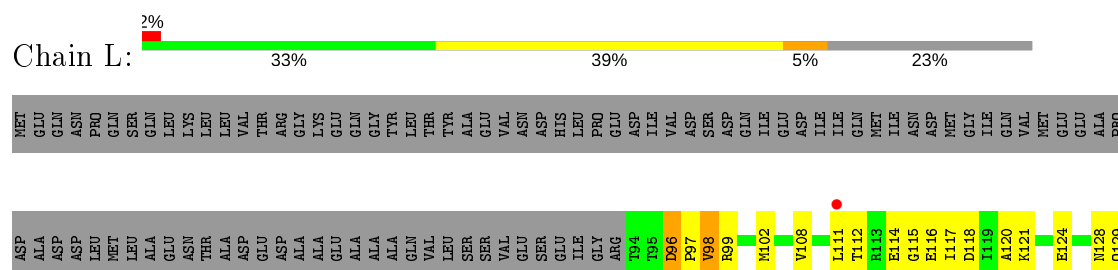
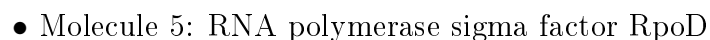
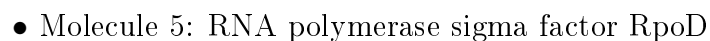






• Molecule 3: DNA-directed RNA polymerase subunit beta'







4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	186.10Å 206.44Å 307.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.96 – 3.57 31.02 – 3.57	Depositor EDS
% Data completeness (in resolution range)	94.6 (29.96-3.57) 92.3 (31.02-3.57)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.12 (at 3.56Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.217 , 0.252 0.227 , 0.259	Depositor DCC
R_{free} test set	2000 reflections (1.50%)	wwPDB-VP
Wilson B-factor (Å ²)	112.0	Xtriage
Anisotropy	0.227	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 57.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	55782	wwPDB-VP
Average B, all atoms (Å ²)	147.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.86% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 4C6, 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	0.80	0/2524	0.91	2/3421 (0.1%)
1	B	0.86	3/1697 (0.2%)	0.97	4/2300 (0.2%)
1	G	0.80	0/1777	0.92	1/2408 (0.0%)
1	H	0.83	1/1681 (0.1%)	0.94	1/2278 (0.0%)
2	C	0.84	1/10739 (0.0%)	0.89	9/14489 (0.1%)
2	I	0.79	2/10735 (0.0%)	0.87	10/14484 (0.1%)
3	D	0.90	6/9188 (0.1%)	0.95	18/12404 (0.1%)
3	J	0.78	2/9128 (0.0%)	0.89	10/12322 (0.1%)
4	E	0.73	0/693	0.75	0/935
4	K	1.03	0/629	0.89	0/847
5	F	0.83	1/3864 (0.0%)	0.87	3/5194 (0.1%)
5	L	0.83	1/3872 (0.0%)	0.83	2/5205 (0.0%)
All	All	0.83	17/56527 (0.0%)	0.90	60/76287 (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	C	0	3
2	I	0	4
3	D	0	2
3	J	0	1
5	F	0	1
All	All	0	11

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	799	ARG	CB-CG	-10.16	1.25	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	339	ARG	CZ-NH2	8.46	1.44	1.33
3	D	799	ARG	CG-CD	8.42	1.73	1.51
2	I	1296	ASP	CG-OD2	8.26	1.44	1.25
3	D	339	ARG	CB-CG	-8.08	1.30	1.52

The worst 5 of 60 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	1269	ARG	NE-CZ-NH1	15.48	128.04	120.30
2	I	1269	ARG	NE-CZ-NH2	-11.03	114.79	120.30
2	I	484	LEU	CA-CB-CG	9.30	136.69	115.30
2	I	1269	ARG	CD-NE-CZ	9.27	136.57	123.60
2	C	1161	LEU	CA-CB-CG	-9.26	94.01	115.30

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	109	ALA	Peptide
2	C	1264	GLN	Peptide
2	C	236	LYS	Peptide
3	D	1184	ASP	Peptide
3	D	901	ARG	Peptide

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	2490	0	2542	204	0
1	B	1677	0	1703	105	0
1	G	1755	0	1773	122	0
1	H	1662	0	1687	131	0
2	C	10570	0	10582	606	0
2	I	10566	0	10576	626	0
3	D	9050	0	9218	641	0
3	J	8990	0	9173	595	0
4	E	691	0	695	24	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	K	627	0	634	47	0
5	F	3813	0	3880	203	0
5	L	3821	0	3884	244	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	I	1	0	0	0	0
6	J	1	0	0	0	0
7	D	2	0	0	0	0
7	J	2	0	0	0	0
8	D	31	0	25	1	0
8	J	31	0	25	2	0
All	All	55782	0	56397	3255	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 29.

The worst 5 of 3255 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:G:190:ALA:HB2	1:G:200:LYS:HB2	1.29	1.14
1:A:45:ARG:HG2	1:B:38:THR:HB	1.31	1.12
3:D:660:GLU:HB3	3:D:685:ILE:HD12	1.30	1.08
2:I:1269:ARG:HD3	3:J:343:LEU:HD21	1.35	1.08
3:J:1280:VAL:HG21	3:J:1304:ARG:HE	1.17	1.07

There are no symmetry-related clashes.

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	317/329 (96%)	243 (77%)	48 (15%)	26 (8%)	1 10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	213/329 (65%)	193 (91%)	15 (7%)	5 (2%)	6	38
1	G	225/329 (68%)	195 (87%)	21 (9%)	9 (4%)	3	26
1	H	212/329 (64%)	193 (91%)	15 (7%)	4 (2%)	8	42
2	C	1338/1342 (100%)	1210 (90%)	111 (8%)	17 (1%)	12	49
2	I	1338/1342 (100%)	1207 (90%)	112 (8%)	19 (1%)	11	48
3	D	1157/1407 (82%)	1031 (89%)	101 (9%)	25 (2%)	6	39
3	J	1146/1407 (81%)	1032 (90%)	92 (8%)	22 (2%)	8	42
4	E	87/91 (96%)	81 (93%)	4 (5%)	2 (2%)	6	38
4	K	77/91 (85%)	73 (95%)	3 (4%)	1 (1%)	12	49
5	F	462/613 (75%)	424 (92%)	30 (6%)	8 (2%)	9	44
5	L	463/613 (76%)	425 (92%)	30 (6%)	8 (2%)	9	44
All	All	7035/8222 (86%)	6307 (90%)	582 (8%)	146 (2%)	7	40

5 of 146 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	62	ASP
1	A	107	ILE
1	A	114	ASP
1	A	136	GLU
1	A	195	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	278/286 (97%)	231 (83%)	47 (17%)	2	13
1	B	186/286 (65%)	172 (92%)	14 (8%)	13	45
1	G	193/286 (68%)	170 (88%)	23 (12%)	5	27
1	H	183/286 (64%)	170 (93%)	13 (7%)	14	48
2	C	1155/1157 (100%)	1048 (91%)	107 (9%)	9	38

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	I	1154/1157 (100%)	1055 (91%)	99 (9%)	10	41
3	D	964/1168 (82%)	867 (90%)	97 (10%)	7	35
3	J	962/1168 (82%)	869 (90%)	93 (10%)	8	36
4	E	72/75 (96%)	67 (93%)	5 (7%)	15	49
4	K	67/75 (89%)	63 (94%)	4 (6%)	19	54
5	F	417/540 (77%)	376 (90%)	41 (10%)	8	36
5	L	418/540 (77%)	378 (90%)	40 (10%)	8	37
All	All	6049/7024 (86%)	5466 (90%)	583 (10%)	8	37

5 of 583 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	1361	THR
1	G	168	ILE
3	J	1366	HIS
5	F	98	VAL
5	F	529	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 66 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	132	HIS
2	I	1038	GLN
4	K	7	GLN
2	I	139	ASN
2	I	658	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 8 are 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
8	4C6	D	2004	-	32,34,34	1.23	3 (9%)	41,51,51	0.97	1 (2%)
8	4C6	J	2004	-	32,34,34	0.83	1 (3%)	41,51,51	1.32	6 (14%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	4C6	D	2004	-	-	0/7/45/45	0/4/4/4
8	4C6	J	2004	-	-	0/7/45/45	0/4/4/4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	D	2004	4C6	C2-S	-5.01	1.71	1.79
8	D	2004	4C6	C4-C3	-2.16	1.46	1.48
8	D	2004	4C6	S-N	-2.05	1.60	1.64
8	J	2004	4C6	C2-S	-2.02	1.75	1.79

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	J	2004	4C6	C4-C3-C2	4.75	134.41	129.47
8	D	2004	4C6	C4-C3-C2	4.24	133.87	129.47
8	J	2004	4C6	C7-C21-C	-2.81	85.26	88.04

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	J	2004	4C6	O3-S-C2	2.66	112.70	108.74
8	J	2004	4C6	C7-C1-N	2.40	135.43	129.93

There are no chirality outliers.

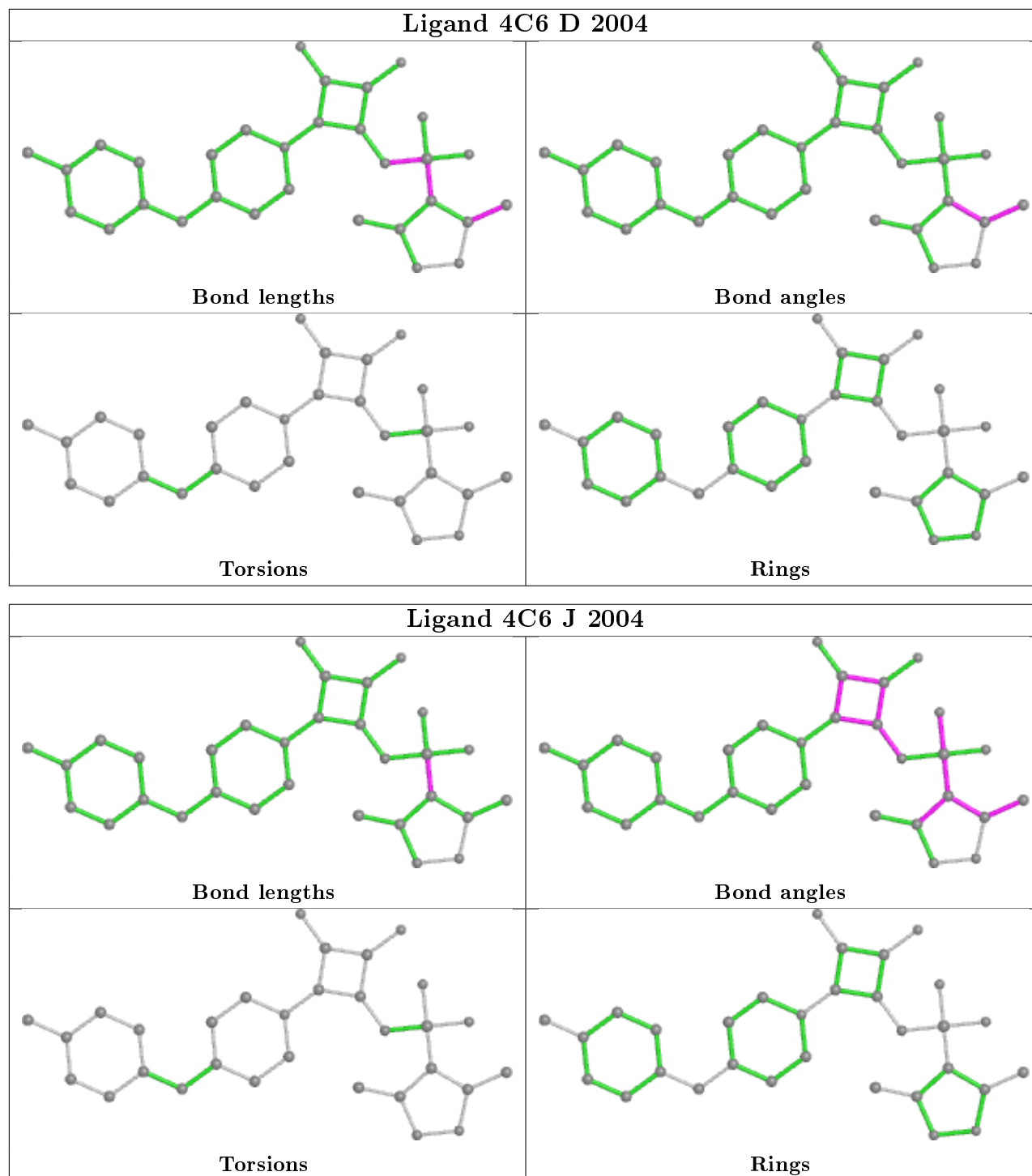
There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	D	2004	4C6	1	0
8	J	2004	4C6	2	0

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 than 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.



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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	319/329 (96%)	-0.33	5 (1%) 72 55	105, 138, 176, 185	0
1	B	217/329 (65%)	-0.16	4 (1%) 68 51	114, 172, 192, 198	0
1	G	227/329 (68%)	-0.38	0 100 100	138, 159, 175, 192	0
1	H	216/329 (65%)	-0.06	6 (2%) 53 35	130, 174, 192, 201	0
2	C	1340/1342 (99%)	-0.37	14 (1%) 82 69	88, 126, 202, 228	0
2	I	1340/1342 (99%)	-0.19	34 (2%) 57 39	108, 154, 212, 314	0
3	D	1163/1407 (82%)	-0.36	3 (0%) 94 88	90, 119, 164, 197	0
3	J	1152/1407 (81%)	-0.26	12 (1%) 82 69	103, 137, 181, 211	0
4	E	89/91 (97%)	-0.23	0 100 100	129, 159, 178, 184	0
4	K	79/91 (86%)	0.63	9 (11%) 5 3	186, 221, 249, 254	0
5	F	468/613 (76%)	-0.24	15 (3%) 47 31	113, 159, 233, 249	0
5	L	469/613 (76%)	-0.21	11 (2%) 60 42	127, 168, 244, 260	0
All	All	7079/8222 (86%)	-0.27	113 (1%) 72 55	88, 143, 205, 314	0

The worst 5 of 113 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	I	982	GLY	9.3
2	I	1003	THR	7.0
5	F	167	ASP	5.6
2	I	1002	LEU	5.1
2	I	979	LEU	4.9

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 ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

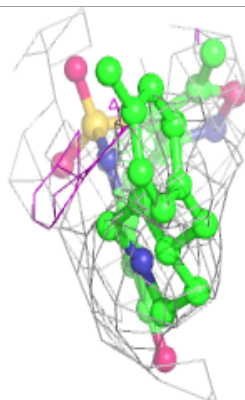
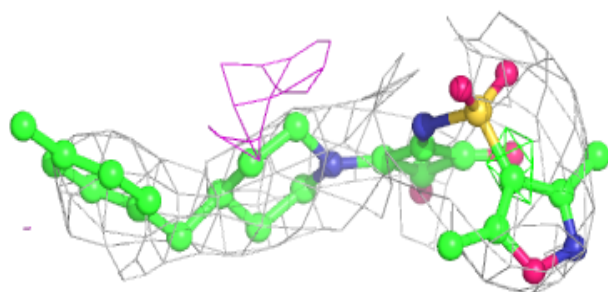
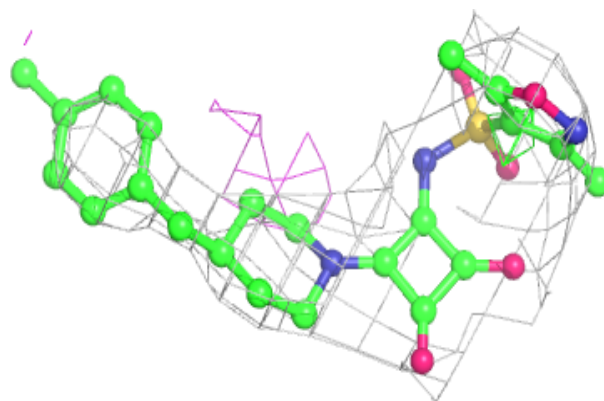
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
6	MG	I	1401	1/1	0.85	0.37	127,127,127,127	0
6	MG	C	1401	1/1	0.86	0.35	127,127,127,127	0
6	MG	D	2001	1/1	0.93	0.27	127,127,127,127	0
6	MG	J	2001	1/1	0.95	0.23	127,127,127,127	0
8	4C6	J	2004	31/31	0.95	0.32	127,127,128,141	0
7	ZN	J	2003	1/1	0.96	0.20	127,127,127,127	0
7	ZN	J	2002	1/1	0.96	0.12	190,190,190,190	0
8	4C6	D	2004	31/31	0.97	0.34	127,127,127,127	0
7	ZN	D	2002	1/1	0.98	0.18	164,164,164,164	0
7	ZN	D	2003	1/1	0.99	0.22	127,127,127,127	0

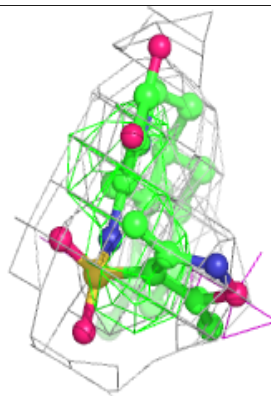
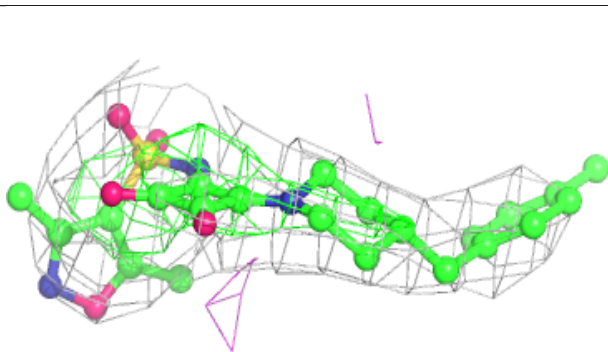
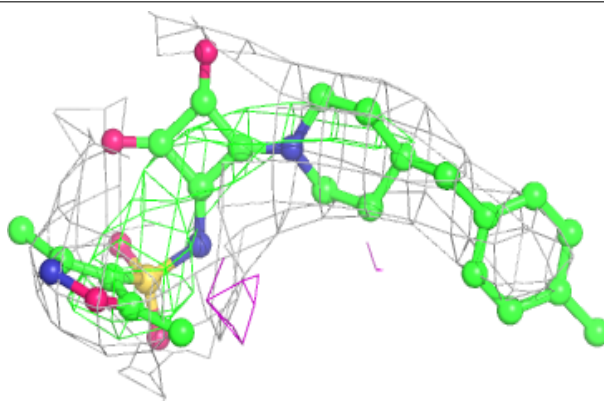
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around 4C6 J 2004:

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

**Electron density around 4C6 D 2004:**

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



6.5 Other polymers [i](#)

There are no such residues in this entry.