



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 27, 2023 – 01:42 AM EDT

PDB ID : 6BYU  
Title : X-ray crystal structure of Escherichia coli RNA polymerase (RpoB-H526Y) and ppApp complex  
Authors : Murakami, K.S.; Molodtsov, V.  
Deposited on : 2017-12-21  
Resolution : 3.60 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35.1  
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) : 2.35.1

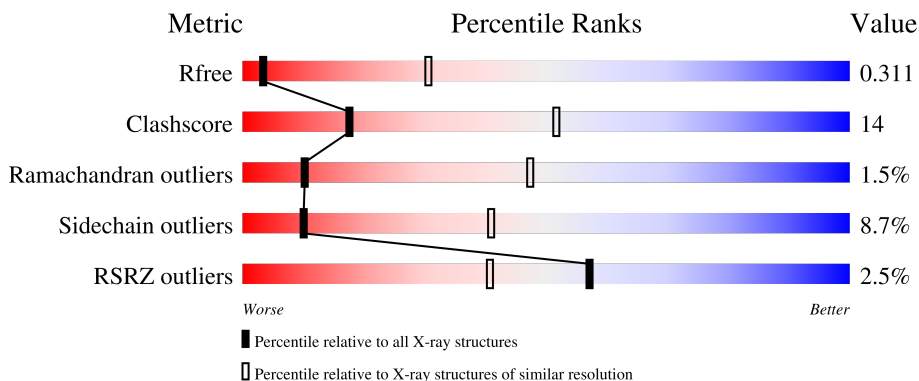
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.60 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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	
1	B	329	
1	G	329	
1	H	329	
2	C	1342	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	I	1342	<p>3% 64% 32% .</p>
3	D	1407	<p>% 52% 26% 5% 17%</p>
3	J	1407	<p>% 52% 26% . 18%</p>
4	E	91	<p>% 66% 30% ..</p>
4	K	91	<p>16% 47% 38% . 13%</p>
5	F	613	<p>4% 51% 23% . 24%</p>
5	L	613	<p>3% 57% 17% . 23%</p>

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 54996 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
			Total	C	N	O	S			
1	A	228	Total 1768	C 1102	N 312	O 348	S 6	0	0	0
1	B	217	Total 1672	C 1044	N 295	O 327	S 6	0	0	0
1	G	224	Total 1730	C 1076	N 308	O 340	S 6	0	0	0
1	H	217	Total 1667	C 1041	N 293	O 327	S 6	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	1340	Total 10561	C 6626	N 1837	O 2055	S 43	0	0	0
2	I	1340	Total 10557	C 6624	N 1836	O 2054	S 43	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	526	TYR	HIS	engineered mutation	UNP P0A8V2
I	526	TYR	HIS	engineered mutation	UNP P0A8V2

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	1163	Total 9063	C 5696	N 1622	O 1699	S 46	0	0	0
3	J	1155	Total 8998	C 5656	N 1612	O 1684	S 46	0	0	0

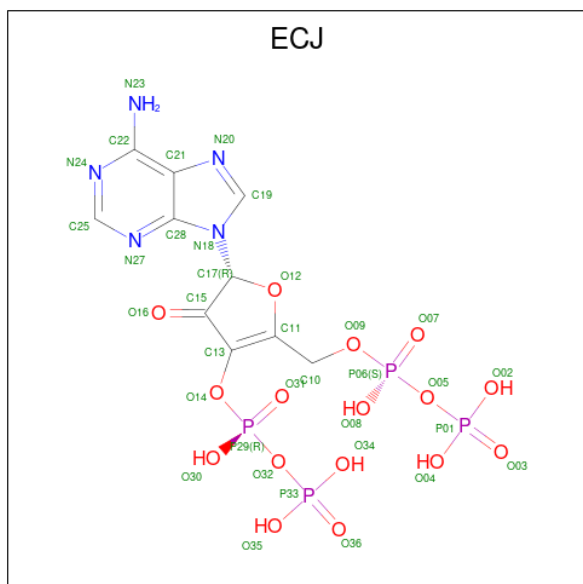
- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	89	Total	C	N	O	S	0	0	0
			685	418	126	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	467	Total	C	N	O	S	0	0	0
			3796	2379	677	717	23			
5	L	469	Total	C	N	O	S	0	0	0
			3796	2379	677	717	23			

- Molecule 6 is (5R)-5-(6-amino-9H-purin-9-yl)-2-({[(S)-hydroxy(phosphonooxy)phosphoryl]oxy}methyl)-4-oxo-4,5-dihydrofuran-3-yl trihydrogen diphosphate (three-letter code: ECJ) (formula: C<sub>10</sub>H<sub>13</sub>N<sub>5</sub>O<sub>16</sub>P<sub>4</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	C	1	Total	C	N	O	P	0	0
			35	10	5	16	4		
6	I	1	Total	C	N	O	P	0	0
			35	10	5	16	4		

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	1	Total Mg 1 1	0	0
7	J	1	Total Mg 1 1	0	0

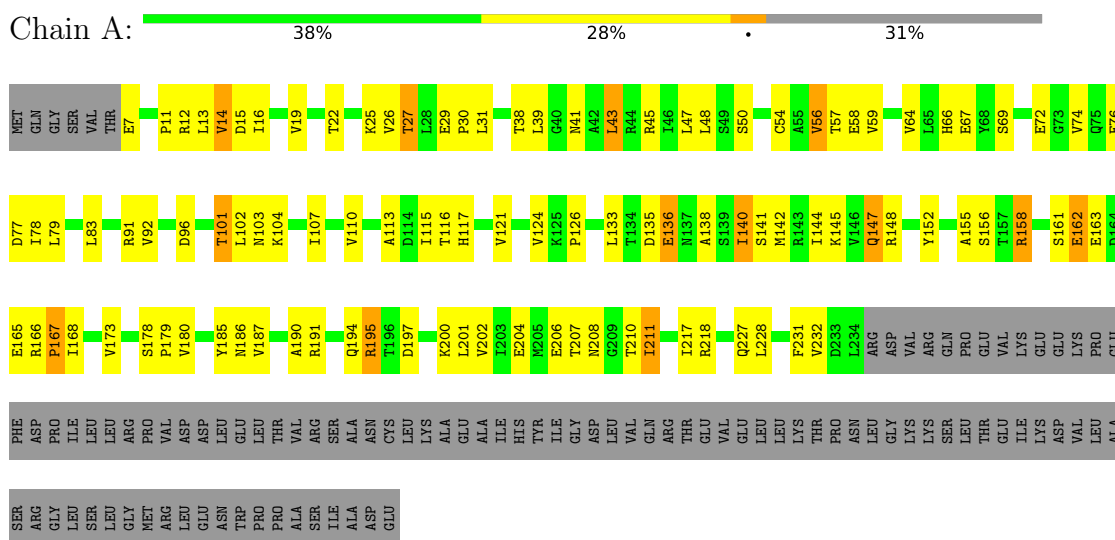
- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

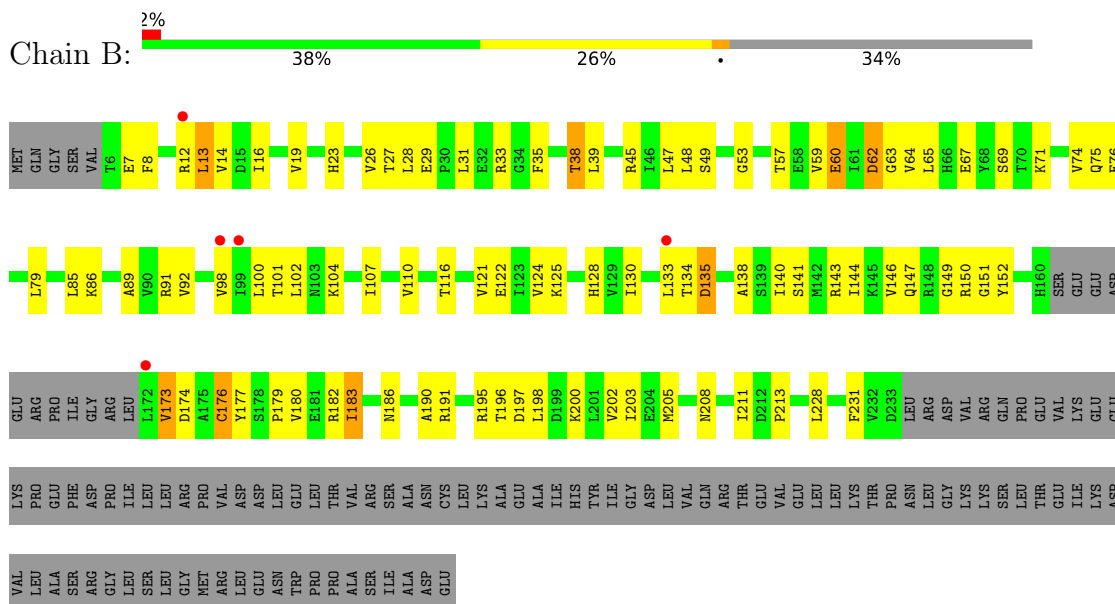
### 3 Residue-property plots [i](#)

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



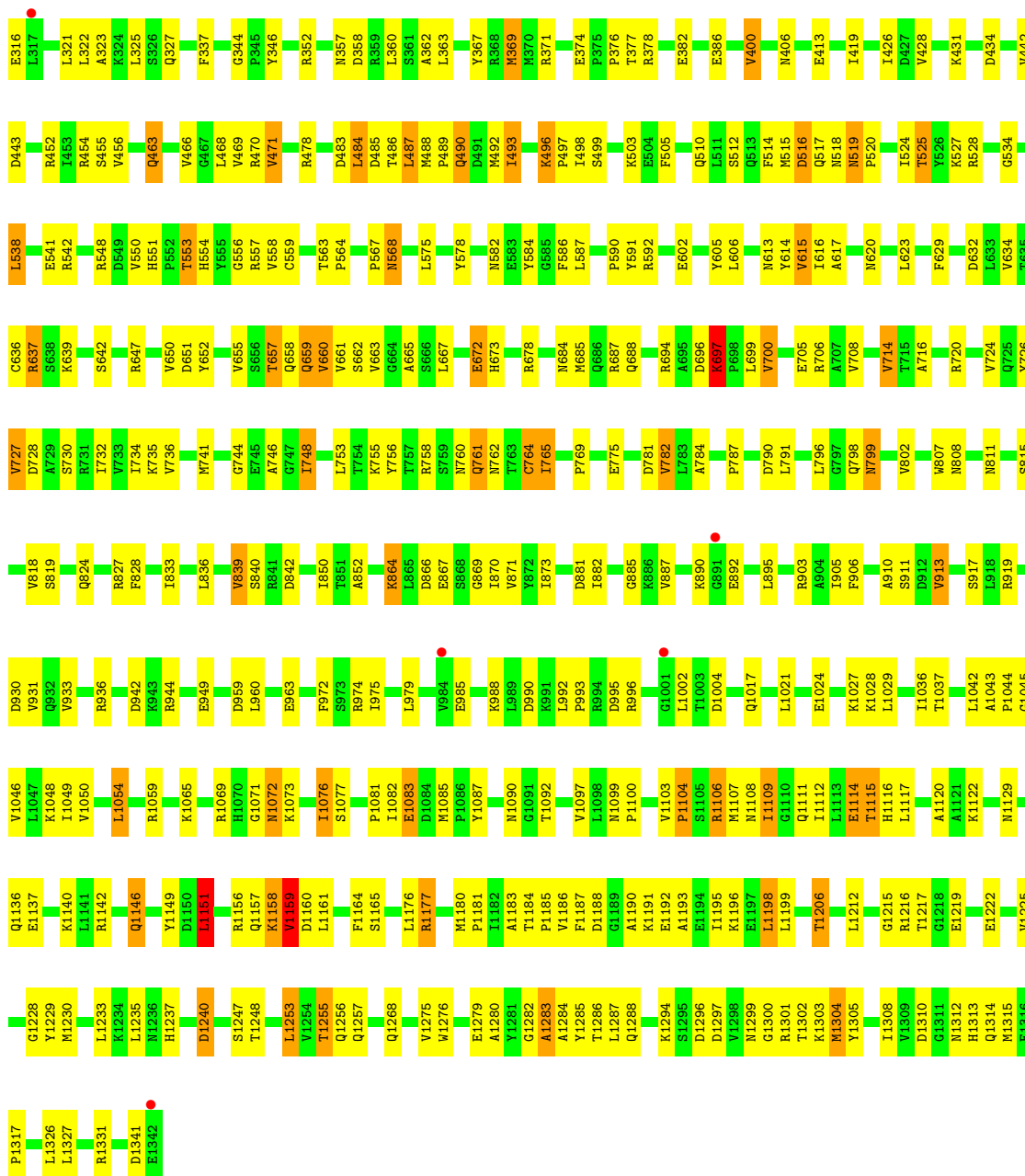
- Molecule 1: DNA-directed RNA polymerase subunit alpha



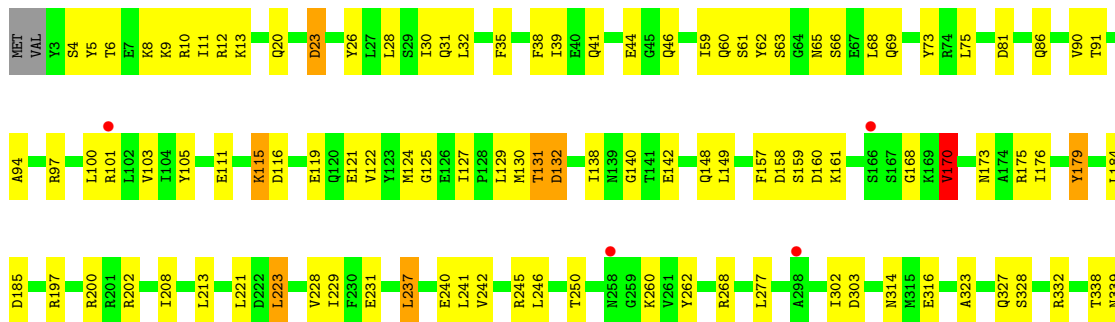
- Molecule 1: DNA-directed RNA polymerase subunit alpha

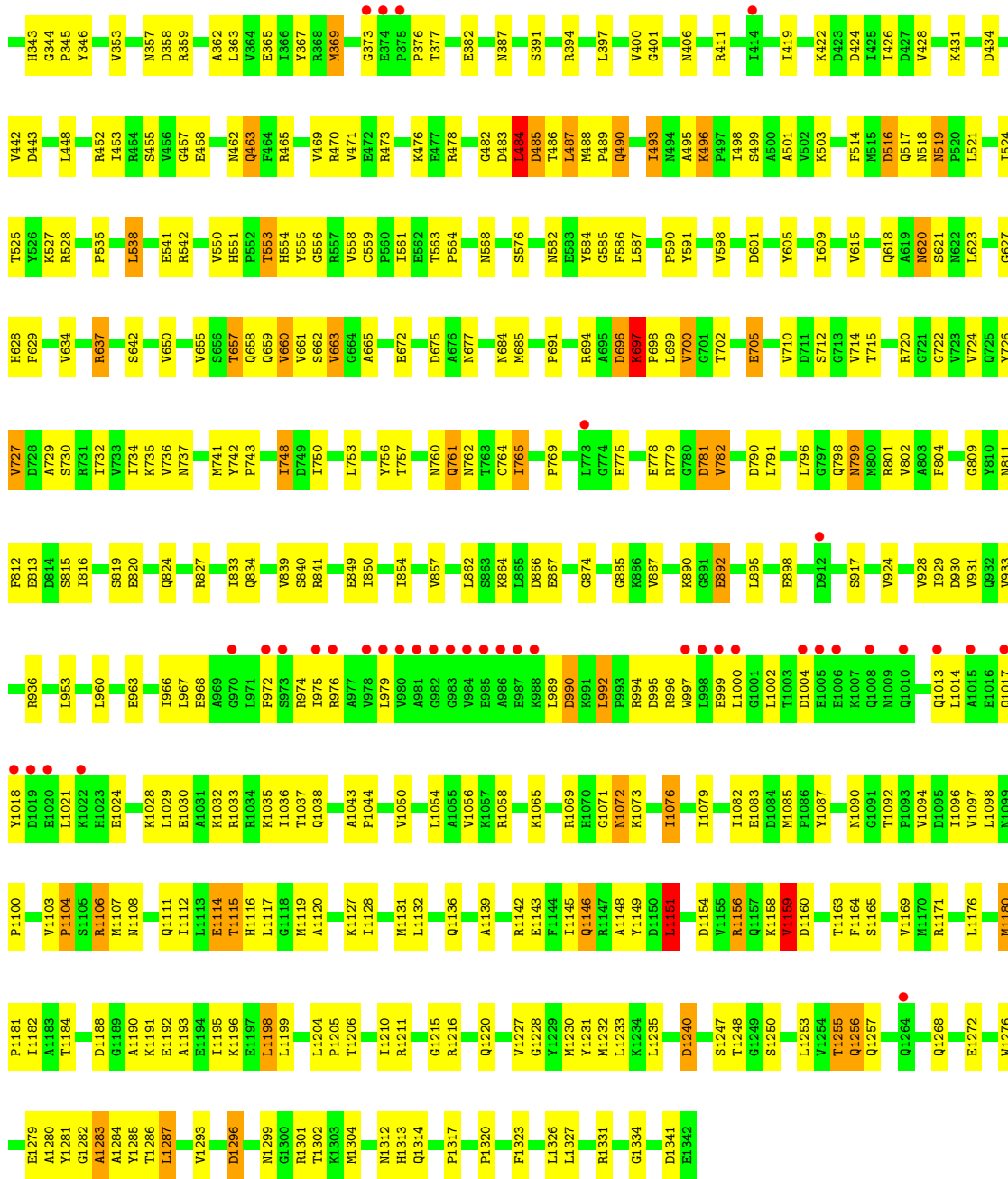




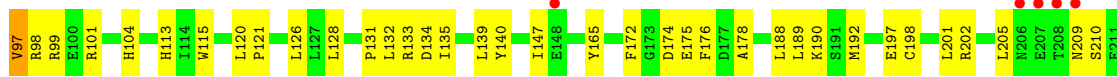
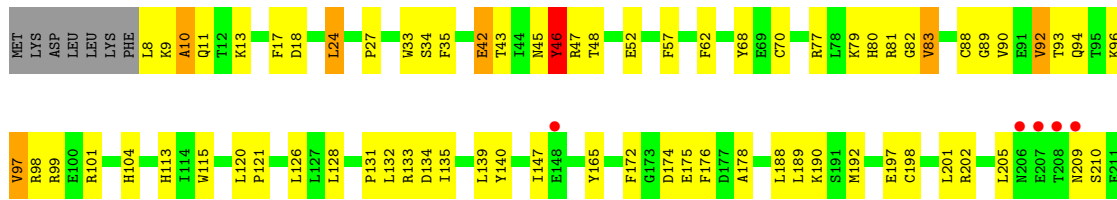


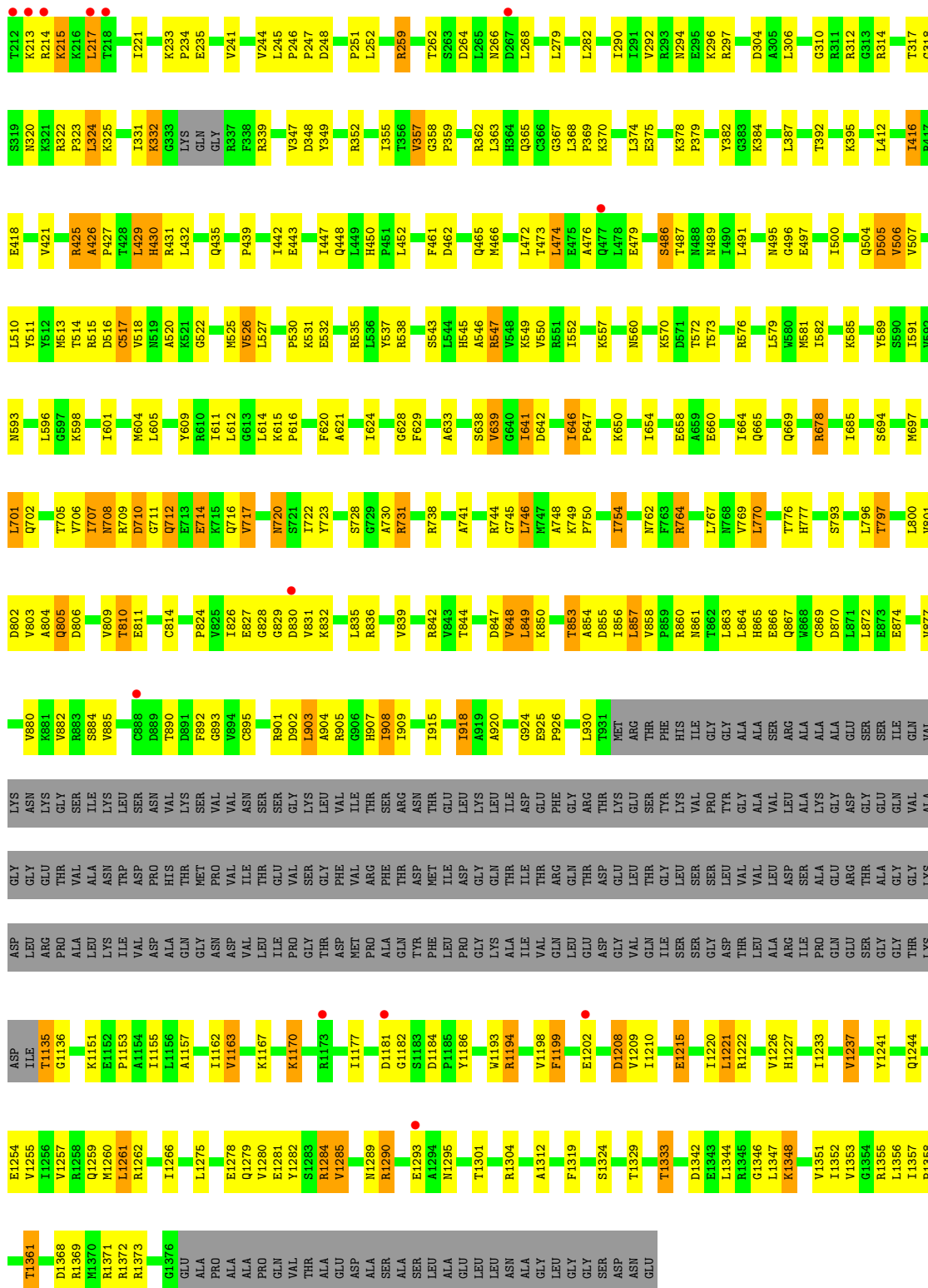
• Molecule 2: DNA-directed RNA polymerase subunit beta





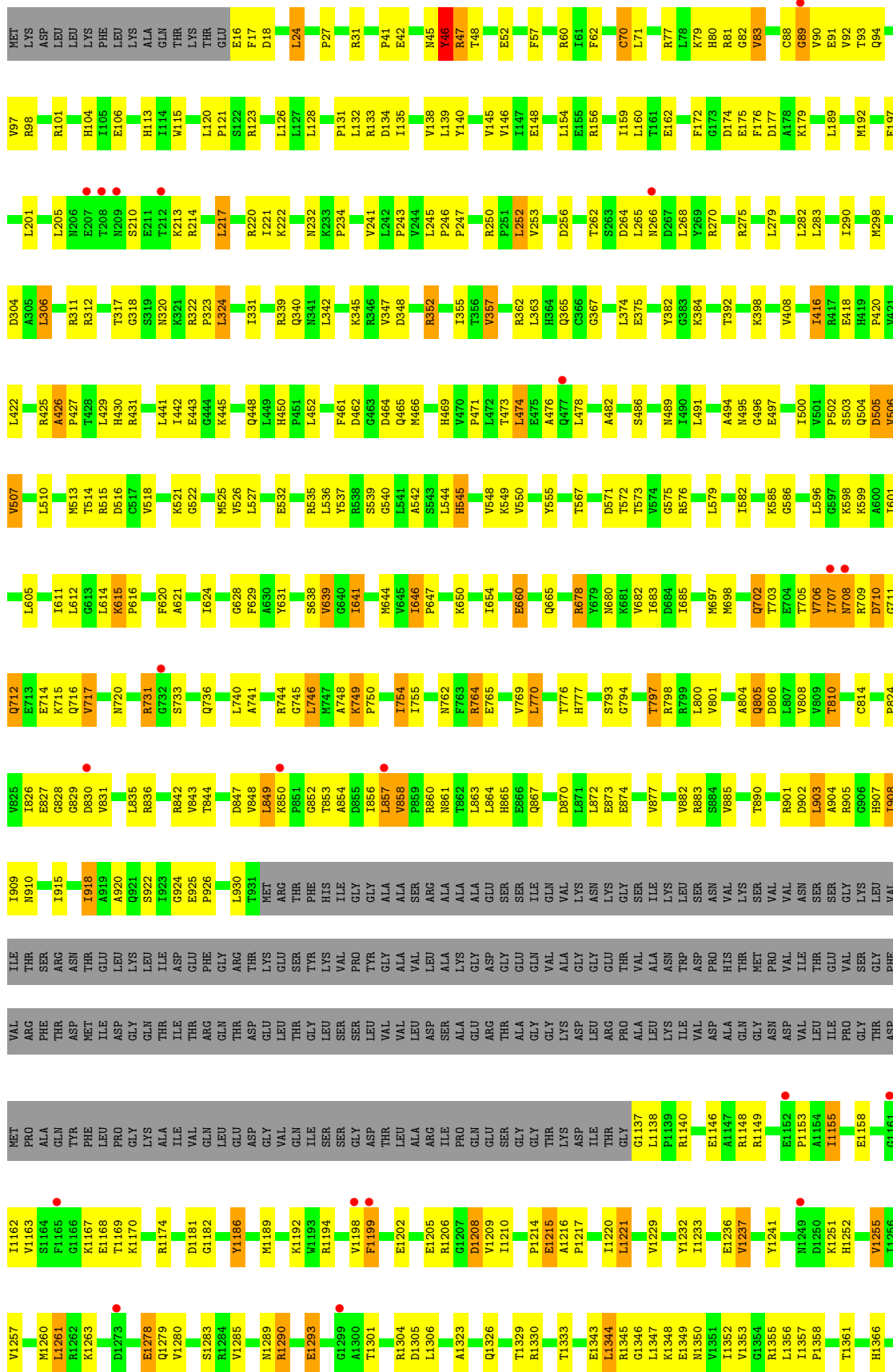
• Molecule 3: DNA-directed RNA polymerase subunit beta'

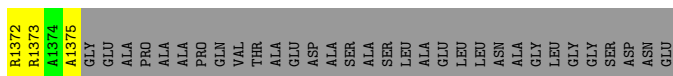




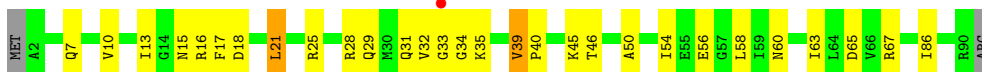
• Molecule 3: DNA-directed RNA polymerase subunit beta'



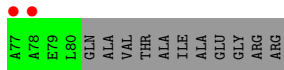
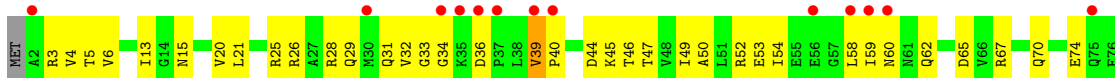




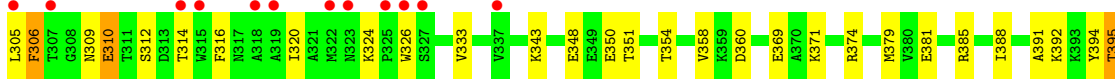
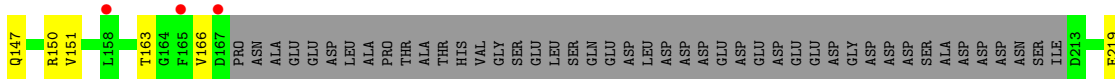
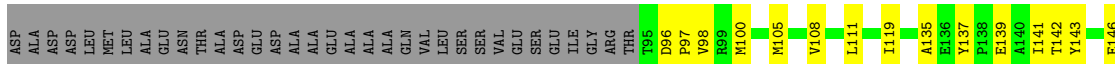
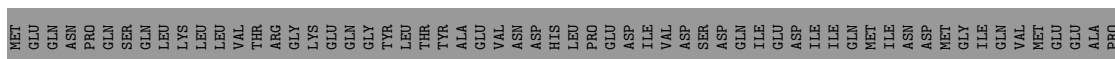
- Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 5: RNA polymerase sigma factor RpoD



Molecule 5: RNA polymerase sigma factor RpoD



MET	ASP	R157	K236	L341	L483	Q589
GLU	ALA	L198	ALA	L342	A484	I590
GLN	ASP	L198	LYS	Q442	A485	E591
ASN	ASP	L163	GLY	K343	R486	L595
PRO	LEU	G164	ARG	T351	M487	H600
GLN	MET	F165	SER	T354	L488	P601
LEU	LEU	V166	HIS	A243	M489	S602
GLN	ALA	D167	A243	T244	A490	R603
LEU	GLU	PRO	A245	Q357	D491	L607
LYS	ASN	ALA	E248	V358	D492	D612
LEU	THR	LEU	I249	K359	R495	D613
LEU	VAL	GLU	L250	D360	G512	
THR	ASP	GLU	K251	M379	D513	
ARG	ASP	LEU	L252	V380	D514	
GLY	ALA	ALA	S253	E381	E515	
LYS	ALA	PRO	A382	A382	I523	
GLU	GLU	THR	L261	N383	P631	
GLY	ALA	ALA	V262	L384	L532	
TYR	ALA	THR	P263	T395	T536	
LEU	GLN	HIS	K264	N396	L540	
THR	LEU	VAL	Q265	L402	T544	
THR	LEU	GLY	F266	M409	V547	
TYR	LEU	SER	D267	L410	A556	
ALA	SER	GLU	V268	G411	K557	
GLU	SER	LEU	L269	M413	V558	
VAL	VAL	SER	M273	D417	L559	
ASN	GLU	SER	M277	T429	R560	
ASN	ASP	GLU	M277	I435	M561	
ASP	SER	LEU	R281	R448	R562	
HIS	ASP	ASP	R285	T449	M567	
LEU	ASP	ASP	E293	I452	N568	
PRO	GLY	ASP	F302	R461	D570	
GLU	ASP	GLU	I303	I457	Y571	
ILE	GLU	ASP	T304	E458	T572	
ILE	GLY	ASP	E310	N461	L573	
ASP	ASP	ASP	T311	K462	E574	
ILE	ASP	ASP	S312	R465	E575	
GLN	ASP	ASP	D313	I466	G577	
MET	ALA	ALA	T314	M470	F580	
ILE	ASP	ASP	M315	N471	D581	
ASN	ASP	ASP	F316	Q472	V582	
ASP	ASP	ASP	N317	T479	T583	
MET	ASN	ASN	A318	E482	R584	
GLY	SER	SER	A319	V337	E585	
ILE	ILE	ILE	N322	I587	I587	
GLN	VAL	VAL	N323	R232	R588	
MET	GLU	GLU	L224			
GLU	MET	GLU	T231			
ALA	ALA	ALA	R232			
PRO	PRO	PRO	H338			

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	186.25Å 203.66Å 308.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.60 30.02 – 3.60	Depositor EDS
% Data completeness (in resolution range)	98.5 (30.00-3.60) 98.5 (30.02-3.60)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.58 (at 3.56Å)	Xtrriage
Refinement program	PHENIX (1.12_2829)	Depositor
R, $R_{free}$	0.244 , 0.311 0.244 , 0.311	Depositor DCC
$R_{free}$ test set	1979 reflections (1.48%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	119.5	Xtrriage
Anisotropy	0.210	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 87.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	54996	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	153.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.95% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, ZN, ECJ

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	0.30	0/1790	0.57	0/2426
1	B	0.27	0/1692	0.54	1/2293 (0.0%)
1	G	0.27	0/1751	0.56	0/2373
1	H	0.27	0/1686	0.55	1/2285 (0.0%)
2	C	0.26	0/10730	0.50	0/14479
2	I	0.26	0/10726	0.49	0/14474
3	D	0.26	0/9201	0.49	0/12420
3	J	0.26	0/9137	0.49	0/12337
4	E	0.26	0/687	0.48	0/928
4	K	0.23	0/629	0.47	0/847
5	F	0.25	0/3847	0.45	0/5171
5	L	0.25	0/3846	0.44	0/5171
All	All	0.26	0/55722	0.50	2/75204 (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	C	0	1
2	I	0	1
All	All	0	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	13	LEU	CA-CB-CG	5.80	128.64	115.30
1	B	13	LEU	CA-CB-CG	5.48	127.90	115.30



There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	1159	VAL	Peptide
2	I	1159	VAL	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	1768	0	1793	74	0
1	B	1672	0	1694	64	0
1	G	1730	0	1756	73	0
1	H	1667	0	1689	60	0
2	C	10561	0	10555	305	0
2	I	10557	0	10549	301	0
3	D	9063	0	9234	291	0
3	J	8998	0	9154	287	0
4	E	685	0	684	18	0
4	K	627	0	634	19	0
5	F	3796	0	3858	96	1
5	L	3796	0	3848	71	1
6	C	35	0	0	0	0
6	I	35	0	0	1	0
7	D	1	0	0	0	0
7	J	1	0	0	0	0
8	D	2	0	0	0	0
8	J	2	0	0	0	0
All	All	54996	0	55448	1510	1

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 14.

The worst 5 of 1510 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:45:ARG:HG2	1:H:38:THR:HB	1.49	0.94

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:I:696:ASP:HB2	2:I:798:GLN:HG2	1.51	0.92
2:C:696:ASP:HB2	2:C:798:GLN:HG2	1.53	0.91
2:I:10:ARG:HD3	2:I:1181:PRO:HG2	1.54	0.88
3:D:418:GLU:HG3	4:E:45:LYS:H	1.42	0.85

All (1) 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)
5:F:219:GLU:OE1	5:L:232:ARG:NH2[1_565]	2.18	0.02

## 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	226/329 (69%)	202 (89%)	16 (7%)	8 (4%)	3	30
1	B	213/329 (65%)	195 (92%)	16 (8%)	2 (1%)	17	57
1	G	222/329 (68%)	196 (88%)	20 (9%)	6 (3%)	5	35
1	H	213/329 (65%)	194 (91%)	17 (8%)	2 (1%)	17	57
2	C	1338/1342 (100%)	1203 (90%)	114 (8%)	21 (2%)	9	46
2	I	1338/1342 (100%)	1203 (90%)	112 (8%)	23 (2%)	9	45
3	D	1157/1407 (82%)	1038 (90%)	101 (9%)	18 (2%)	9	46
3	J	1151/1407 (82%)	1036 (90%)	100 (9%)	15 (1%)	12	50
4	E	87/91 (96%)	81 (93%)	4 (5%)	2 (2%)	6	38
4	K	77/91 (85%)	72 (94%)	4 (5%)	1 (1%)	12	50
5	F	461/613 (75%)	432 (94%)	25 (5%)	4 (1%)	17	57
5	L	463/613 (76%)	426 (92%)	33 (7%)	4 (1%)	17	57
All	All	6946/8222 (84%)	6278 (90%)	562 (8%)	106 (2%)	10	47

5 of 106 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	167	PRO
1	A	195	ARG
2	C	62	TYR
2	C	170	VAL
2	C	516	ASP

### 5.3.2 Protein sidechains [i](#)

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	196/286 (68%)	180 (92%)	16 (8%)	11	42
1	B	184/286 (64%)	173 (94%)	11 (6%)	19	54
1	G	191/286 (67%)	180 (94%)	11 (6%)	20	55
1	H	183/286 (64%)	167 (91%)	16 (9%)	10	41
2	C	1152/1157 (100%)	1047 (91%)	105 (9%)	9	39
2	I	1151/1157 (100%)	1049 (91%)	102 (9%)	9	40
3	D	968/1168 (83%)	869 (90%)	99 (10%)	7	34
3	J	959/1168 (82%)	864 (90%)	95 (10%)	8	35
4	E	71/75 (95%)	66 (93%)	5 (7%)	15	48
4	K	67/75 (89%)	60 (90%)	7 (10%)	7	33
5	F	413/540 (76%)	389 (94%)	24 (6%)	20	55
5	L	410/540 (76%)	384 (94%)	26 (6%)	18	53
All	All	5945/7024 (85%)	5428 (91%)	517 (9%)	10	41

5 of 517 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	918	ILE
3	J	1261	LEU
3	J	910	ASN
3	D	797	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	D	749	LYS

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

Mol	Chain	Res	Type
3	J	594	GLN
3	J	910	ASN
5	L	446	GLN
3	D	716	GLN
3	D	702	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 6 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
6	ECJ	I	1401	3,2	30,37,37	2.70	6 (20%)	32,59,59	3.22	8 (25%)
6	ECJ	C	1401	3,2	30,37,37	2.73	7 (23%)	32,59,59	2.91	8 (25%)

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
6	ECJ	I	1401	3,2	-	5/21/43/43	0/3/3/3
6	ECJ	C	1401	3,2	-	5/21/43/43	0/3/3/3

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	1401	ECJ	O16-C15	9.97	1.39	1.22
6	I	1401	ECJ	O16-C15	9.84	1.39	1.22
6	C	1401	ECJ	C13-C11	7.00	1.49	1.34
6	I	1401	ECJ	C13-C11	6.87	1.49	1.34
6	I	1401	ECJ	O12-C11	4.86	1.45	1.36

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	I	1401	ECJ	O16-C15-C13	-13.34	109.93	127.33
6	C	1401	ECJ	O16-C15-C13	-12.31	111.27	127.33
6	I	1401	ECJ	C15-C13-C11	-8.41	100.82	109.47
6	C	1401	ECJ	C15-C13-C11	-7.21	102.06	109.47
6	I	1401	ECJ	N27-C25-N24	-5.80	119.61	128.68

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

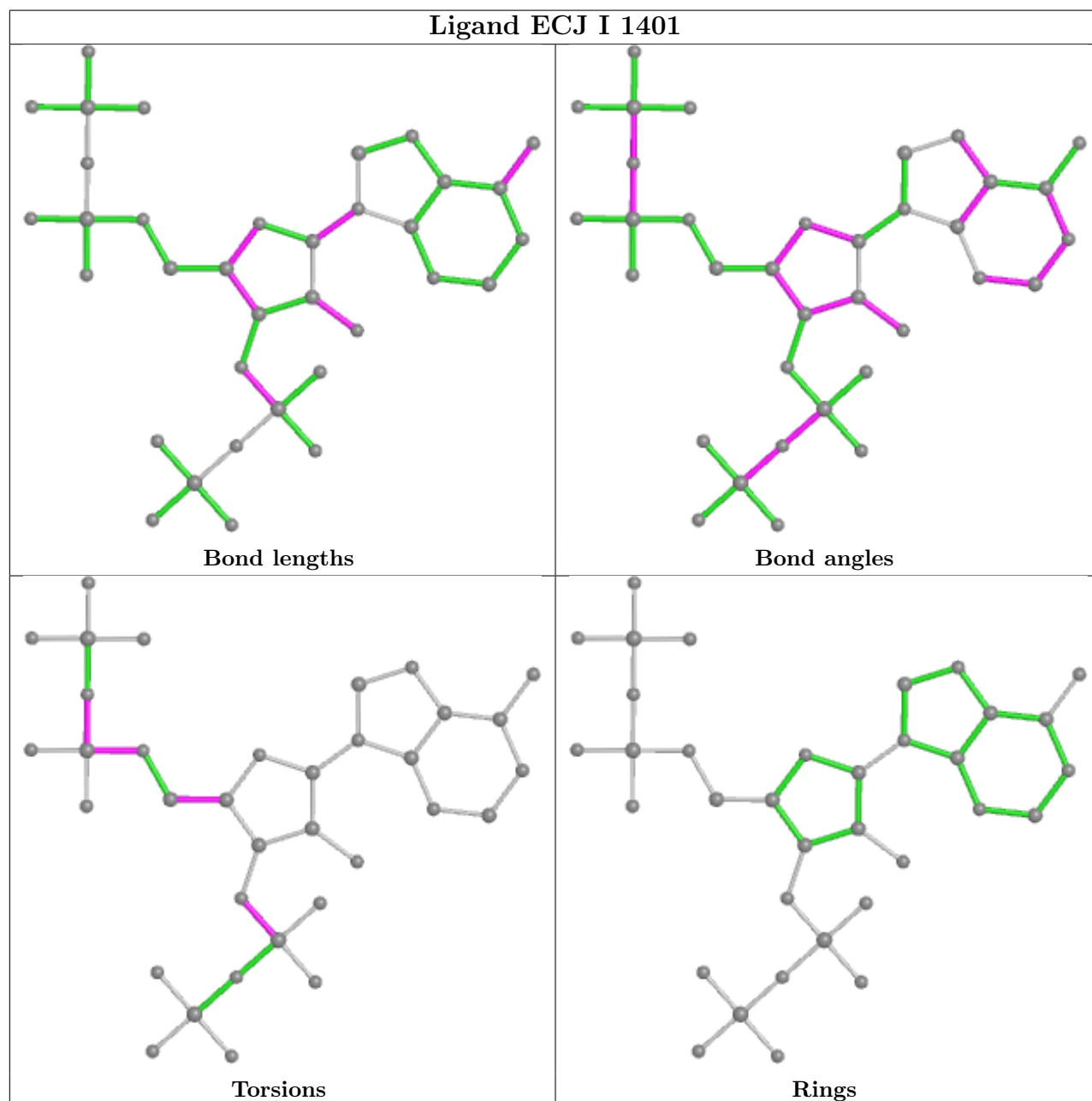
Mol	Chain	Res	Type	Atoms
6	C	1401	ECJ	O09-C10-C11-C13
6	I	1401	ECJ	O09-C10-C11-C13
6	I	1401	ECJ	P01-O05-P06-O09
6	I	1401	ECJ	C10-O09-P06-O05
6	C	1401	ECJ	O09-C10-C11-O12

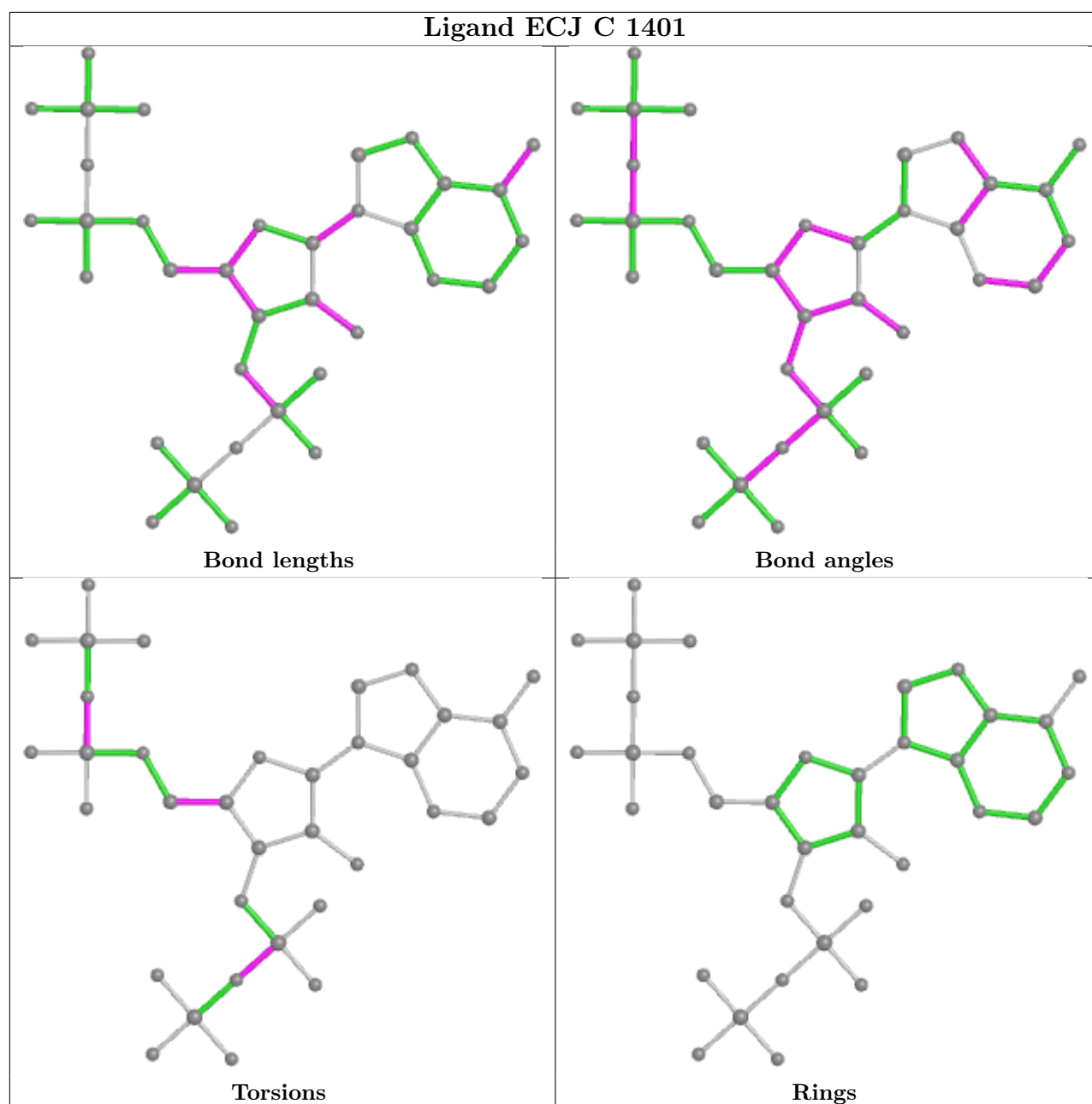
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	I	1401	ECJ	1	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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

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, 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	A	228/329 (69%)	-0.38	0 <b>100</b> <b>100</b>	90, 124, 199, 234	0
1	B	217/329 (65%)	0.05	5 (2%) 60 44	85, 176, 231, 252	0
1	G	224/329 (68%)	-0.09	3 (1%) 77 63	137, 176, 229, 280	0
1	H	217/329 (65%)	0.15	12 (5%) 25 15	150, 198, 235, 274	0
2	C	1340/1342 (99%)	-0.30	15 (1%) 80 68	67, 119, 204, 265	0
2	I	1340/1342 (99%)	-0.12	43 (3%) 47 32	100, 147, 231, 373	0
3	D	1163/1407 (82%)	-0.22	18 (1%) 73 60	63, 114, 199, 265	0
3	J	1155/1407 (82%)	-0.09	21 (1%) 68 53	82, 145, 217, 279	0
4	E	89/91 (97%)	-0.23	1 (1%) 80 68	107, 148, 174, 186	0
4	K	79/91 (86%)	0.94	15 (18%) 1 0	182, 265, 298, 321	0
5	F	467/613 (76%)	0.05	23 (4%) 29 18	97, 182, 330, 392	0
5	L	469/613 (76%)	-0.01	21 (4%) 33 21	121, 192, 300, 372	0
All	All	6988/8222 (84%)	-0.13	177 (2%) 57 41	63, 145, 240, 392	0

The worst 5 of 177 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	I	981	ALA	8.3
5	F	300	LYS	8.3
2	I	982	GLY	7.8
5	F	301	ASN	7.3
5	F	326	TRP	6.6

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

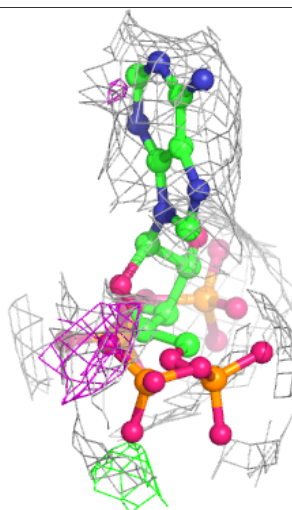
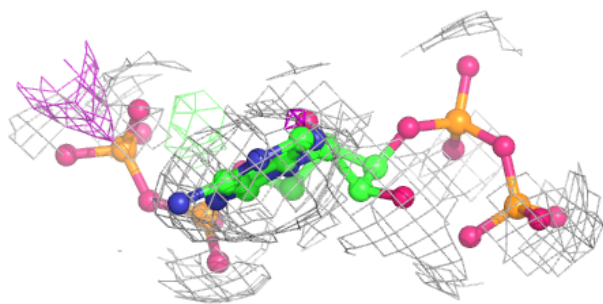
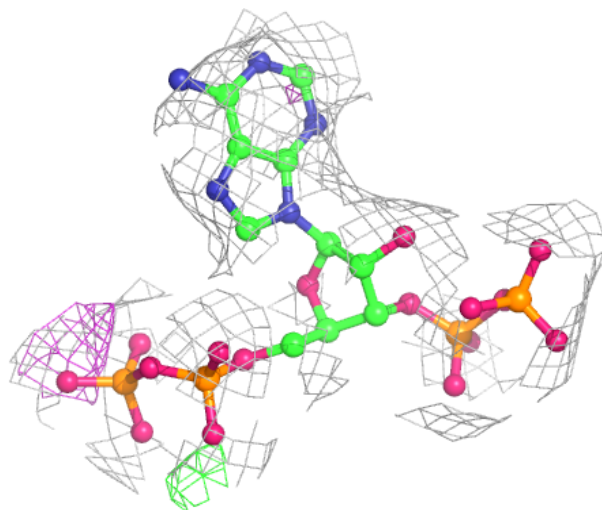
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

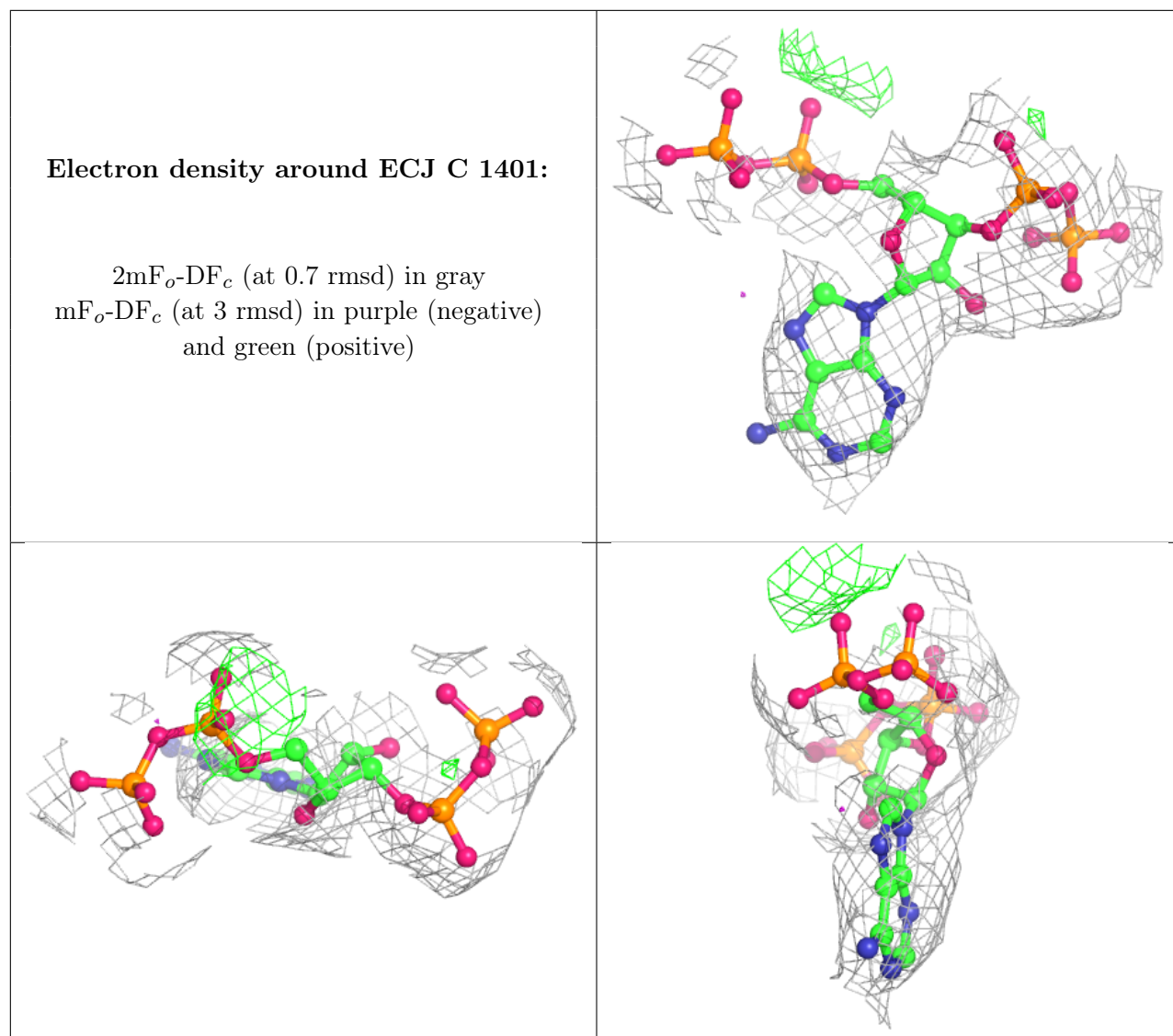
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
8	ZN	D	2003	1/1	0.83	0.64	377,377,377,377	0
6	ECJ	I	1401	35/35	0.88	0.18	63,90,174,177	35
6	ECJ	C	1401	35/35	0.88	0.20	74,99,183,185	35
7	MG	J	2001	1/1	0.92	0.31	81,81,81,81	0
7	MG	D	2001	1/1	0.93	0.16	47,47,47,47	0
8	ZN	J	2003	1/1	0.96	0.12	128,128,128,128	0
8	ZN	J	2002	1/1	0.98	0.04	155,155,155,155	0
8	ZN	D	2002	1/1	0.99	0.11	152,152,152,152	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 ECJ I 1401:**

$2mF_o-DF_c$  (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.