



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 08:38 am GMT

PDB ID : 4KMU  
Title : X-ray crystal structure of the Escherichia coli RNA polymerase in complex with Rifampin  
Authors : Murakami, K.S.  
Deposited on : 2013-05-08  
Resolution : 3.85 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

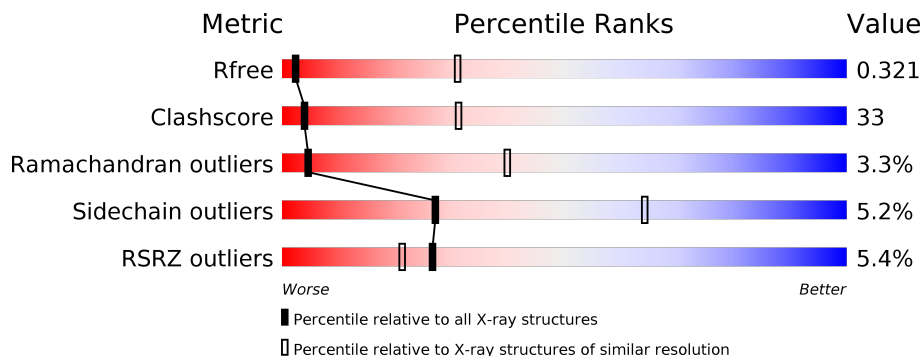
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.85 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))
$R_{free}$	100719	1009 (4.16-3.56)
Clashscore	112137	1029 (4.12-3.60)
Ramachandran outliers	110173	1017 (4.14-3.58)
Sidechain outliers	110143	1010 (4.14-3.58)
RSRZ outliers	101464	1023 (4.16-3.56)

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. 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>3%</div> <div> <div></div> <div>51%</div> <div>42%</div> <div>5%</div> </div> </div>
1	B	329	<div> <div>5%</div> <div> <div></div> <div>38%</div> <div>27%</div> <div>33%</div> </div> </div>
1	F	329	<div> <div>6%</div> <div> <div></div> <div>43%</div> <div>24%</div> <div>30%</div> </div> </div>
1	G	329	<div> <div>7%</div> <div> <div></div> <div>36%</div> <div>27%</div> <div>34%</div> </div> </div>
2	C	1342	<div> <div>3%</div> <div> <div></div> <div>46%</div> <div>48%</div> <div>5%</div> </div> </div>
2	H	1342	<div> <div>6%</div> <div> <div></div> <div>49%</div> <div>45%</div> <div>5%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	1407	<div><div></div><div>3%</div><div>32%</div><div>45%</div><div>5%</div><div>18%</div></div>
3	I	1407	<div><div></div><div>5%</div><div>35%</div><div>42%</div><div>5%</div><div>18%</div></div>
4	E	91	<div><div></div><div>54%</div><div>41%</div><div>• • •</div></div>
4	J	91	<div><div></div><div>4%</div><div>46%</div><div>33%</div><div>• •</div><div>16%</div></div>
5	X	613	<div><div></div><div>6%</div><div>44%</div><div>37%</div><div>•</div><div>16%</div></div>
5	Y	613	<div><div></div><div>5%</div><div>37%</div><div>36%</div><div>•</div><div>25%</div></div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 56315 atoms, of which 116 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	323	Total	C	N	O	S	0	0	0
			2514	1571	443	492	8			
1	B	221	Total	C	N	O	S	0	0	0
			1706	1065	300	335	6			
1	F	229	Total	C	N	O	S	0	0	0
			1775	1106	313	350	6			
1	G	217	Total	C	N	O	S	0	0	0
			1671	1045	293	327	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1335	Total	C	N	O	S	0	0	0
			10523	6601	1836	2043	43			
2	H	1335	Total	C	N	O	S	0	0	0
			10523	6601	1836	2043	43			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1160	Total	C	N	O	S	0	0	0
			9060	5695	1621	1697	47			
3	I	1160	Total	C	N	O	S	0	0	0
			9060	5695	1621	1697	47			

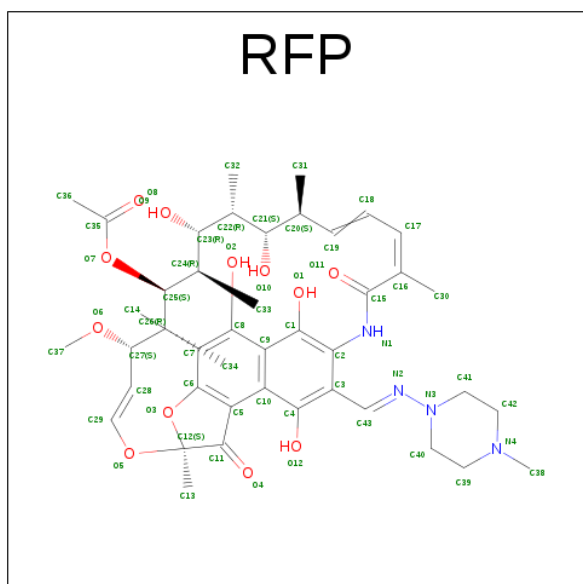
- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	J	76	Total	C	N	O	S	0	0	0
			605	368	115	121	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	X	517	Total	C	N	O	S	0	0	0
			4198	2621	745	806	26			
5	Y	458	Total	C	N	O	S	0	0	0
			3732	2335	671	703	23			

- Molecule 6 is RIFAMPICIN (three-letter code: RFP) (formula:  $C_{43}H_{58}N_4O_{12}$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	C	1	Total	C	H	N	O	0	0
			117	43	58	4	12		
6	H	1	Total	C	H	N	O	0	0
			117	43	58	4	12		

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

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

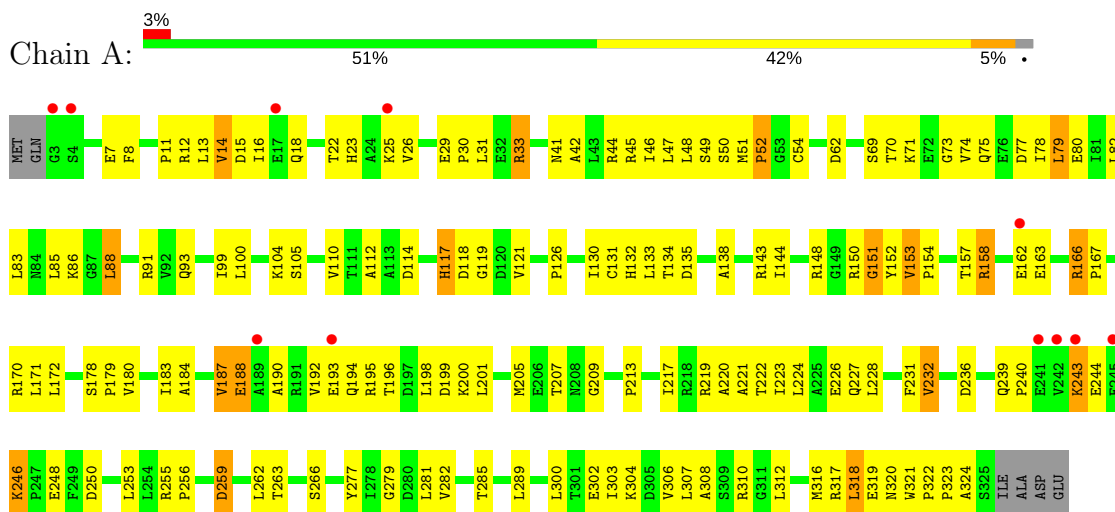
- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	I	1	Total 1	Mg 1	0	0
8	D	1	Total 1	Mg 1	0	0

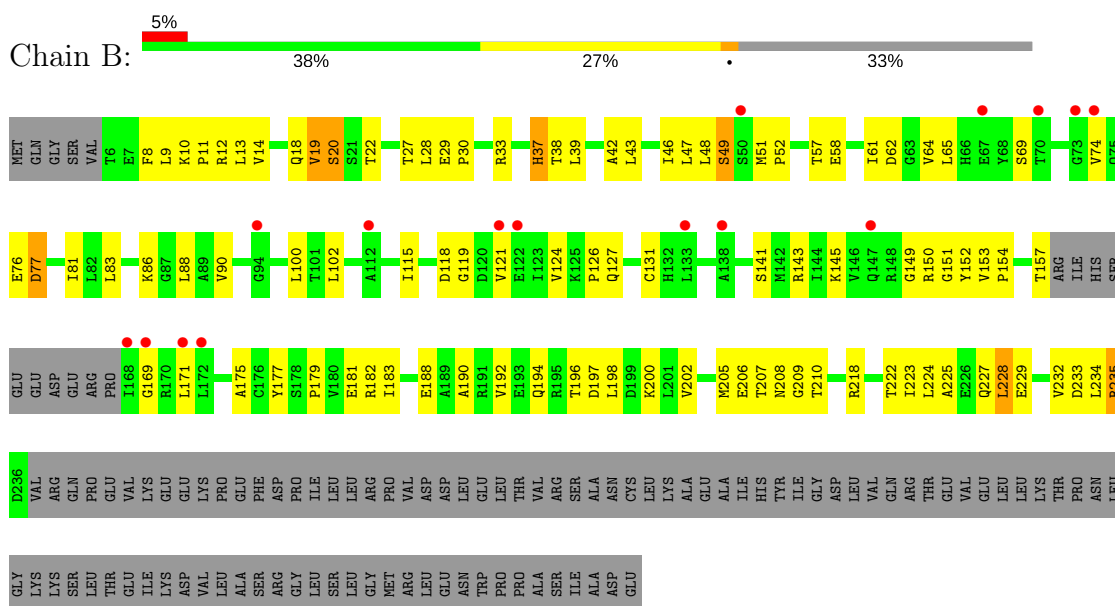
### 3 Residue-property plots [i](#)

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

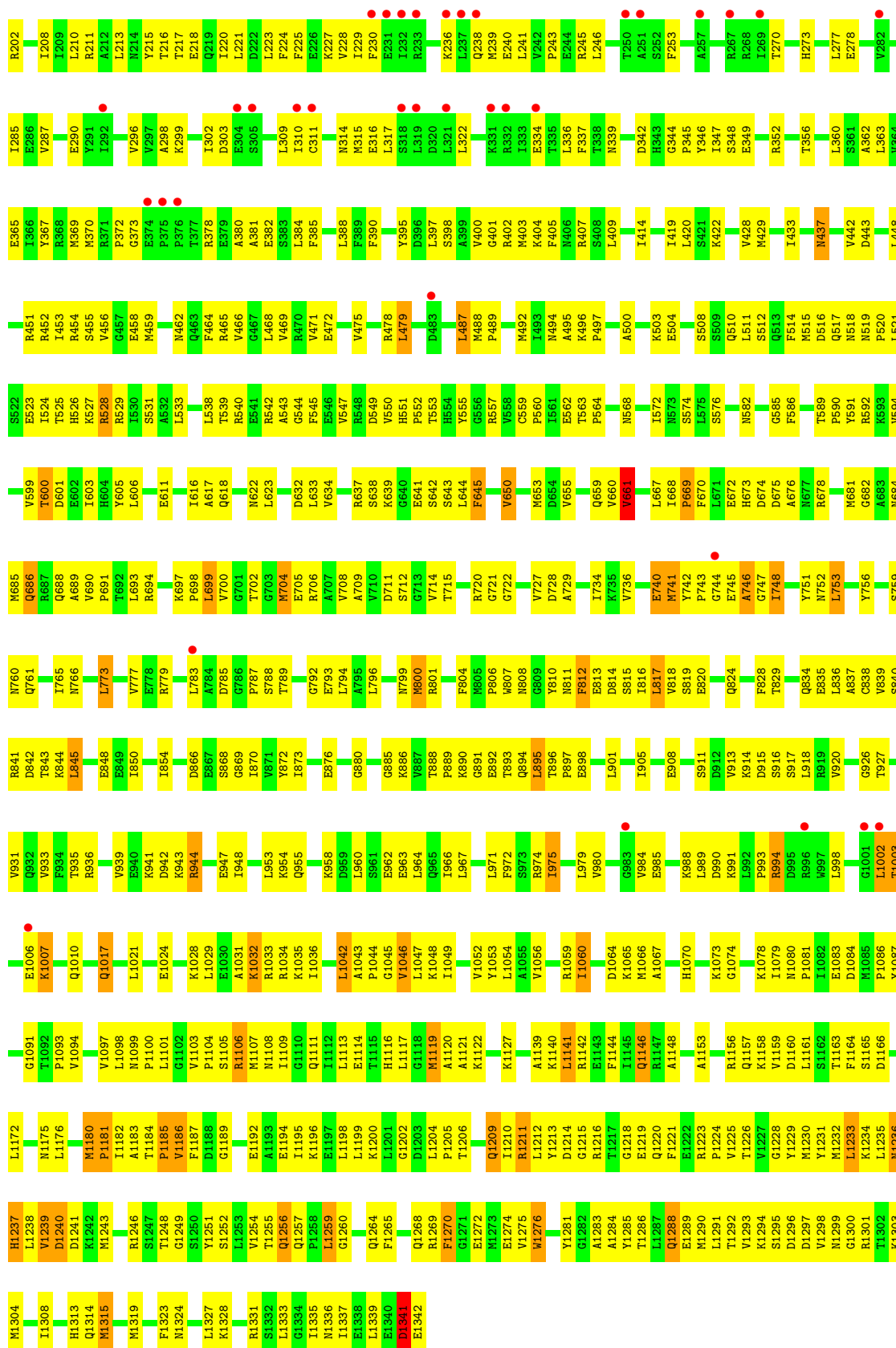


#### • Molecule 1: DNA-directed RNA polymerase subunit alpha



#### • Molecule 1: DNA-directed RNA polymerase subunit alpha

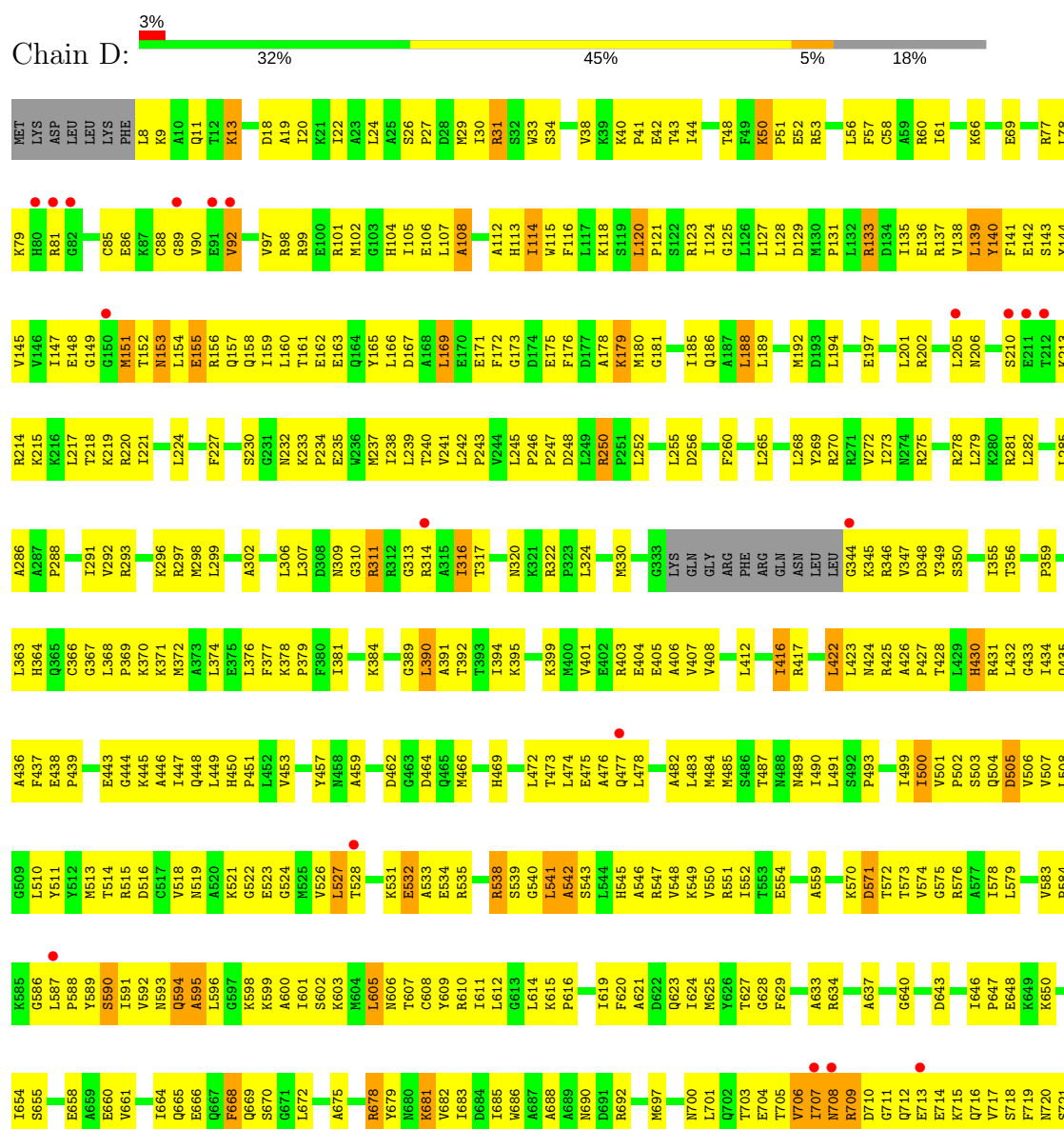


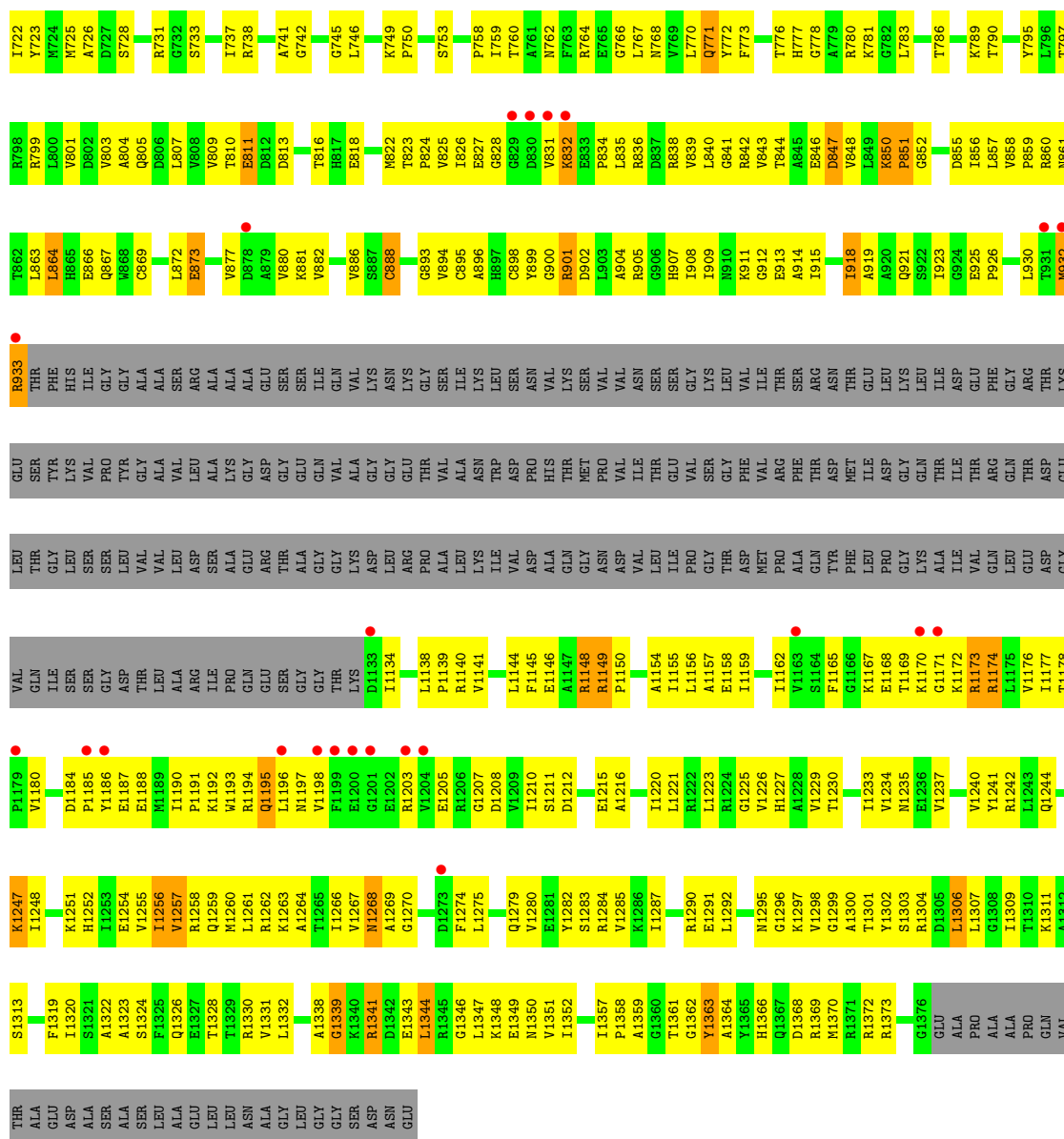




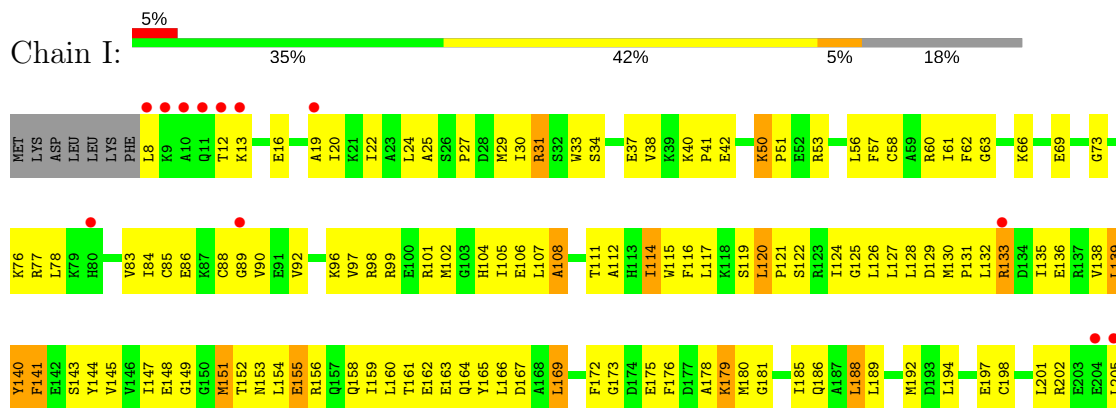


### • Molecule 3: DNA-directed RNA polymerase subunit beta'

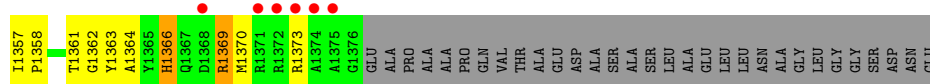




• Molecule 3: DNA-directed RNA polymerase subunit beta'



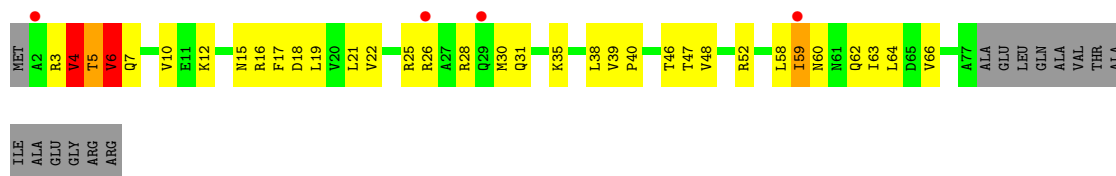
E1281	R1203	GLU	ARG	ASP	GLU	D882	I722	D643	T572	L423	R346	N206
Y1282	E1204	SER	THR	GLY	GLY	V803	Y723	M644	T573	R424	V347	E207
S1283	V1205	ALA	GLY	GLU	ILE	A804	M724	V645	V574	R425	D348	T208
R1284	E1206	GLY	GLY	GLU	GLY	K805	M725	G575	G575	R425	Y349	N209
V1285	G1207	ALA	LYS	VAL	ALA	D806	I646	R576	R576	P502	S350	S210
K1286	D1208	ASP	ASP	GLY	GLY	L807	I647	P647	A577	P503	I355	E211
I1287	D1209	LEU	LEU	GLY	LYS	V808	E648	E648	I578	T428	A286	T212
	V1209				ASN	R809	K649	K650	L579	H430	A287	K213
	S1211				LYS	T810	S733	R650	V583	G433	P288	R214
					LYS	E811	A735	E858	P584	L361	I361	K215
	A1216				ALA	D812	A736	L508	L508	I434	I291	L217
					ASN	D813	I737	K585	G509	R362	V292	T218
					LYS	T816	R738	E860	L510	L363	R293	
					LEU		R739	V661		H384	N294	
					SER					Q365	E295	
					ASP					I365	E296	
					PRO					L366	K221	
					THR					I434	K216	
					ALA					R362	L217	
					GLY					L363	T218	
					MET					H384		
					PRO					Q365		
					VAL					E296		
					ASN					K366		
					ASP					G367		
					VAL					L368		
					ILE					N298		
					GLY					L299		
					GLY					Q300		
					LYS					K370		
					LYS					K371		
					LEU					M372		
					LEU					L376		
					LEU					E235		
					LEU					E236		
					LEU					M237		
					LEU					L238		
					LEU					L239		
					LEU					L240		
					LEU					L241		
					LEU					L242		
					LEU					L243		
					LEU					L244		
					LEU					L245		
					LEU					L246		
					LEU					L247		
					LEU					L248		
					LEU					L249		
					LEU					L250		
					LEU					L251		
					LEU					L252		
					LEU					L253		
					LEU					L254		
					LEU					L255		
					LEU					L256		
					LEU					L257		
					LEU					L258		
					LEU					L259		
					LEU					L260		
					LEU					L261		
					LEU					L262		
					LEU					L263		
					LEU					L264		
					LEU					L265		
					LEU					L266		
					LEU					L267		
					LEU					L268		
					LEU					L269		
					LEU					L270		
					LEU					L271		
					LEU					L272		
					LEU					L273		
					LEU					L274		
					LEU					L275		
					LEU					L276		
					LEU					L277		
					LEU					L278		
					LEU					L279		
					LEU					L280		
					LEU					L281		
					LEU					L282		
					LEU					L283		
					LEU					L284		
					LEU					L285		
					LEU					L286		
					LEU					L287		
					LEU					L288		
					LEU					L289		
					LEU					L290		
					LEU					L291		
					LEU					L292		
					LEU					L293		
					LEU					L294		
					LEU					L295		
					LEU					L296		
					LEU					L297		
					LEU					L298		
					LEU					L299		
					LEU					L300		
					LEU					L301		
					LEU					L302		
					LEU					L303		
					LEU					L304		
					LEU					L305		
					LEU					L306		
					LEU					L307		
					LEU					L308		
					LEU					L309		
					LEU					L310		
					LEU					L311		
					LEU					L312		
					LEU					L313		
					LEU					L314		
					LEU					L315		
					LEU					L316		
					LEU					L317		
					LEU					L318		
					LEU					L319		
					LEU					L320		
					LEU					L321		
					LEU					L322		
					LEU					L323		
					LEU					L324		
					LEU					L325		
					LEU					L326		
					LEU					L327		
					LEU					L328		
					LEU					L329		
					LEU					L330		
					LEU					L331		
					LEU					L332		
					LEU					L333		
					LEU					L334		
					LEU					L335		
					LEU					L336		
					LEU					L337		
					LEU					L338		
					LEU					L339		
					LEU					L340		
					LEU					L341		
					LEU					L342		
					LEU					L343		
					LEU					L344		
					LEU					L345		
					LEU					L346		
					LEU					L347		
					LEU					L348		
					LEU					L349		
					LEU					L350		
					LEU					L351		
					LEU					L352		
					LEU					L353		
					LEU					L354		
					LEU					L355		
					LEU					L356		
					LEU					L357		
					LEU					L358		
					LEU					L359		
					LEU					L360		
					LEU					L361		
					LEU					L362		
					LEU					L363		
					LEU					L364		
					LEU					L365		
					LEU					L366		
					LEU					L367		
					LEU					L368		
					LEU					L369		
					LEU					L370		
					LEU					L371		
					LEU					L372		
					LEU					L373		
					LEU					L374		
					LEU					L375		
					LEU					L376		
					LEU					L377		
					LEU					L378		
					LEU					L379		
					LEU					L380		
					LEU					L381		
					LEU					L382		
					LEU					L383		
					LEU					L384		
					LEU					L385		
					LEU					L386		
					LEU					L387		
					LEU					L388		



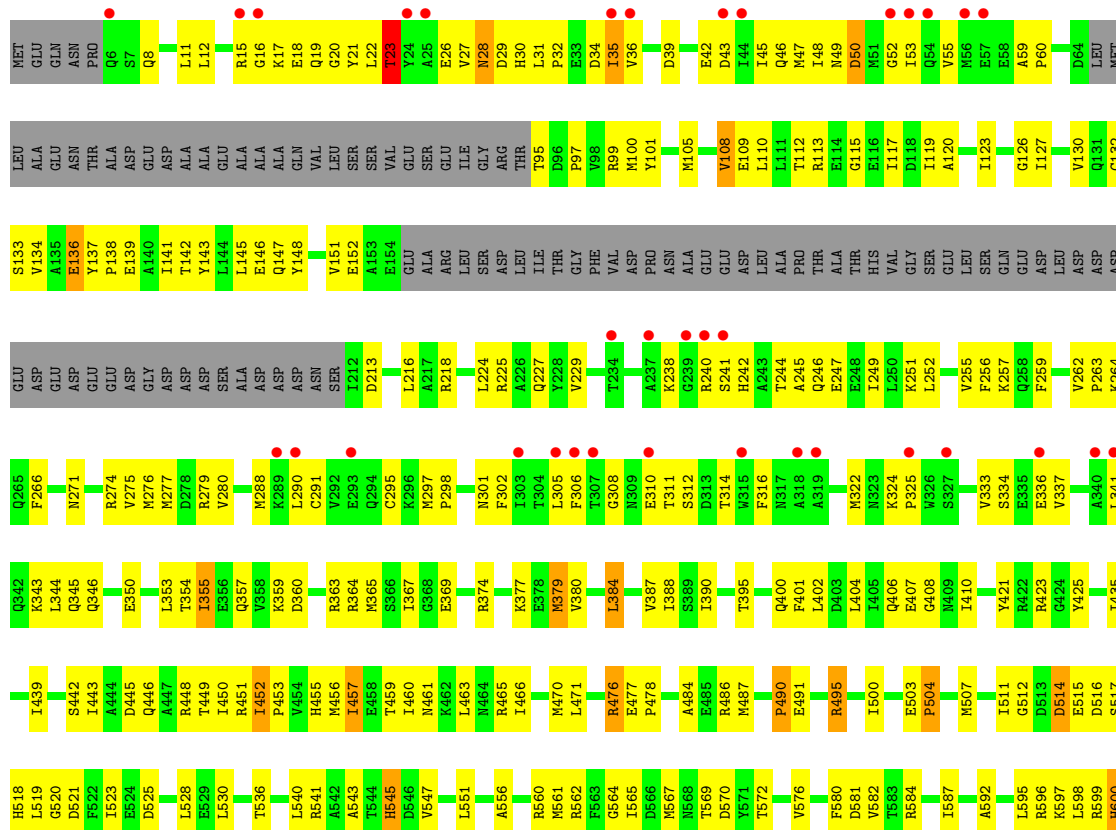
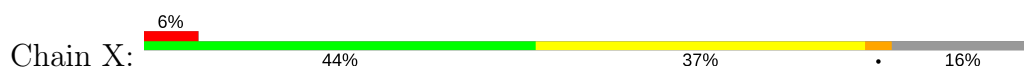
- Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 5: RNA polymerase sigma factor RpoD





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	184.52Å 203.87Å 307.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.97 – 3.85 30.75 – 3.81	Depositor EDS
% Data completeness (in resolution range)	92.4 (29.97-3.85) 85.5 (30.75-3.81)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.25 (at 3.75Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.264 , 0.321 0.264 , 0.321	Depositor DCC
$R_{free}$ test set	5088 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	112.9	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 42.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.20$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	56315	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.67% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, RFP, 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	0.19	0/2548	0.38	0/3454
1	B	0.20	0/1725	0.42	0/2337
1	F	0.20	0/1797	0.41	0/2436
1	G	0.20	0/1690	0.41	0/2290
2	C	0.21	0/10690	0.42	0/14423
2	H	0.22	0/10690	0.42	0/14423
3	D	0.20	0/9198	0.42	0/12413
3	I	0.20	0/9198	0.42	0/12413
4	E	0.19	0/710	0.38	0/956
4	J	0.19	0/607	0.39	0/817
5	X	0.20	0/4253	0.39	0/5719
5	Y	0.20	0/3783	0.39	0/5083
All	All	0.21	0/56889	0.41	0/76764

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2514	0	2566	170	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1706	0	1738	108	0
1	F	1775	0	1800	77	0
1	G	1671	0	1706	92	0
2	C	10523	0	10546	801	0
2	H	10523	0	10546	701	0
3	D	9060	0	9257	808	0
3	I	9060	0	9257	751	0
4	E	708	0	719	51	0
4	J	605	0	612	44	0
5	X	4198	0	4250	243	0
5	Y	3732	0	3809	211	0
6	C	59	58	56	9	0
6	H	59	58	56	14	0
7	D	2	0	0	0	0
7	I	2	0	0	0	0
8	D	1	0	0	0	0
8	I	1	0	0	0	0
All	All	56199	116	56918	3754	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 33.

The worst 5 of 3754 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:H:1119:MET:HG2	2:H:1228:GLY:HA2	1.20	1.17
3:D:1173:ARG:HA	3:D:1174:ARG:HB2	1.26	1.14
3:I:610:ARG:HG3	3:I:864:LEU:HD13	1.29	1.14
2:C:700:VAL:HG11	2:C:1114:GLU:HG3	1.30	1.12
3:I:1173:ARG:HA	3:I:1174:ARG:HB2	1.28	1.12

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	321/329 (98%)	254 (79%)	52 (16%)	15 (5%)	3	30
1	B	217/329 (66%)	188 (87%)	23 (11%)	6 (3%)	6	43
1	F	227/329 (69%)	194 (86%)	28 (12%)	5 (2%)	8	47
1	G	213/329 (65%)	188 (88%)	20 (9%)	5 (2%)	7	47
2	C	1333/1342 (99%)	1066 (80%)	225 (17%)	42 (3%)	5	40
2	H	1333/1342 (99%)	1065 (80%)	222 (17%)	46 (4%)	4	38
3	D	1154/1407 (82%)	919 (80%)	193 (17%)	42 (4%)	4	38
3	I	1154/1407 (82%)	925 (80%)	192 (17%)	37 (3%)	5	40
4	E	88/91 (97%)	76 (86%)	7 (8%)	5 (6%)	2	27
4	J	74/91 (81%)	64 (86%)	5 (7%)	5 (7%)	1	23
5	X	511/613 (83%)	444 (87%)	54 (11%)	13 (2%)	6	45
5	Y	454/613 (74%)	410 (90%)	33 (7%)	11 (2%)	7	46
All	All	7079/8222 (86%)	5793 (82%)	1054 (15%)	232 (3%)	4	40

5 of 232 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	52	PRO
1	B	20	SER
1	B	52	PRO
2	C	21	VAL
2	C	39	ILE

### 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	281/286 (98%)	270 (96%)	11 (4%)	37	70
1	B	189/286 (66%)	184 (97%)	5 (3%)	51	77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	197/286 (69%)	191 (97%)	6 (3%)	46	75
1	G	185/286 (65%)	180 (97%)	5 (3%)	50	77
2	C	1150/1157 (99%)	1084 (94%)	66 (6%)	24	62
2	H	1150/1157 (99%)	1084 (94%)	66 (6%)	24	62
3	D	971/1168 (83%)	911 (94%)	60 (6%)	21	59
3	I	971/1168 (83%)	913 (94%)	58 (6%)	22	60
4	E	74/75 (99%)	72 (97%)	2 (3%)	50	77
4	J	65/75 (87%)	63 (97%)	2 (3%)	45	74
5	X	460/540 (85%)	442 (96%)	18 (4%)	37	70
5	Y	407/540 (75%)	388 (95%)	19 (5%)	30	66
All	All	6100/7024 (87%)	5782 (95%)	318 (5%)	27	64

5 of 318 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	E	6	VAL
2	H	46	GLN
3	I	1306	LEU
5	X	136	GLU
5	X	607	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 97 such sidechains are listed below:

Mol	Chain	Res	Type
4	E	31	GLN
5	X	437	GLN
5	Y	242	HIS
5	X	28	ASN
5	X	242	HIS

### 5.3.3 RNA ⓘ

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 carbohydrates 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	RFP	C	1401	-	62,63,63	2.04	10 (16%)	91,94,94	2.07	24 (26%)
6	RFP	H	1401	-	62,63,63	2.10	9 (14%)	91,94,94	1.85	24 (26%)

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	RFP	C	1401	-	-	0/60/85/85	0/1/5/5
6	RFP	H	1401	-	-	0/60/85/85	0/1/5/5

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	1401	RFP	C12-C11	-3.63	1.36	1.54
6	H	1401	RFP	C12-C11	-3.59	1.36	1.54
6	H	1401	RFP	O7-C25	-3.58	1.39	1.44
6	C	1401	RFP	O7-C25	-2.91	1.40	1.44
6	C	1401	RFP	C2-N1	2.06	1.47	1.43

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	C	1401	RFP	C2-C3-C43	-5.85	115.01	123.29
6	H	1401	RFP	C17-C18-C19	-4.59	112.60	124.55
6	C	1401	RFP	O4-C11-C5	-4.29	122.02	131.84
6	H	1401	RFP	O4-C11-C5	-4.12	122.40	131.84
6	C	1401	RFP	C40-N3-N2	-3.63	93.88	113.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	1401	RFP	9	0
6	H	1401	RFP	14	0

## 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	323/329 (98%)	0.09	11 (3%) 46 36	0, 73, 165, 263	0
1	B	221/329 (67%)	0.39	16 (7%) 16 12	3, 97, 189, 266	0
1	F	229/329 (69%)	0.50	20 (8%) 11 9	16, 121, 201, 266	0
1	G	217/329 (65%)	0.48	23 (10%) 7 6	39, 111, 186, 215	0
2	C	1335/1342 (99%)	-0.04	46 (3%) 46 36	0, 48, 166, 284	0
2	H	1335/1342 (99%)	0.19	87 (6%) 20 14	1, 86, 201, 341	0
3	D	1160/1407 (82%)	-0.02	42 (3%) 43 34	0, 40, 157, 284	0
3	I	1160/1407 (82%)	0.14	70 (6%) 23 17	1, 52, 180, 322	0
4	E	90/91 (98%)	-0.34	0 100 100	0, 40, 109, 159	0
4	J	76/91 (83%)	0.23	4 (5%) 27 22	5, 76, 155, 167	0
5	X	517/613 (84%)	0.24	35 (6%) 18 14	3, 99, 228, 365	0
5	Y	458/613 (74%)	0.17	33 (7%) 16 12	2, 102, 219, 328	0
All	All	7121/8222 (86%)	0.12	387 (5%) 26 21	0, 70, 190, 365	0

The worst 5 of 387 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	1002	LEU	16.4
2	H	1001	GLY	12.6
3	I	10	ALA	10.7
5	X	319	ALA	10.3
2	C	231	GLU	9.5

### 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. 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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	RFP	H	1401	59/59	0.83	0.36	1.33	20,20,20,20	0
6	RFP	C	1401	59/59	0.90	0.29	0.81	20,20,20,20	0
7	ZN	I	1502	1/1	0.96	0.17	-0.54	49,49,49,49	0
7	ZN	D	1502	1/1	0.97	0.16	-0.71	8,8,8,8	0
7	ZN	I	1501	1/1	0.97	0.05	-1.46	60,60,60,60	0
7	ZN	D	1501	1/1	0.99	0.07	-1.48	54,54,54,54	0
8	MG	D	1503	1/1	0.89	0.16	-	24,24,24,24	0
8	MG	I	1503	1/1	0.97	0.70	-	20,20,20,20	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.