



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Feb 21, 2018 – 02:31 PM EST

PDB ID : 6FBV  
EMDB ID: : EMD-4230  
Title : Single particle cryo em structure of Mycobacterium tuberculosis RNA polymerase in complex with Fidaxomicin  
Authors : Das, K.; Lin, W.; Ebright, E.  
Deposited on : 2017-12-19  
Resolution : 3.50 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB 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/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

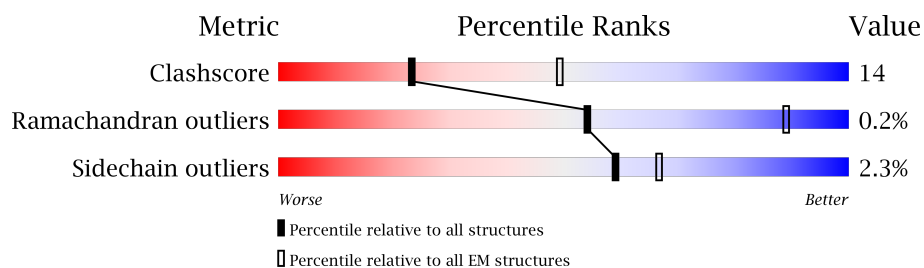
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.50 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)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	347	
1	B	347	
2	C	1178	
3	D	1316	
4	E	110	
5	F	528	

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 25146 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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					AltConf	Trace
1	A	223	Total	C	N	O	S	0	0
			1700	1068	294	336	2		
1	B	230	Total	C	N	O	S	0	0
			1749	1099	301	347	2		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	LEU	ILE	conflict	UNP P9WGZ1
B	3	LEU	ILE	conflict	UNP P9WGZ1

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	1118	Total	C	N	O	S	0	0
			8671	5427	1521	1684	39		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	1270	Total	C	N	O	S	0	0
			9951	6224	1808	1878	41		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
4	E	83	Total	C	N	O	0	0
			654	417	108	129		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	293	Total	C	N	O	S	0	0
			2343	1465	424	445	9		

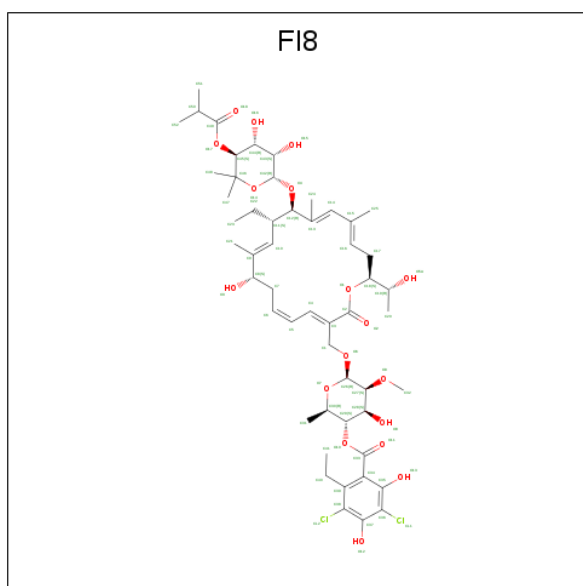
- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
6	D	2	Total	Zn	0
			2	2	

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

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

- Molecule 8 is Fidaxomicin (three-letter code: FI8) (formula: C<sub>52</sub>H<sub>74</sub>Cl<sub>2</sub>O<sub>18</sub>).



Mol	Chain	Residues	Atoms				AltConf
8	D	1	Total	C	Cl	O	0
			72	52	2	18	

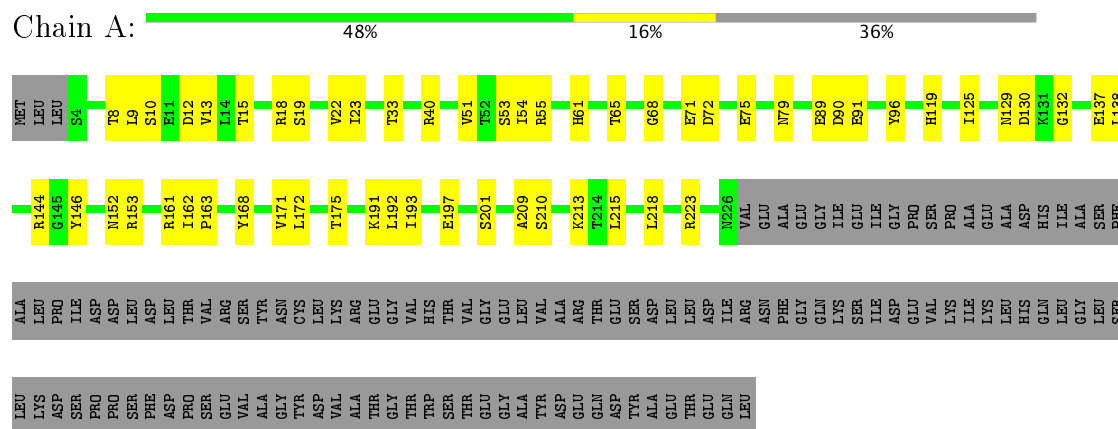
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		AltConf
9	D	3	Total	O	0
			3	3	

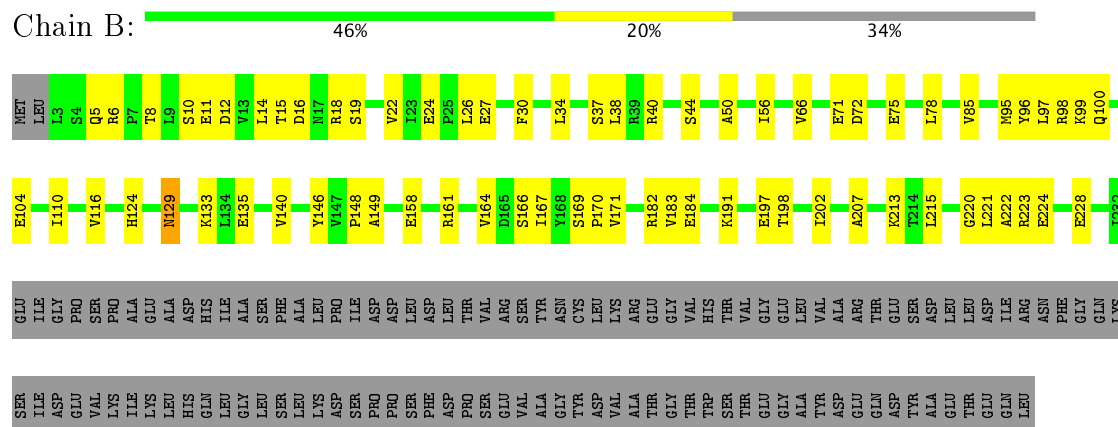
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

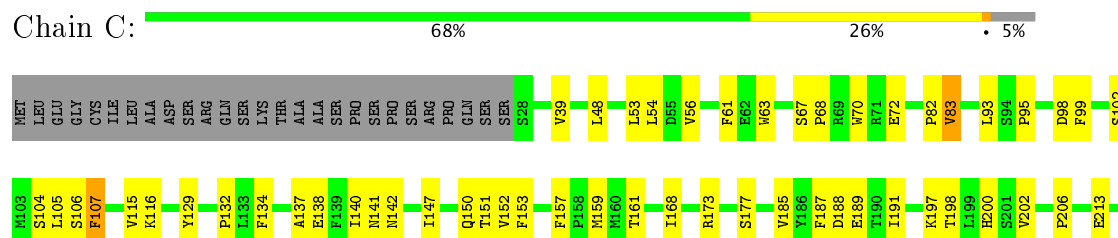
- Molecule 1: DNA-directed RNA polymerase subunit alpha



- Molecule 1: DNA-directed RNA polymerase subunit alpha



- Molecule 2: DNA-directed RNA polymerase subunit beta

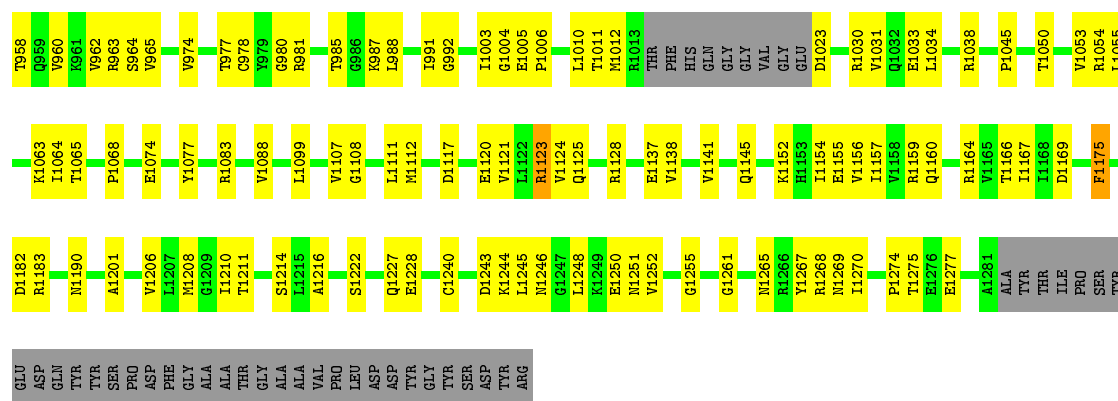


L985	V873	E767	T628	L496	E335	D217
L988	A874	E767	G629	L500	E336	K218
R956	G882	D772	B630	L502	L345	D220
L1131	D883	I773	L632	L507	L348	T221
L1133	K884	P774	D837	P508	T367	V222
E1136	G880	N775	A638	L512	F373	G223
V1137	N891	I776	V653	L513	F373	V224
E1145	K892	E786	L658	L514	R377	L226
GLU	V894	R791	M672	K518	L378	K229
LEU	L899	R797	I689	V519	I384	K230
ARG	D903	K805	I702	V523	P232	K231
GLU	P904	V806	A703	V524	Q388	Q233
GLU	P905	T807	D704	S525	I389	V234
ASP	D909	P808	C707	B526	M393	V236
ASP	G910	K809	T707	E527	V400	L237
LEU	V913	B810	T708	V529	R401	L238
GLU	D914	T815	D709	V540	M404	A240
ARG	I915	B818	D710	V541	V409	Q247
ALA	I916	R819	G711	P547	I412	L248
ALA	L917	L820	M713	V555	T413	E250
ASN	N918	L821	N718	R563	P414	R251
LEU	T919	R822	L719	I418	Q415	F252
GLY	H920	A823	L720	V571	L417	L257
ILE	G921	T824	V721	E575	L418	K258
ASN	V922	P825	G728	E578	F430	R259
LEU	R924	A829	Y731	Q585	F431	
SER	G925	R830	E732	N586	P445	
ARG	N926	R831	D733	S588	L449	
ASN	N927	R832	N739	V587	T450	
GLU	I928	R833	L741	S589	S464	
SER	G929	D834	E744	A592	R465	
ALA	Q930	T835	D745	L592	E466	
VAL	T934	K846	L746	R604	R467	
GLU	H935	T850	E748	I752	R473	
ASP	L936	R851	D749	E756	C484	
LEU	W944	R852	V749	I757	P612	
ALA	K1101	F853	L747	S854	R613	
V1102	V1102	R855	S749	R608	I486	
E900	I1106	L861	I752	M611	R310	
V1107	V1107	V865	E756	P612	V311	
K1108	K1108	N866	I757	R613	V311	
I974	N1111	E867	I758	A759	V316	
A99	I1112	R870	A759	R760		
P100	P1118	V871	R760	L764		
V101	E1119	N872	L764			
T102	L1125					
H103						
I104						

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

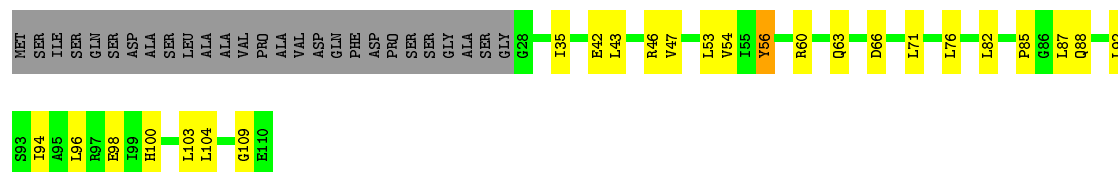
Chain D:  67% 28%

Q233	Q344	P454	L558	Q637	F831
L234	K345	F455	S561	Q637	I832
D237	I348	M456	S567	I697	P833
N239	N352	K458	L574	A701	P835
L244	K355	L460	A575	V709	P840
R247	K356	D462	M576	D714	Y849
Y248	L357	M463	P577	K715	T853
G249	L360	M464	L579	L716	H654
E250	P863	A466	N580	F721	R857
Y251	F252	M468	M581	Y722	Y872
F252	I366	L469	L588	Y723	T873
M256	K371	K470	T589	V729	L874
E259	Q375	A472	D595	T730	T875
S260	Q375	K473	E598	V731	R876
L264	V378	M474	A602	P739	L877
I265	L381	M475	L605	P740	V880
E275	I384	R480	H606	K741	V886
S276	G385	V483	V611	K742	H889
L277	R387	V486	V612	L746	Q892
R279	P394	S401	S613	R762	L897
V280	K400	D404	A616	L765	V893
I281	S401	L507	M620	E769	V899
R282	R421	E513	R624	W778	L901
K285	K420	L516	G625	L789	R904
G286	R421	Q523	V626	Y793	L910
K289	Y424	L524	V629	I799	A920
L290	S425	H525	R630	I800	Y921
K294	G426	P526	A631	D804	A922
K297	S428	L527	L637	T808	R923
V317	V429	E530	T638	L816	T924
P318	L430	F536	Q639	L817	L925
V319	L438	Q540	R641	A818	G926
I320	C441	M541	L642	G819	D931
R321	G442	V543	M662	M820	E932
P322	P444	L547	V663	K821	V938
E323	L446	L547	T666	L823	E939
R327	M447	E554	T667	M826	R940
M327	L451	R556	R670	P827	D946
L219	L451	I557	V671	E830	P947
E220			L676		L952
D221					
I222					
L229					
K232					



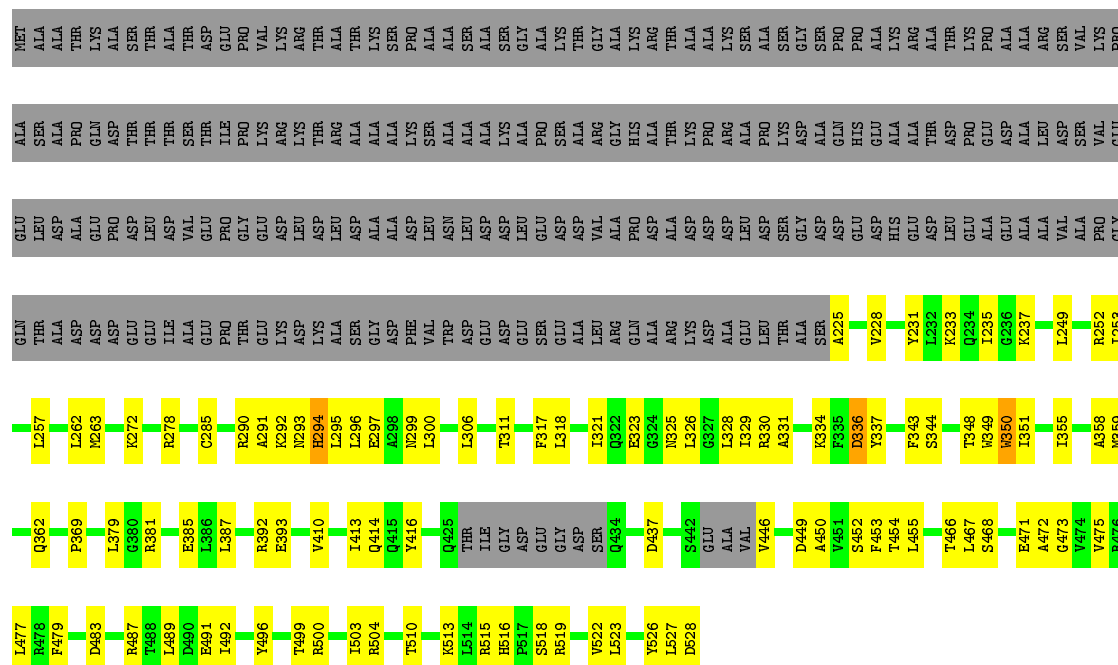
• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E: 53% 22% 25%



• Molecule 5: RNA polymerase sigma factor SigA

Chain F: 37% 18% 45%



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	68895	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	1.4	Depositor
Minimum defocus (nm)	-1.0	Depositor
Maximum defocus (nm)	-2.0	Depositor
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor



## 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: MG, ZN, FI8

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  >2	RMSZ	# Z  >2
1	A	0.37	0/1726	0.61	0/2348
1	B	0.33	0/1775	0.58	0/2414
2	C	0.37	0/8830	0.61	0/11972
3	D	0.34	0/10116	0.58	0/13673
4	E	0.33	0/667	0.58	0/908
5	F	0.27	0/2371	0.52	0/3193
All	All	0.35	0/25485	0.59	0/34508

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	1700	0	1734	46	0
1	B	1749	0	1781	52	0
2	C	8671	0	8610	239	0
3	D	9951	0	10024	315	0
4	E	654	0	648	20	0
5	F	2343	0	2377	87	0
6	D	2	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	D	1	0	0	0	0
8	D	72	0	0	7	0
9	D	3	0	0	0	0
All	All	25146	0	25174	700	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 14.

The worst 5 of 700 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)
3:D:282:ARG:HH21	3:D:282:ARG:CB	1.28	1.43
3:D:282:ARG:HB2	3:D:282:ARG:NH2	1.08	1.40
3:D:285:LYS:HA	3:D:285:LYS:HE3	1.29	1.10
3:D:282:ARG:NH2	3:D:282:ARG:CB	2.04	0.96
3:D:277:LEU:O	3:D:281:ILE:HG13	1.67	0.95

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 PDB entries followed by that with respect to all EM entries.

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	221/347 (64%)	203 (92%)	18 (8%)	0	100	100
1	B	228/347 (66%)	214 (94%)	14 (6%)	0	100	100
2	C	1116/1178 (95%)	1036 (93%)	76 (7%)	4 (0%)	38	77
3	D	1266/1316 (96%)	1196 (94%)	68 (5%)	2 (0%)	51	85
4	E	81/110 (74%)	76 (94%)	5 (6%)	0	100	100
5	F	287/528 (54%)	276 (96%)	11 (4%)	0	100	100
All	All	3199/3826 (84%)	3001 (94%)	192 (6%)	6 (0%)	54	85

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	835	PRO
2	C	82	PRO
2	C	831	GLU
3	D	394	PRO
2	C	1048	PRO

### 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 PDB entries followed by that with respect to all EM entries.

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	192/297 (65%)	190 (99%)	2 (1%)	80	91
1	B	197/297 (66%)	194 (98%)	3 (2%)	70	88
2	C	947/998 (95%)	926 (98%)	21 (2%)	57	83
3	D	1058/1095 (97%)	1032 (98%)	26 (2%)	53	81
4	E	70/90 (78%)	69 (99%)	1 (1%)	71	89
5	F	249/427 (58%)	240 (96%)	9 (4%)	40	74
All	All	2713/3204 (85%)	2651 (98%)	62 (2%)	59	82

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	84	ARG
3	D	345	ARG
5	F	350	TRP
3	D	275	GLU
3	D	420	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	247	GLN
5	F	299	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	C	1062	GLN
2	C	141	ASN
2	C	1035	HIS

### 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 carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 3 are monoatomic - leaving 1 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	FI8	D	1904	-	75,75,75	1.19	6 (8%)	89,109,109	1.60	13 (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	FI8	D	1904	-	-	0/75/118/118	0/3/4/4

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	D	1904	FI8	C44-C45	-2.95	1.46	1.53
8	D	1904	FI8	C1-C3	2.12	1.53	1.50
8	D	1904	FI8	C4-C3	2.81	1.44	1.36
8	D	1904	FI8	C5-C4	2.96	1.52	1.43
8	D	1904	FI8	C14-C15	3.50	1.53	1.46

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	D	1904	FI8	C11-C10-C9	-4.38	116.40	127.66
8	D	1904	FI8	C7-C6-C5	-3.68	120.31	125.40
8	D	1904	FI8	C1-C3-C4	-3.38	121.23	125.10
8	D	1904	FI8	C37-C38-C39	-3.05	119.96	122.55
8	D	1904	FI8	C31-C30-C29	-3.02	108.71	113.36

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	D	1904	FI8	7	0

## 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.