



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 10, 2017 – 01:30 PM EST

PDB ID : 5TW1
Title : Crystal structure of a Mycobacterium smegmatis transcription initiation complex with RbpA
Authors : Hubin, E.A.; Darst, S.A.; Campbell, E.A.
Deposited on : 2016-11-10
Resolution : 2.76 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

<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 : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20028442
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : rb-20028442

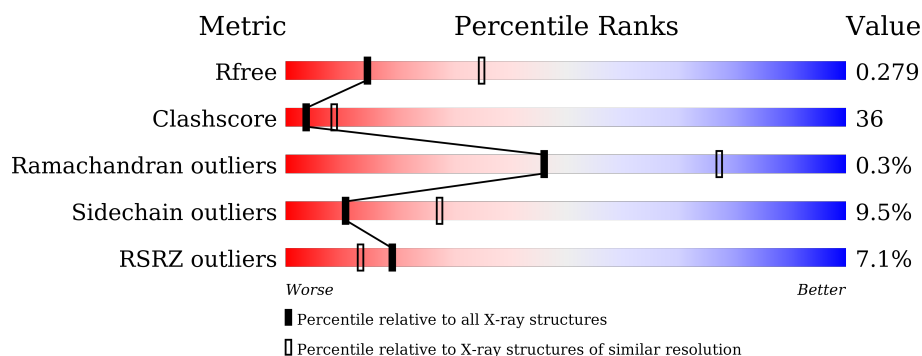
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.76 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}	91344	3340 (2.80-2.72)
Clashscore	102246	3829 (2.80-2.72)
Ramachandran outliers	100387	3767 (2.80-2.72)
Sidechain outliers	100360	3770 (2.80-2.72)
RSRZ outliers	91569	3352 (2.80-2.72)

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. 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	J	114	<div> <div>9%</div> <div>34%</div> <div>33%</div> <div>5%</div> <div>27%</div> </div>
2	A	350	<div> <div>%</div> <div>23%</div> <div>36%</div> <div>•</div> <div>38%</div> </div>
2	B	350	<div> <div>10%</div> <div>26%</div> <div>36%</div> <div>5%</div> <div>33%</div> </div>
2	T	350	<div> <div>13%</div> <div>9%</div> <div>5%</div> <div>•</div> <div>85%</div> </div>
3	C	1169	<div> <div>8%</div> <div>44%</div> <div>45%</div> <div>6%</div> <div>6%</div> </div>
4	D	1317	<div> <div>3%</div> <div>50%</div> <div>40%</div> <div>•</div> <div>5%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	107	
6	F	466	
7	O	31	
8	P	26	
9	G	17	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
10	SO4	C	1203	-	-	X	-
10	SO4	D	2005	-	-	X	-
10	SO4	F	505	-	-	-	X
12	ZN	D	2001	-	-	-	X

2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 26644 atoms, of which 36 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 RNA polymerase-binding protein RbpA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	J	83	Total	C	N	O	S	0	0	0
			667	419	118	128	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	218	Total	C	N	O	S	0	0	0
			1617	1020	276	318	3			
2	B	233	Total	C	N	O	S	0	0	0
			1667	1054	289	322	2			
2	T	53	Total	C	N	O	S	0	0	0
			374	236	65	72	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	1099	Total	C	N	O	S	0	0	0
			8250	5164	1448	1603	35			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	1248	Total	C	N	O	S	0	0	0
			9588	6016	1727	1805	40			

- Molecule 5 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	76	Total	C	N	O		0	0	0
			592	378	100	114				

- Molecule 6 is a protein called RNA polymerase sigma factor SigA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	302	Total	C	N	O	S	0	0	0
			2396	1502	433	454	7			

- Molecule 7 is a DNA chain called DNA (31-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	O	31	Total	C	N	O	P	0	0	0
			634	305	114	185	30			

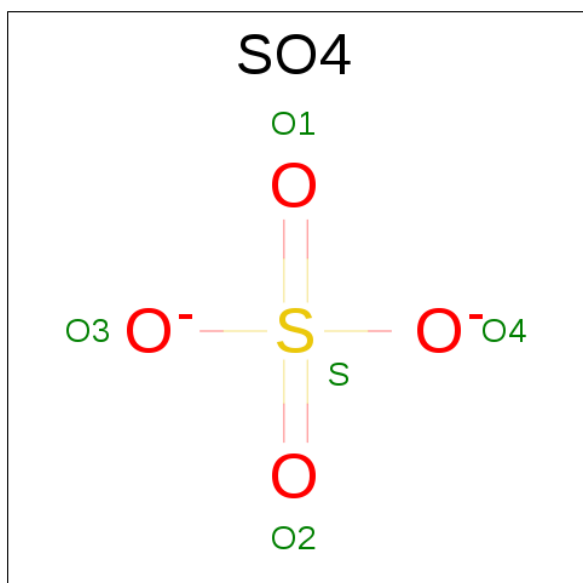
- Molecule 8 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	P	26	Total	C	N	O	P	0	0	0
			526	254	94	153	25			

- Molecule 9 is a protein called Unknown peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
9	G	17	Total	C	N	O	0	0	0
			85	51	17	17			

- Molecule 10 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



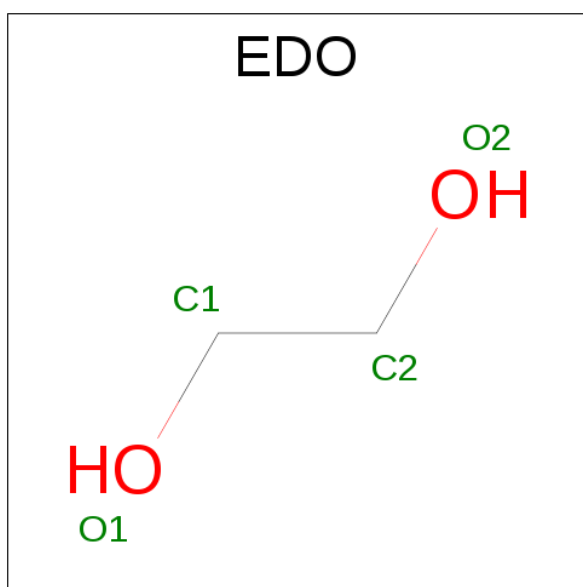
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	C	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	C	1	Total	O	S	0	0
			5	4	1		
10	C	1	Total	O	S	0	0
			5	4	1		
10	D	1	Total	O	S	0	0
			5	4	1		
10	D	1	Total	O	S	0	0
			5	4	1		
10	D	1	Total	O	S	0	0
			5	4	1		
10	D	1	Total	O	S	0	0
			5	4	1		
10	F	1	Total	O	S	0	0
			5	4	1		
10	F	1	Total	O	S	0	0
			5	4	1		
10	F	1	Total	O	S	0	0
			5	4	1		
10	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 11 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	C	1	Total C H O 10 2 6 2	0	0
11	C	1	Total C H O 10 2 6 2	0	0
11	D	1	Total C H O 10 2 6 2	0	0
11	D	1	Total C H O 10 2 6 2	0	0
11	D	1	Total C H O 10 2 6 2	0	0
11	F	1	Total C H O 10 2 6 2	0	0

- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	D	2	Total Zn 2 2	0	0

- Molecule 13 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	D	1	Total Mg 1 1	0	0

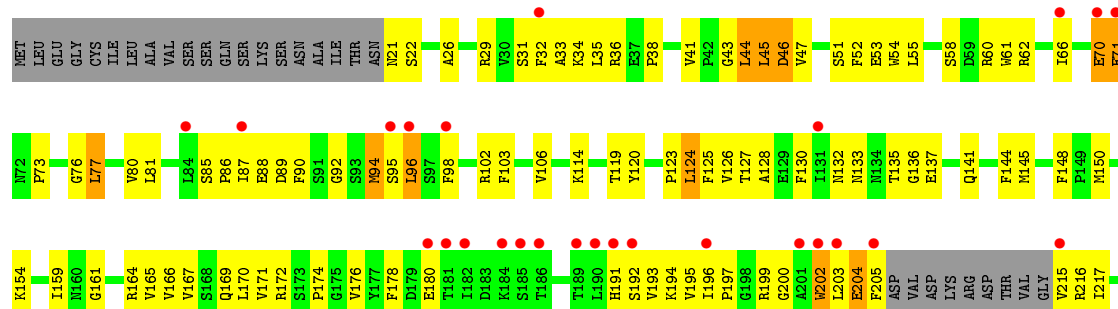
- Molecule 14 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	J	2	Total O 2 2	0	0
14	A	2	Total O 2 2	0	0
14	B	2	Total O 2 2	0	0
14	C	30	Total O 30 30	0	0
14	D	57	Total O 57 57	0	0
14	E	3	Total O 3 3	0	0
14	F	16	Total O 16 16	0	0

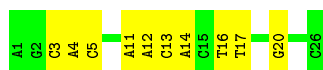
Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	O	10	Total	O	0	0
			10	10		
14	P	3	Total	O	0	0
			3	3		



D1244	S1171	ASP	GLY	V929	R769	P704	L627	L558	P454	Y344	Y251	R173	R88
N1247	S1172	GLY	VAL	D930	R770	P705	S628	P559	F455	R345	P252	R174	R89
G1248	S1173	GLU	THR	A931	S771	T706	W629	L560	V456	I348	T253	L177	E90
L1249	T1174	GLY	THR	G933	S772	T707	W630	S562	M457	I349	G256	E183	R91
K1250	E1175	GLY	GLY	N934	I776	V708	A631	S563	K459	L354	K257	E184	N92
E1251	F1176	ALA	ASP	V935	S777	V709	L633	L566	L460	I358	A258	A186	G93
N1252	L1177	ASP	ASP	V936	S778	Q710	R634	S567	Q467	I359	K262	E187	H94
V1253	P1178	ILE	VAL	V937	A780	T711	W635	P573	N468	A362	D269	G188	E96
I1254	G1179	GLY	VAL	R938	T781	D712	L637	L574	I469	P363	I270	K190	T102
K1257	L1181	V1106	G1027	L943	V784	K716	T638	A575	A472	E364	I271	K191	H103
L1258	T1182	D1110	G1028	G944	G785	D717	L640	N576	A473	I365	A272	D192	I104
P1260	E1183	Q1111	L1029	P945	T792	T722	E647	P577	L487	I366	K277	V193	W105
I1265	E1186	Q1112	V1032	R946	L788	A718	A648	P578	L488	I367	I278	E194	F107
E1266	F1187	L1112	Q1033	P947	T789	R719	E652	L579	E488	P364	E273	K195	F108
E1267	E1188	M1113	E1034	V948	T790	F720	W653	L588	E489	N368	E274	K196	F109
I1271	E1189	E1114	L1035	V949	T791	F721	N633	T589	V490	I368	E275	K197	K108
Q1272	E1190	G1115	F1036	L951	T792	T723	E649	P581	M373	I369	E276	V197	G109
P1275	M1191	A1116	E1037	L952	T793	A724	E650	G584	A492	E376	I279	G201	V110
R1276	R1192	A1117	A1038	L953	T794	R725	E651	L585	E493	P365	I281	E202	S112
P1277	R1193	D1118	R1039	A954	T795	S726	W652	L586	L504	R397	K285	R203	R113
V1278	V1194	P1119	V1040	G955	T796	T727	N634	T588	H505	K400	G286	R209	L114
A1279	V1195	P1041	P1042	T957	T797	T728	E654	T590	L507	D404	Q287	R214	L117
A1283	ALA	R1042	H1043	T958	T798	T729	W655	L591	E513	G411	K288	R219	L118
T1283	GLU	Q1044	K1045	V959	T799	S730	E656	L592	P514	R412	K289	R219	D119
T1284	GLY	A1045	A1046	V960	T800	F731	E657	G594	Q615	G412	K290	R219	L120
ILE	GLY	V1125	D1049	R962	T801	A732	E658	T596	L516	R413	K291	R219	L121
GLU	GLY	Q1127	V1054	G972	T802	T733	W659	G597	V517	R414	K292	R219	L122
PRO	PRO	E1127	R1055	G973	T803	T734	E660	G598	E518	R415	K293	R219	L123
ASP	P1206	E1128	L1056	C977	T804	T735	W661	T597	G519	R416	K294	R219	L124
GLN	V1207	E1129	E1057	R980	T805	T736	E662	T598	Q520	R417	K295	R219	L125
GLN	L1208	Q1130	E1058	S981	T806	T737	W663	T599	Q521	R418	K296	R219	L126
T1209	M1209	Q1131	T892	T982	T807	T738	E664	L601	A521	R419	K297	R219	L127
T1210	G1210	I1133	T893	T983	T808	T739	E665	A602	Q522	R420	K298	R219	L128
SER	I1211	Q1140	V898	L983	T809	T740	W666	Q603	Q523	R421	K299	R219	L129
PRO	E1141	E1141	T899	T984	T810	T741	E667	T604	Q524	R422	K300	R219	L130
ASP	T1212	E1142	L900	G985	T811	T742	W668	D605	H525	R423	K301	R219	L131
PHE	A1213	V1142	A901	V985	T812	T743	E669	A606	P526	R424	K302	R219	L132
GLY	S1215	Q1146	E902	V986	T813	T744	W670	P607	L527	R425	K303	R219	L133
GLN	S1220	I1150	T906	L987	T814	T745	E671	Q608	V528	R426	K304	R219	L134
ALA	M1221	H1151	GLY	G991	T815	T746	W672	Q609	F536	R427	K305	R219	L135
GLY	L1222	D1152	THR	G992	T816	T747	E673	G610	D537	R428	K306	R219	L136
ALA	S1223	K1153	LEU	E993	T817	T748	W674	V611	G538	R429	K307	R219	L137
VAL	E1229	H1154	I910	G994	T818	T749	E675	V612	P539	R430	K308	R219	L138
PRO	T1230	I1155	R911	P1005	T819	T750	W676	V613	D540	R431	K309	R219	L139
LEU	R1232	E1156	D912	P1006	T820	T751	E677	V614	M541	R432	K310	R219	L140
ASP	V1233	V1157	A913	L1009	T821	T752	W678	V615	A542	R433	K311	R219	L141
THR	L1234	R1087	V915	T1010	T822	T753	E679	V616	V543	R434	K312	R219	L142
GLY	T1235	R1088	A919	MET	T823	T754	W680	V617	Q544	R435	K313	R219	L143
T1236	R1164	R1089	R920	ARG	T824	T755	W681	V618	M545	R436	K314	R219	L144
T1237	R1165	I1090	T844	THR	T825	T756	W682	V619	A546	R437	K315	R219	L145
SER	D1238	THR	T845	PHE	T826	T757	W683	V620	V547	R438	K316	R219	L146
ASP	T1167	HIS	L946	HIS	T827	T758	W684	V621	F450	R439	K317	R219	L147
T1168	I1168	GLU	T924	GLN	T828	T759	W685	V622	F451	R440	K318	R219	L148
					T829	T760	W686	V623	F452	R441	K319	R219	L149
					T830	T761	W687	V624	F453	R442	K320	R219	L150
					T831	T762	W688	V625	F454	R443	K321	R219	L151
					T832	T763	W689	V626	F455	R444	K322	R219	L152
					T833	T764	W690	V627	F456	R445	K323	R219	L153
					T834	T765	W691	V628	F457	R446	K324	R219	L154
					T835	T766	W692	V629	F458	R447	K325	R219	L155
					T836	T767	W693	V630	F459	R448	K326	R219	L156
					T837	T768	W694	V631	F460	R449	K327	R219	L157
					T838	T769	W695	V632	F461	R450	K328	R219	L158
					T839	T770	W696	V633	F462	R451	K329	R219	L159
					T840	T771	W697	V634	F463	R452	K330	R219	L160
					T841	T772	W698	V635	F464	R453	K331	R219	L161
					T842	T773	W699	V636	F465	R454	K332	R219	L162
					T843	T774	W700	V637	F466	R455	K333	R219	L163
					T844	T775	W701	V638	F467	R456	K334	R219	L164
					T845	T776	W702	V639	F468	R457	K335	R219	L165
					T846	T777	W703	V640	F469	R458	K336	R219	L166
					T847	T778	W704	V641	F470	R459	K337	R219	L167
					T848	T779	W705	V642	F471	R460	K338	R219	L168
					T849	T780	W706	V643	F472	R461	K339	R219	L169
					T850	T781	W707	V644	F473	R462	K340	R219	L170



- Molecule 9: Unknown peptide

Chain G:  100%

There are no outlier residues recorded for this chain.

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	133.01Å 161.63Å 139.21Å 90.00° 107.72° 90.00°	Depositor
Resolution (Å)	54.91 – 2.76 54.91 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.5 (54.91-2.76) 99.4 (54.91-2.75)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.16 (at 2.77Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.239 , 0.280 0.236 , 0.279	Depositor DCC
R_{free} test set	1974 reflections (1.38%)	DCC
Wilson B-factor (Å ²)	78.4	Xtriage
Anisotropy	0.150	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 53.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.003 for l,-k,h	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	26644	wwPDB-VP
Average B, all atoms (Å ²)	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.46% 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: MG, EDO, ZN, SO4

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	J	0.27	0/681	0.47	0/923
2	A	0.27	0/1641	0.50	0/2236
2	B	0.26	0/1693	0.49	0/2316
2	T	0.24	0/376	0.39	0/511
3	C	0.28	0/8394	0.50	1/11410 (0.0%)
4	D	0.30	0/9742	0.49	0/13189
5	E	0.28	0/604	0.49	0/822
6	F	0.26	0/2426	0.45	1/3273 (0.0%)
7	O	0.68	0/710	0.96	0/1095
8	P	0.69	0/589	0.96	0/906
All	All	0.31	0/26856	0.53	2/36681 (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
1	J	0	1
3	C	0	5
All	All	0	6

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	890	LEU	CA-CB-CG	5.46	127.85	115.30
6	F	330	ARG	C-N-CA	-5.05	109.07	121.70

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	368	ARG	Peptide
3	C	433	GLN	Peptide
3	C	982	SER	Peptide
3	C	985	PRO	Peptide
1	J	70	PRO	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	J	667	0	649	51	0
2	A	1617	0	1636	179	0
2	B	1667	0	1636	193	0
2	T	374	0	344	29	0
3	C	8250	0	7989	692	0
4	D	9588	0	9552	631	0
5	E	592	0	583	53	0
6	F	2396	0	2422	142	0
7	O	634	0	350	33	0
8	P	526	0	296	16	0
9	G	85	0	19	0	0
10	C	15	0	0	2	0
10	D	20	0	0	2	0
10	F	25	0	0	1	0
11	C	8	12	12	1	0
11	D	12	18	18	3	0
11	F	4	6	6	1	0
12	D	2	0	0	0	0
13	D	1	0	0	0	0
14	A	2	0	0	0	0
14	B	2	0	0	0	0
14	C	30	0	0	16	0
14	D	57	0	0	17	0
14	E	3	0	0	0	0
14	F	16	0	0	3	0
14	J	2	0	0	0	0
14	O	10	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
14	P	3	0	0	1	0
All	All	26608	36	25512	1851	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 36.

The worst 5 of 1851 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:B:184:GLU:HA	2:B:187:THR:HG22	1.28	1.15
3:C:940:ALA:HB1	3:C:941:ALA:HA	1.28	1.12
3:C:228:LEU:HD21	3:C:268:ILE:HG12	1.29	1.11
3:C:982:SER:HB3	3:C:983:THR:HG23	1.35	1.08
3:C:176:VAL:HG12	3:C:195:VAL:HG22	1.35	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	J	81/114 (71%)	75 (93%)	6 (7%)	0	100	100
2	A	214/350 (61%)	202 (94%)	12 (6%)	0	100	100
2	B	231/350 (66%)	216 (94%)	14 (6%)	1 (0%)	39	72
2	T	51/350 (15%)	51 (100%)	0	0	100	100
3	C	1093/1169 (94%)	1038 (95%)	51 (5%)	4 (0%)	39	72
4	D	1238/1317 (94%)	1189 (96%)	46 (4%)	3 (0%)	52	83
5	E	72/107 (67%)	65 (90%)	5 (7%)	2 (3%)	6	18
6	F	298/466 (64%)	294 (99%)	4 (1%)	0	100	100
All	All	3278/4223 (78%)	3130 (96%)	138 (4%)	10 (0%)	46	77

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	E	76	VAL
3	C	982	SER
3	C	1134	ALA
4	D	902	GLU
4	D	1194	VAL

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	J	71/98 (72%)	65 (92%)	6 (8%)	13	34
2	A	178/297 (60%)	162 (91%)	16 (9%)	12	31
2	B	171/297 (58%)	152 (89%)	19 (11%)	8	20
2	T	35/297 (12%)	29 (83%)	6 (17%)	2	6
3	C	857/984 (87%)	777 (91%)	80 (9%)	11	29
4	D	994/1095 (91%)	904 (91%)	90 (9%)	12	30
5	E	62/86 (72%)	50 (81%)	12 (19%)	2	4
6	F	252/379 (66%)	232 (92%)	20 (8%)	15	37
All	All	2620/3533 (74%)	2371 (90%)	249 (10%)	11	27

5 of 249 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	1028	VAL
4	D	239	VAL
6	F	325	LEU
3	C	1054	PHE
4	D	70	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	1068	GLN
4	D	416	ASN
5	E	66	ASN
3	C	1120	GLN
4	D	287	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 21 ligands modelled in this entry, 3 are monoatomic - leaving 18 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	$\# Z > 2$
10	SO4	C	1201	-	4,4,4	0.26	0	6,6,6	0.12	0
10	SO4	C	1202	-	4,4,4	0.24	0	6,6,6	0.13	0
10	SO4	C	1203	-	4,4,4	0.24	0	6,6,6	0.28	0
11	EDO	C	1204	-	3,3,3	0.42	0	2,2,2	0.48	0
11	EDO	C	1205	-	3,3,3	0.50	0	2,2,2	0.28	0
10	SO4	D	2004	-	4,4,4	0.26	0	6,6,6	0.32	0
10	SO4	D	2005	-	4,4,4	0.24	0	6,6,6	0.25	0
10	SO4	D	2006	-	4,4,4	0.24	0	6,6,6	0.25	0
10	SO4	D	2007	-	4,4,4	0.23	0	6,6,6	0.13	0
11	EDO	D	2008	-	3,3,3	0.47	0	2,2,2	0.31	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
11	EDO	D	2009	-	3,3,3	0.44	0	2,2,2	0.36	0
11	EDO	D	2010	-	3,3,3	0.49	0	2,2,2	0.18	0
10	SO4	F	501	6	4,4,4	0.21	0	6,6,6	0.11	0
10	SO4	F	502	-	4,4,4	0.22	0	6,6,6	0.17	0
10	SO4	F	503	-	4,4,4	0.19	0	6,6,6	0.14	0
10	SO4	F	504	-	4,4,4	0.23	0	6,6,6	0.09	0
10	SO4	F	505	-	4,4,4	0.22	0	6,6,6	0.13	0
11	EDO	F	506	-	3,3,3	0.48	0	2,2,2	0.30	0

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
10	SO4	C	1201	-	-	0/0/0/0	0/0/0/0
10	SO4	C	1202	-	-	0/0/0/0	0/0/0/0
10	SO4	C	1203	-	-	0/0/0/0	0/0/0/0
11	EDO	C	1204	-	-	0/1/1/1	0/0/0/0
11	EDO	C	1205	-	-	0/1/1/1	0/0/0/0
10	SO4	D	2004	-	-	0/0/0/0	0/0/0/0
10	SO4	D	2005	-	-	0/0/0/0	0/0/0/0
10	SO4	D	2006	-	-	0/0/0/0	0/0/0/0
10	SO4	D	2007	-	-	0/0/0/0	0/0/0/0
11	EDO	D	2008	-	-	0/1/1/1	0/0/0/0
11	EDO	D	2009	-	-	0/1/1/1	0/0/0/0
11	EDO	D	2010	-	-	0/1/1/1	0/0/0/0
10	SO4	F	501	6	-	0/0/0/0	0/0/0/0
10	SO4	F	502	-	-	0/0/0/0	0/0/0/0
10	SO4	F	503	-	-	0/0/0/0	0/0/0/0
10	SO4	F	504	-	-	0/0/0/0	0/0/0/0
10	SO4	F	505	-	-	0/0/0/0	0/0/0/0
11	EDO	F	506	-	-	0/1/1/1	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	C	1203	SO4	2	0
11	C	1205	EDO	1	0
10	D	2005	SO4	2	0
11	D	2008	EDO	3	0
10	F	502	SO4	1	0
11	F	506	EDO	1	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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	J	83/114 (72%)	0.43	10 (12%) 6 4	64, 97, 141, 158	0
2	A	218/350 (62%)	0.18	5 (2%) 64 57	59, 86, 118, 137	0
2	B	233/350 (66%)	0.87	36 (15%) 3 2	80, 110, 134, 148	0
2	T	53/350 (15%)	3.60	44 (83%) 0 0	113, 149, 165, 169	0
3	C	1099/1169 (94%)	0.60	98 (8%) 12 8	42, 85, 147, 163	0
4	D	1248/1317 (94%)	0.33	38 (3%) 54 47	39, 73, 125, 156	0
5	E	76/107 (71%)	0.35	4 (5%) 30 23	51, 78, 113, 117	0
6	F	302/466 (64%)	0.06	4 (1%) 79 75	38, 72, 122, 137	0
7	O	31/31 (100%)	-0.68	0 100 100	51, 64, 83, 89	0
8	P	26/26 (100%)	-0.78	0 100 100	59, 68, 83, 93	0
9	G	0/17	-	-	-	-
All	All	3369/4297 (78%)	0.46	239 (7%) 19 13	38, 82, 141, 169	0

The worst 5 of 239 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	228	LEU	8.6
2	T	275	LEU	8.0
2	T	269	VAL	7.3
3	C	215	VAL	6.5
2	T	256	LEU	6.4

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 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	LLDF	B-factors(Å ²)	Q<0.9
10	SO4	F	505	5/5	0.84	0.43	8.69	80,84,117,167	0
12	ZN	D	2001	1/1	0.99	0.22	2.97	66,66,66,66	0
11	EDO	D	2010	4/4	0.87	0.28	1.86	62,75,77,89	0
12	ZN	D	2002	1/1	0.79	0.33	1.80	222,222,222,222	0
11	EDO	C	1204	4/4	0.90	0.23	1.22	53,63,75,77	0
10	SO4	D	2006	5/5	0.88	0.21	0.85	80,103,121,126	0
11	EDO	D	2008	4/4	0.79	0.21	0.36	74,89,105,105	0
11	EDO	F	506	4/4	0.87	0.21	0.32	73,88,95,95	0
11	EDO	C	1205	4/4	0.92	0.19	0.02	74,93,101,112	0
10	SO4	C	1202	5/5	0.84	0.19	-0.13	105,111,115,178	0
10	SO4	D	2005	5/5	0.94	0.17	-0.64	80,85,97,125	0
11	EDO	D	2009	4/4	0.87	0.15	-0.77	71,86,103,103	0
10	SO4	D	2007	5/5	0.92	0.13	-1.12	93,98,113,141	0
10	SO4	D	2004	5/5	0.96	0.16	-	69,72,83,87	0
10	SO4	C	1201	5/5	0.90	0.17	-	111,114,124,138	0
10	SO4	F	501	5/5	0.97	0.06	-	97,98,116,122	0
10	SO4	C	1203	5/5	0.75	0.27	-	103,113,139,197	0
10	SO4	F	503	5/5	0.95	0.08	-	85,92,98,113	0
10	SO4	F	502	5/5	0.95	0.11	-	84,86,102,108	0
13	MG	D	2003	1/1	0.94	0.28	-	82,82,82,82	0
10	SO4	F	504	5/5	0.90	0.24	-	112,114,117,128	0

6.5 Other polymers ⓘ

There are no such residues in this entry.