wwPDB EM Validation Summary Report

Aug 11, 2024 – 12:31 AM JST

PDB ID	:	8K60
EMDB ID	:	EMD-36914
Title	:	Cryo-EM structure of Streptomyces coelicolor transcription initiation complex with the global transcription factor AfsR
Authors	:	Lin, W.; Shi, J.
Deposited on	:	2023-07-24
Resolution	:	3.40 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
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:

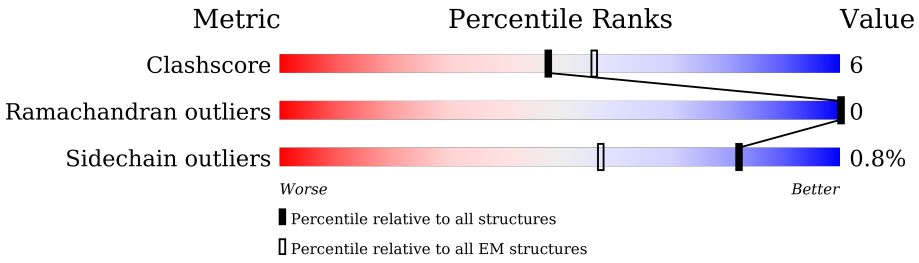
EMDB validation analysis	:	FAILED
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.40 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	340	57% 10% 34%
1	B	340	61% 8% 31%
1	K	340	13% · 83%
2	C	1161	82% 14% ·
3	D	1299	85% 12% ·
4	E	90	78% 8% 14%
5	F	511	53% 8% 38%
6	G	59	56% 39% 5%
7	H	59	69% 27% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	I	993	
8	J	993	

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 31659 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	226	Total	C	N	O	S	0	0
			1742	1102	302	334	4		
1	B	233	Total	C	N	O	S	0	0
			1785	1125	308	347	5		
1	K	57	Total	C	N	O	S	0	0
			438	274	77	85	2		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	1116	Total	C	N	O	S	0	0
			8692	5450	1513	1698	31		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	1259	Total	C	N	O	S	0	0
			9812	6137	1778	1856	41		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
4	E	77	Total	C	N	O	0	0
			594	379	98	117		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	315	Total	C	N	O	S	0	0
			2465	1546	437	475	7		

- Molecule 6 is a DNA chain called Non-template strand DNA for AfsS promoter.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	G	56	Total	C	N	O	P	0	0
			1148	546	204	342	56		

- Molecule 7 is a DNA chain called Template strand DNA for AfsS promoter.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	H	57	Total	C	N	O	P	0	0
			1180	557	232	334	57		

- Molecule 8 is a protein called Regulatory protein AfsR.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	I	250	Total	C	N	O	S	0	0
			1904	1184	359	357	4		
8	J	249	Total	C	N	O	S	0	0
			1896	1178	358	356	4		

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

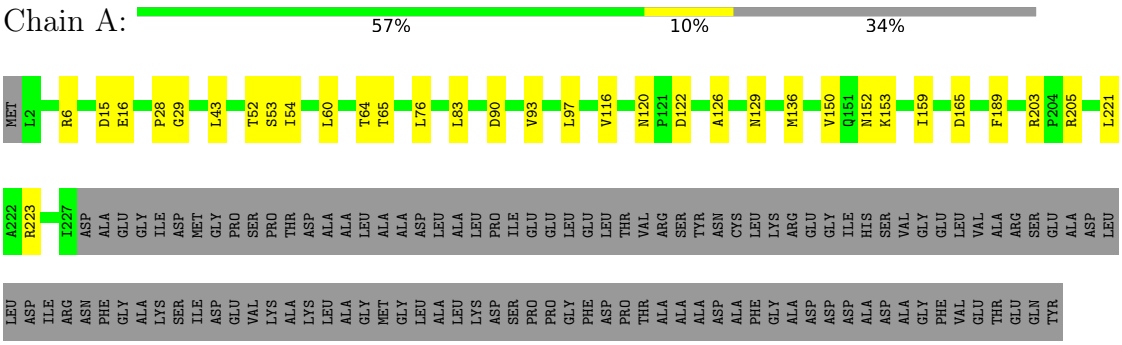
Mol	Chain	Residues	Atoms		AltConf
10	D	2	Total	Zn	0
			2	2	

3 Residue-property plots

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. 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.

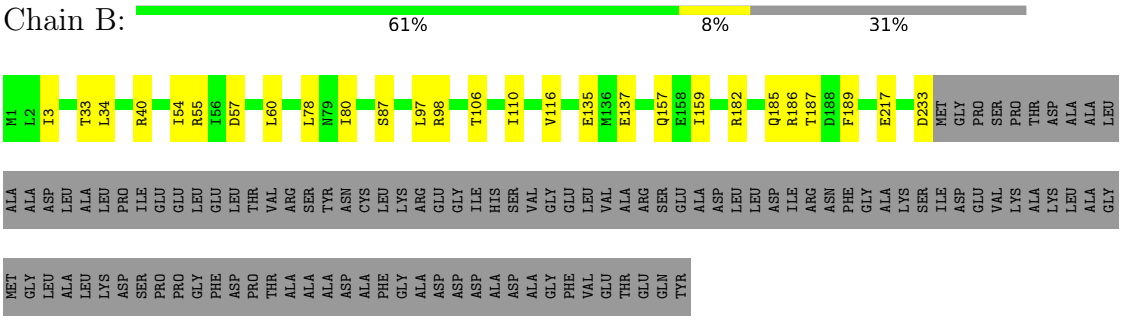
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:



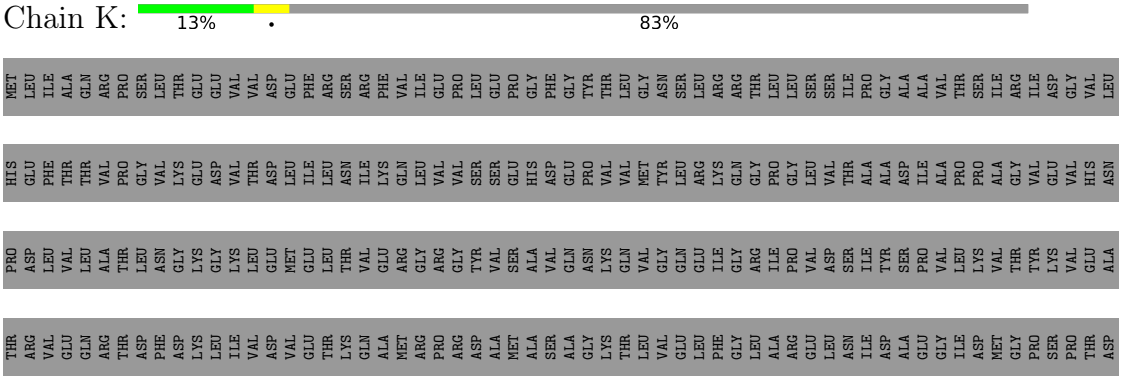
• Molecule 1: DNA-directed RNA polymerase subunit alpha

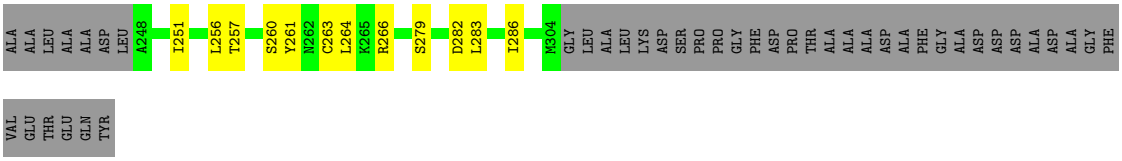
Chain B:



• Molecule 1: DNA-directed RNA polymerase subunit alpha

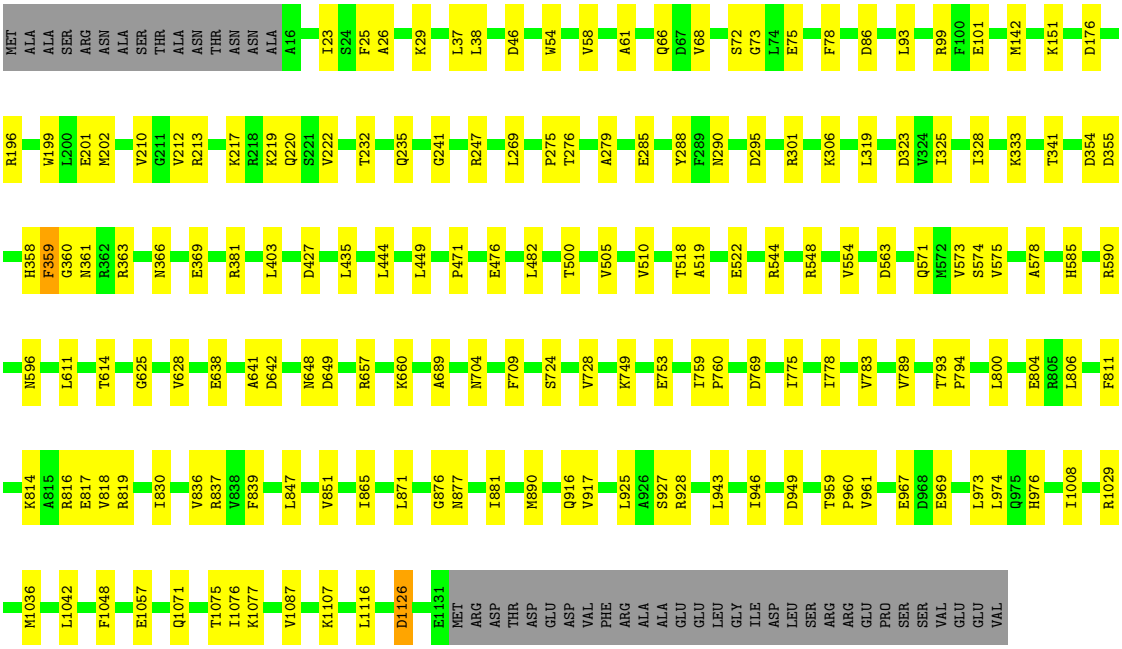
Chain K:





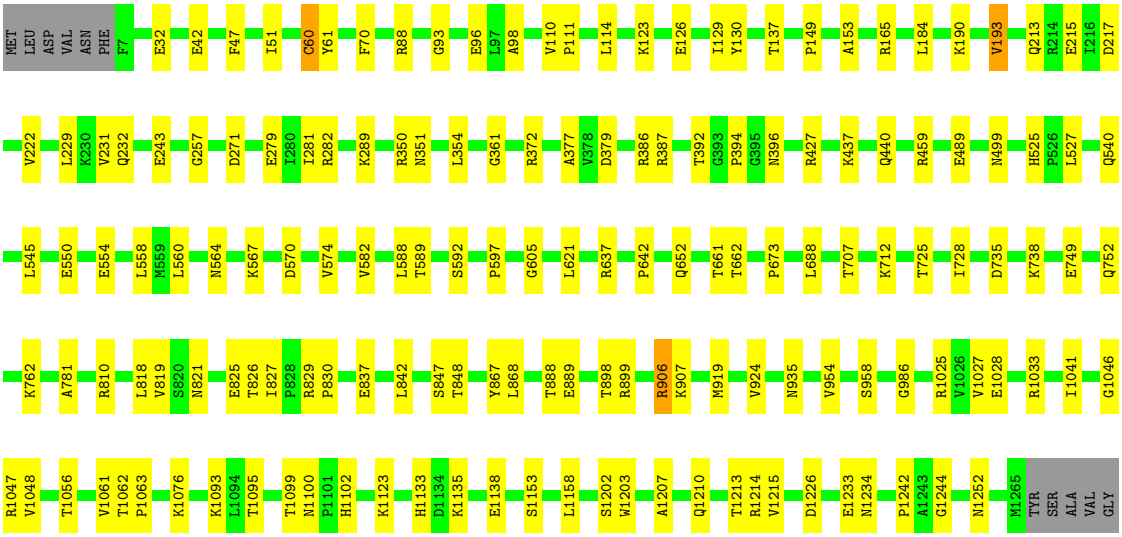
• Molecule 2: DNA-directed RNA polymerase subunit beta

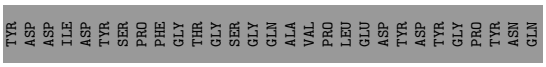
Chain C:



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

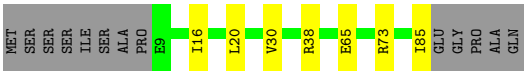
Chain D:





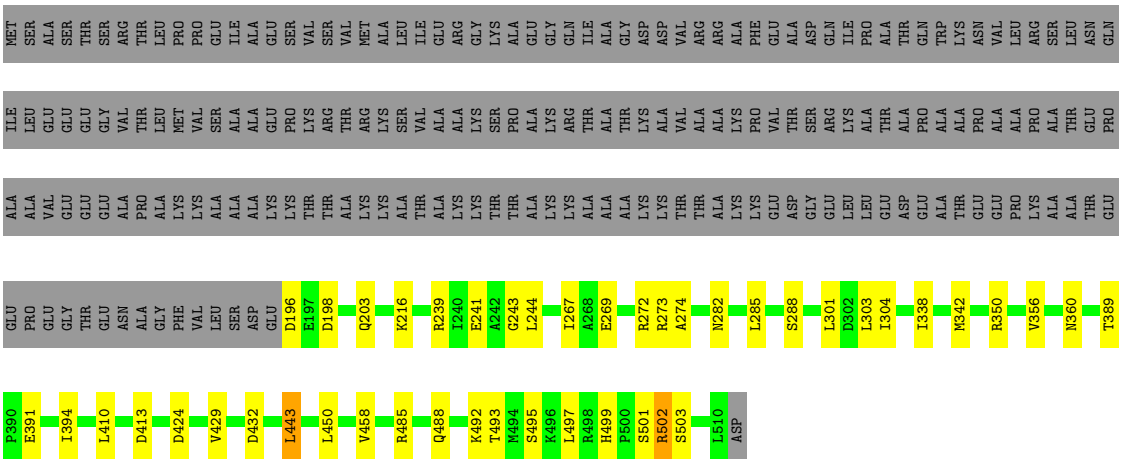
• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E:



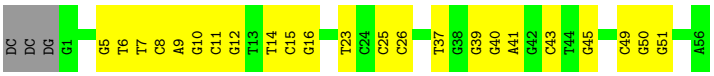
• Molecule 5: RNA polymerase principal sigma factor HrdB

Chain F:



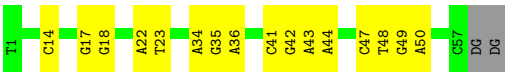
• Molecule 6: Non-template strand DNA for AfsS promoter

Chain G:



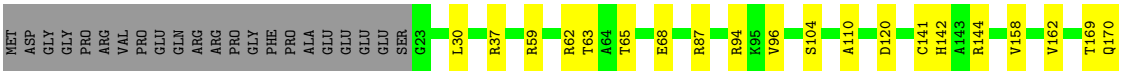
• Molecule 7: Template strand DNA for AfsS promoter

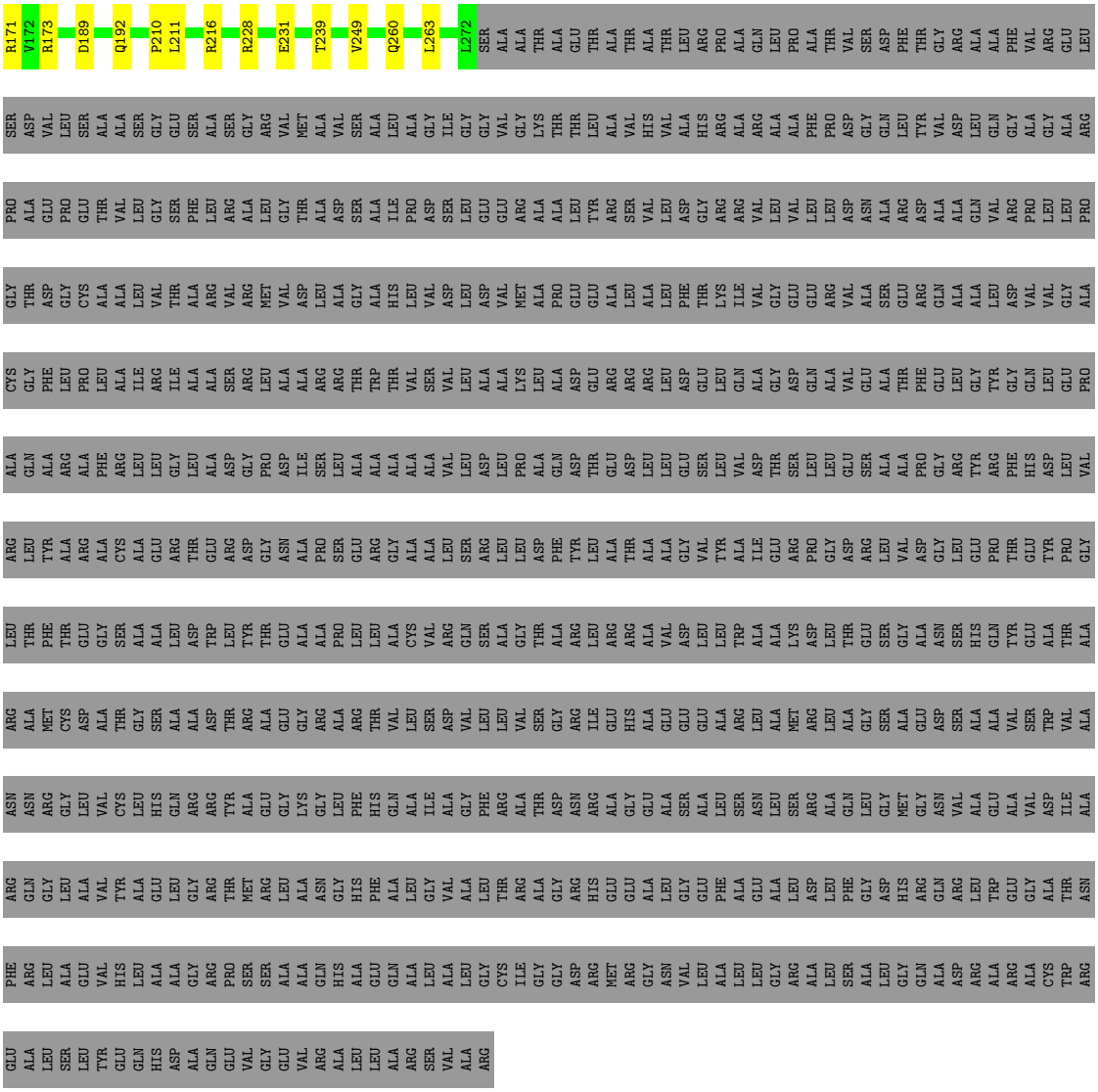
Chain H:



• Molecule 8: Regulatory protein AfsR

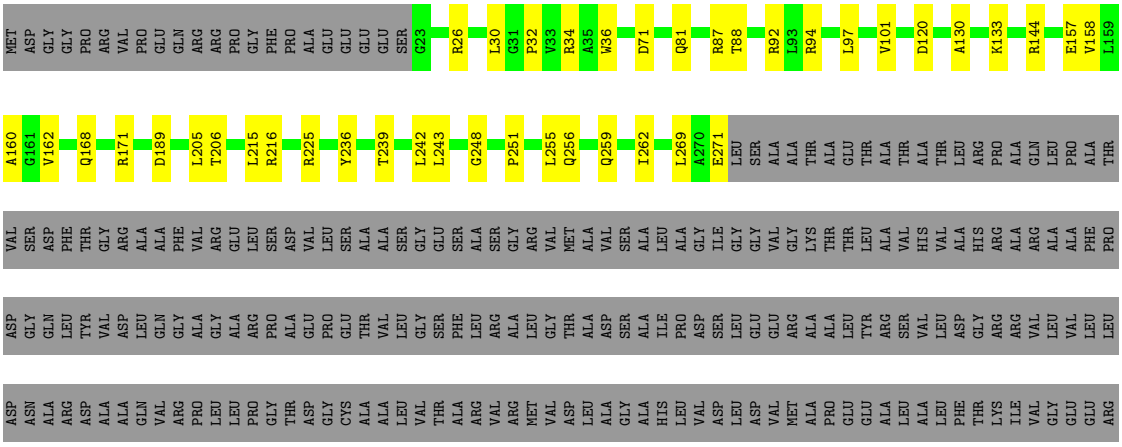
Chain I:

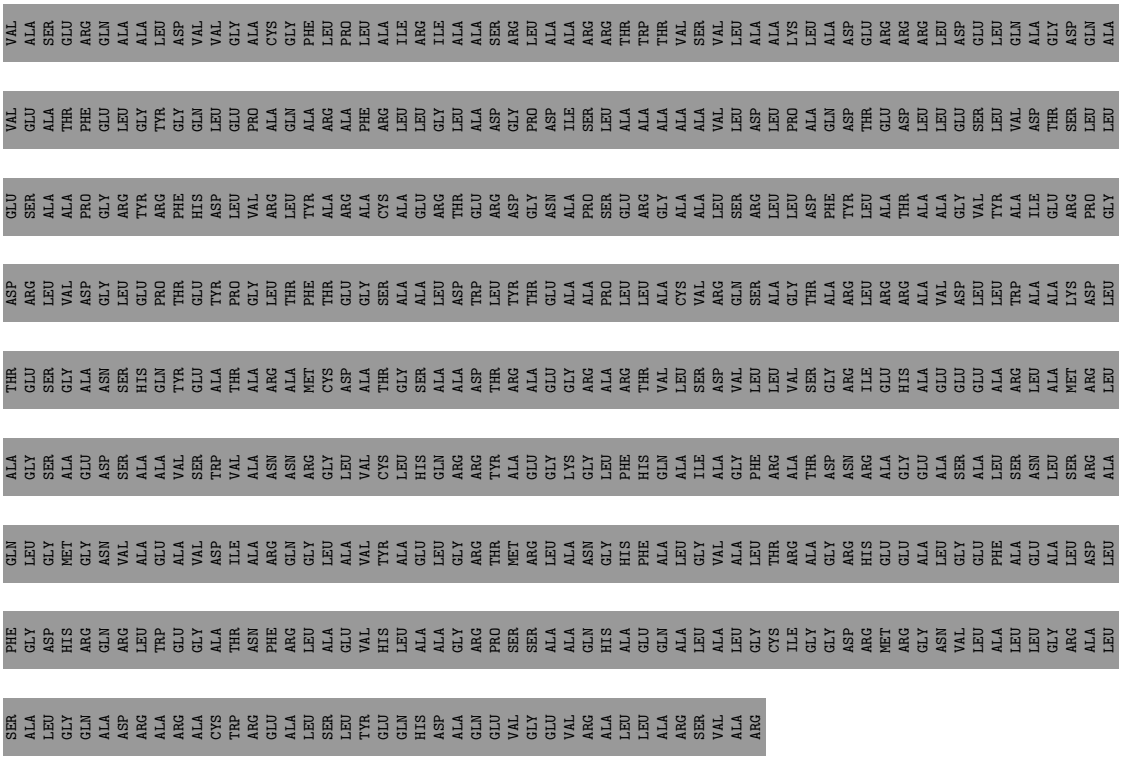




• Molecule 8: Regulatory protein AfsR

Chain J:




VAL ALA SER GLU ARG GLN ALA ALA LEU ASP VAL VAL GLY ALA CYS GLY PHE LEU PRO LEU ALA ILE ARG ILE ALA ALA SER ARG LEU ALA ALA ARG ARG THR TRP THR VAL SER VAL LEU ALA ALA LYS LEU ALA ASP GLU ARG ARG ARG LEU ASP GLU LEU GLN ALA GLY ASP GLN ALA
VAL GLU ALA THR PHE GLU LEU GLY TYR GLY GLN LEU GLU PRO ALA GLN ALA ARG ALA PHE ARG LEU LEU GLY LEU ALA ASP GLY PRO ASP ILE SER LEU ALA ALA ALA ALA VAL LEU ASP LEU PRO ALA GLN ASP THR GLU ASP LEU LEU GLU SER LEU VAL ASP THR SER LEU LEU
GLU SER ALA ALA PRO GLY ARG TYR ARG PHE HIS ASP LEU VAL ARG LEU TYR ALA ARG ALA CYS ALA GLU GLU ARG THR GLU ARG ASP GLY ASN ALA PRO SER GLU ARG GLY ALA ALA LEU SER ARG LEU LEU ASP PHE TYR LEU ALA THR ALA ALA GLY VAL TYR ALA ILE GLU ARG PRO GLY
ASP ARG LEU VAL ASP GLY LEU GLU PRO THR GLU TYR PRO GLY LEU THR PHE THR GLU GLY SER ALA ALA LEU ASP TRP LEU TYR THR GLU ALA ALA PRO LEU LEU ALA CYS VAL ARG GLN SER ALA GLY THR ALA ARG LEU ARG ARG ALA VAL ASP LEU LEU TRP ALA ALA LYS ASP LEU
THR GLU SER GLY ALA ASN SER HIS GLN TYR GLU ALA THR ALA ARG ALA MET CYS ASP ALA THR GLY SER ALA ALA ASP THR ARG ALA GLU GLY GLY ARG ALA ALA ARG THR VAL LEU SER ASP VAL LEU LEU VAL SER GLY ARG ILE GLU HIS ALA GLU GLU GLU ALA ARG LEU ALA MET ARG LEU
ALA GLY SER ALA GLU ASP SER ALA ALA VAL ALA ASN ASN ARG GLY LEU VAL CYS LEU HIS GLN ARG ARG TYR ALA GLU GLY LYS GLY LEU PHE HIS GLN ALA ILE ALA GLY PHE ARG ALA THR ASP ASN ARG ALA GLY GLU ALA SER ALA LEU SER ASN LEU SER ARG ALA
GLN LEU GLY MET GLY ASN VAL ALA GLU ALA VAL ASP ILE ALA ARG GLN GLY LEU ALA VAL TYR ALA GLU GLY ARG THR MET ARG LEU ALA ASN GLY HIS PHE ALA LEU GLY VAL ALA LEU THR ARG ALA ALA GLY ARG HIS GLU GLU ALA LEU GLY GLU PHE ALA GLU GLU ALA LEU ASP LEU
PHE GLY ASP HIS ARG GLN ARG LEU TRP GLU GLY ALA THR ASN PHE ARG LEU ALA GLU VAL HIS LEU ALA ALA GLY ARG PRO SER SER ALA ALA GLN HIS ALA GLU GLN ALA LEU ALA LEU GLY CYS ILE GLY GLY ASP ARG MET ARG GLY ASN VAL LEU ALA LEU LEU GLY ARG ALA LEU
SER ALA LEU GLY GLN ALA ASP ARG ALA ALA CYS TRP ARG GLU ALA LEU SER LEU TYR GLU GLN HIS ASP ALA GLN GLU VAL GLY GLU VAL ARG ALA LEU LEU ALA ARG SER VAL ALA ARG

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	231221	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 ($e^-/Å^2$)	52	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

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

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	0.44	0/1768	0.68	0/2400
1	B	0.39	0/1811	0.66	0/2459
1	K	0.31	0/441	0.56	0/591
2	C	0.47	0/8848	0.63	0/11982
3	D	0.45	0/9963	0.64	0/13446
4	E	0.47	0/604	0.65	0/822
5	F	0.36	0/2500	0.59	0/3377
6	G	0.73	0/1285	1.01	0/1982
7	H	0.83	0/1328	0.98	0/2049
8	I	0.34	0/1931	0.56	0/2620
8	J	0.29	0/1923	0.58	0/2609
All	All	0.47	0/32402	0.67	0/44337

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

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	1742	0	1796	19	0
1	B	1785	0	1821	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	K	438	0	452	6	0
2	C	8692	0	8632	107	0
3	D	9812	0	9951	90	0
4	E	594	0	595	5	0
5	F	2465	0	2487	33	0
6	G	1148	0	633	62	0
7	H	1180	0	637	36	0
8	I	1904	0	1933	21	0
8	J	1896	0	1922	28	0
9	D	1	0	0	0	0
10	D	2	0	0	0	0
All	All	31659	0	30859	382	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 6.

The worst 5 of 382 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)
6:G:8:DC:H2"	6:G:9:DA:H5'	1.22	1.08
6:G:8:DC:H2"	6:G:9:DA:C5'	2.00	0.91
7:H:47:DC:C6	7:H:48:DT:H71	2.07	0.89
6:G:8:DC:C2'	6:G:9:DA:H5'	2.01	0.88
7:H:49:DG:C2	7:H:50:DA:C4	2.64	0.86

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	224/340 (66%)	218 (97%)	6 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	231/340 (68%)	228 (99%)	3 (1%)	0	100	100
1	K	55/340 (16%)	55 (100%)	0	0	100	100
2	C	1114/1161 (96%)	1080 (97%)	34 (3%)	0	100	100
3	D	1257/1299 (97%)	1221 (97%)	36 (3%)	0	100	100
4	E	75/90 (83%)	74 (99%)	1 (1%)	0	100	100
5	F	313/511 (61%)	307 (98%)	6 (2%)	0	100	100
8	I	248/993 (25%)	245 (99%)	3 (1%)	0	100	100
8	J	247/993 (25%)	243 (98%)	4 (2%)	0	100	100
All	All	3764/6067 (62%)	3671 (98%)	93 (2%)	0	100	100

There are no Ramachandran outliers to report.

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 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	194/279 (70%)	192 (99%)	2 (1%)	76	88
1	B	197/279 (71%)	195 (99%)	2 (1%)	76	88
1	K	47/279 (17%)	47 (100%)	0	100	100
2	C	942/979 (96%)	934 (99%)	8 (1%)	81	91
3	D	1047/1087 (96%)	1037 (99%)	10 (1%)	76	88
4	E	63/74 (85%)	63 (100%)	0	100	100
5	F	261/413 (63%)	258 (99%)	3 (1%)	73	86
8	I	187/722 (26%)	186 (100%)	1 (0%)	88	94
8	J	186/722 (26%)	186 (100%)	0	100	100
All	All	3124/4834 (65%)	3098 (99%)	26 (1%)	82	91

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	137	THR
3	D	637	ARG
5	F	502	ARG
3	D	392	THR
3	D	762	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	752	GLN
8	J	192	GLN
3	D	935	ASN
1	K	270	HIS
5	F	371	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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.