



wwPDB X-ray Structure Validation Summary Report

Mar 24, 2022 – 12:58 pm GMT

PDB ID : 5LNH
Title : Structure of full length Unliganded CodY from Bacillus subtilis
Authors : Wilkinson, A.J.; Levdikov, V.M.; Blagova, E.V.
Deposited on : 2016-08-04
Resolution : 3.00 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.27
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.27

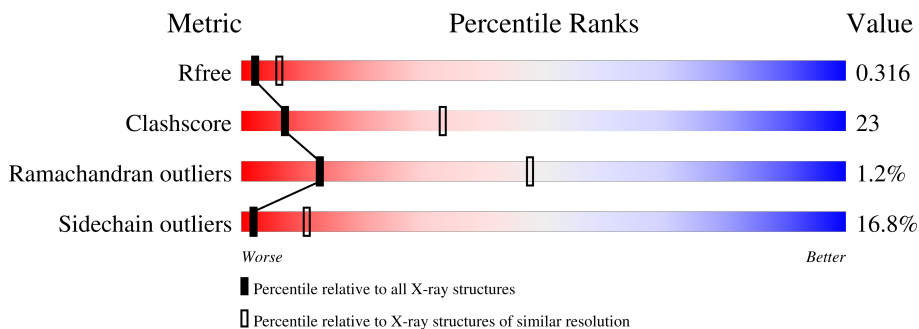
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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.




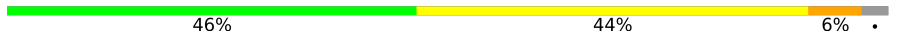

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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 ≥ 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	263	52% 36% 9% .
1	B	263	48% 39% 9% . .
1	C	263	54% 36% 7% . .
1	D	263	51% 36% 9% . .
1	E	263	53% 36% 8% .
1	F	263	58% 30% 9% .
1	G	263	51% 37% 10% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	263	 46% 41% 10% .
1	I	263	 46% 44% 6% .
1	K	263	 56% 36% 5% .

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 20090 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 GTP-sensing transcriptional pleiotropic repressor CodY.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	255	2002	1259	341	395	7	0	0	0
1	B	255	2002	1259	341	395	7	0	0	0
1	C	255	2002	1259	341	395	7	0	0	0
1	D	255	2002	1259	341	395	7	0	0	0
1	E	255	2002	1259	341	395	7	0	0	0
1	F	255	2002	1259	341	395	7	0	0	0
1	G	255	2002	1259	341	395	7	0	0	0
1	H	255	2002	1259	341	395	7	0	0	0
1	I	255	2002	1259	341	395	7	0	0	0
1	K	255	2002	1259	341	395	7	0	0	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	260	HIS	-	expression tag	UNP P39779
A	261	HIS	-	expression tag	UNP P39779
A	262	HIS	-	expression tag	UNP P39779
A	263	HIS	-	expression tag	UNP P39779
A	264	HIS	-	expression tag	UNP P39779
B	260	HIS	-	expression tag	UNP P39779
B	261	HIS	-	expression tag	UNP P39779
B	262	HIS	-	expression tag	UNP P39779
B	263	HIS	-	expression tag	UNP P39779

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	264	HIS	-	expression tag	UNP P39779
C	260	HIS	-	expression tag	UNP P39779
C	261	HIS	-	expression tag	UNP P39779
C	262	HIS	-	expression tag	UNP P39779
C	263	HIS	-	expression tag	UNP P39779
C	264	HIS	-	expression tag	UNP P39779
D	260	HIS	-	expression tag	UNP P39779
D	261	HIS	-	expression tag	UNP P39779
D	262	HIS	-	expression tag	UNP P39779
D	263	HIS	-	expression tag	UNP P39779
D	264	HIS	-	expression tag	UNP P39779
E	260	HIS	-	expression tag	UNP P39779
E	261	HIS	-	expression tag	UNP P39779
E	262	HIS	-	expression tag	UNP P39779
E	263	HIS	-	expression tag	UNP P39779
E	264	HIS	-	expression tag	UNP P39779
F	260	HIS	-	expression tag	UNP P39779
F	261	HIS	-	expression tag	UNP P39779
F	262	HIS	-	expression tag	UNP P39779
F	263	HIS	-	expression tag	UNP P39779
F	264	HIS	-	expression tag	UNP P39779
G	260	HIS	-	expression tag	UNP P39779
G	261	HIS	-	expression tag	UNP P39779
G	262	HIS	-	expression tag	UNP P39779
G	263	HIS	-	expression tag	UNP P39779
G	264	HIS	-	expression tag	UNP P39779
H	260	HIS	-	expression tag	UNP P39779
H	261	HIS	-	expression tag	UNP P39779
H	262	HIS	-	expression tag	UNP P39779
H	263	HIS	-	expression tag	UNP P39779
H	264	HIS	-	expression tag	UNP P39779
I	260	HIS	-	expression tag	UNP P39779
I	261	HIS	-	expression tag	UNP P39779
I	262	HIS	-	expression tag	UNP P39779
I	263	HIS	-	expression tag	UNP P39779
I	264	HIS	-	expression tag	UNP P39779
K	260	HIS	-	expression tag	UNP P39779
K	261	HIS	-	expression tag	UNP P39779
K	262	HIS	-	expression tag	UNP P39779
K	263	HIS	-	expression tag	UNP P39779
K	264	HIS	-	expression tag	UNP P39779

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	E	1	Total O S 5 4 1	0	0
2	E	1	Total O S 5 4 1	0	0
2	F	1	Total O S 5 4 1	0	0
2	G	1	Total O S 5 4 1	0	0
2	H	1	Total O S 5 4 1	0	0
2	I	1	Total O S 5 4 1	0	0
2	K	1	Total O S 5 4 1	0	0

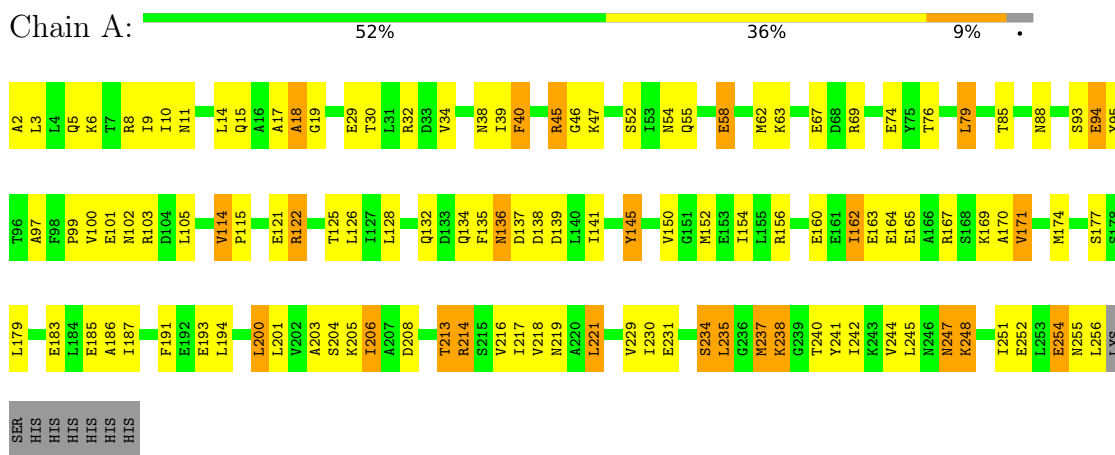
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	4	Total O 4 4	0	0
3	B	2	Total O 2 2	0	0
3	C	1	Total O 1 1	0	0
3	H	1	Total O 1 1	0	0
3	I	1	Total O 1 1	0	0
3	K	1	Total O 1 1	0	0

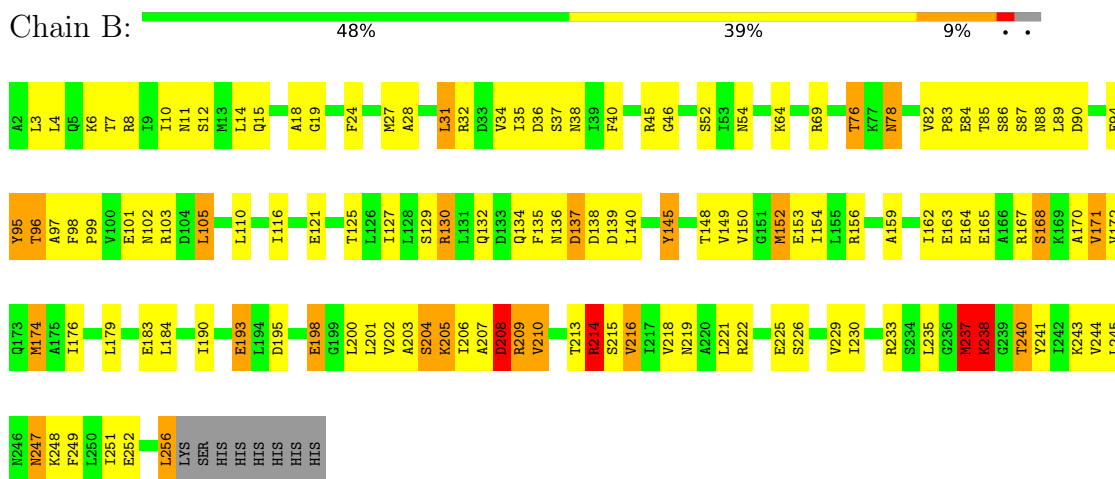
3 Residue-property plots

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.

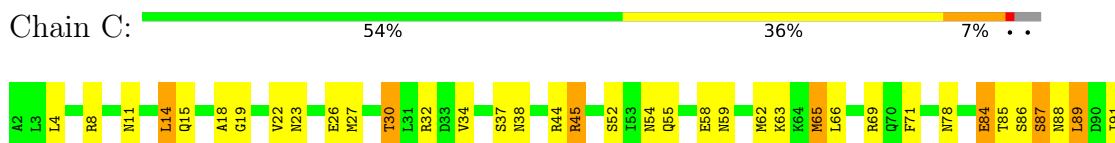
- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

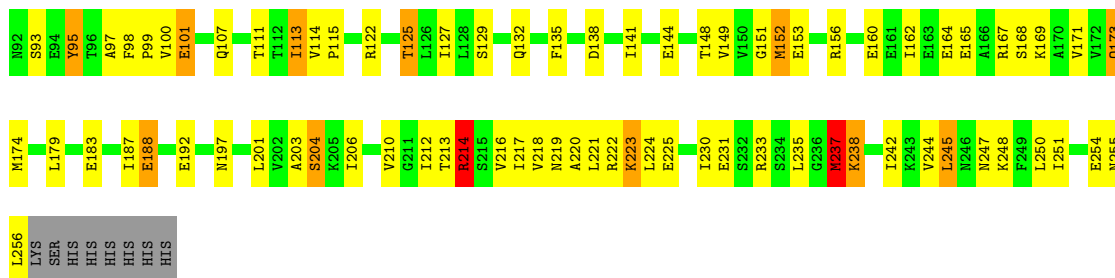


- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY



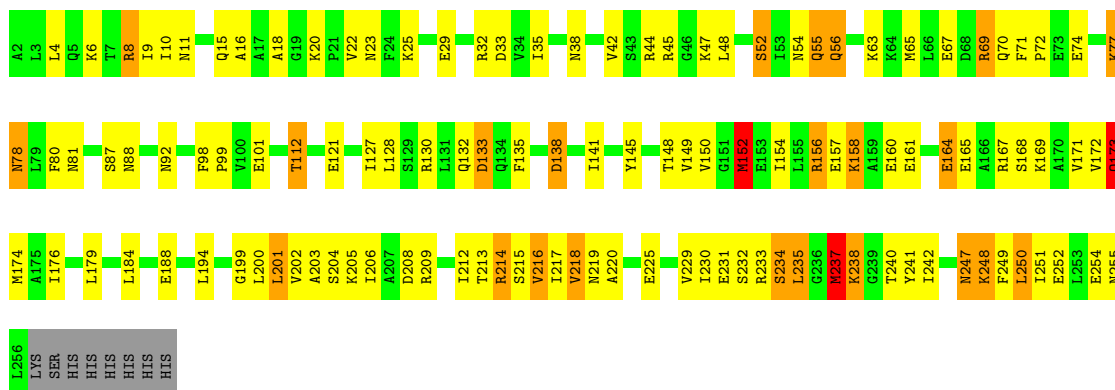
- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY





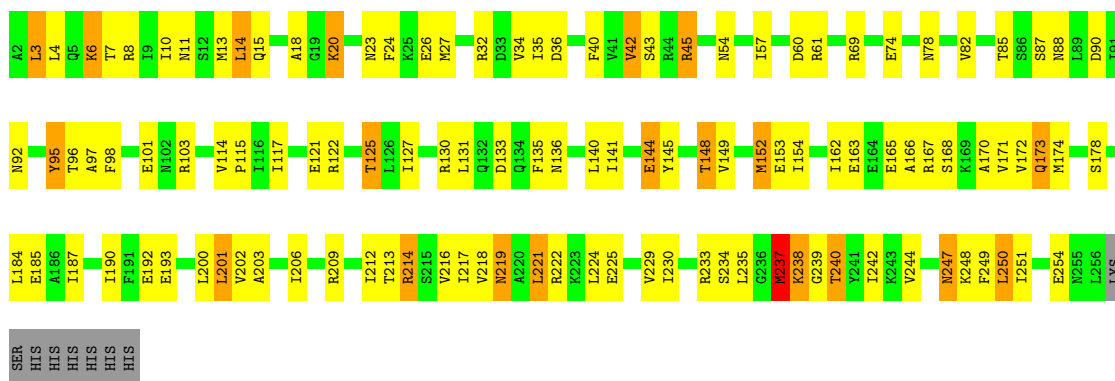
- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain D: 51% 36% 9% ..



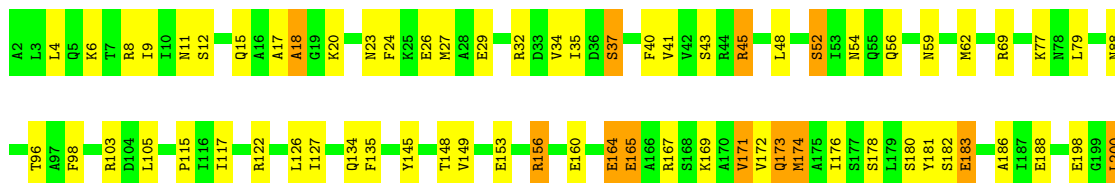
- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

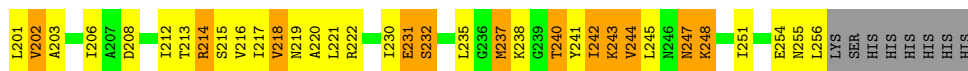
Chain E: 53% 36% 8% .



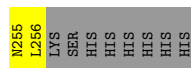
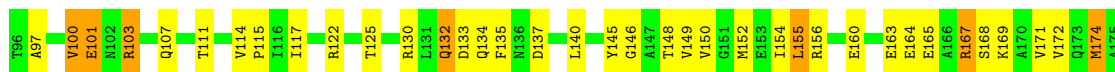
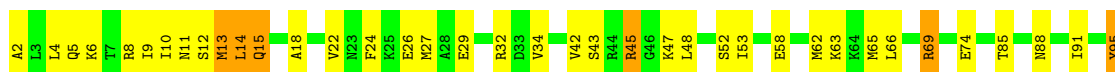
- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain F: 58% 30% 9% .





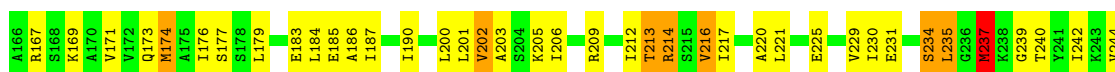
● Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY



● Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY



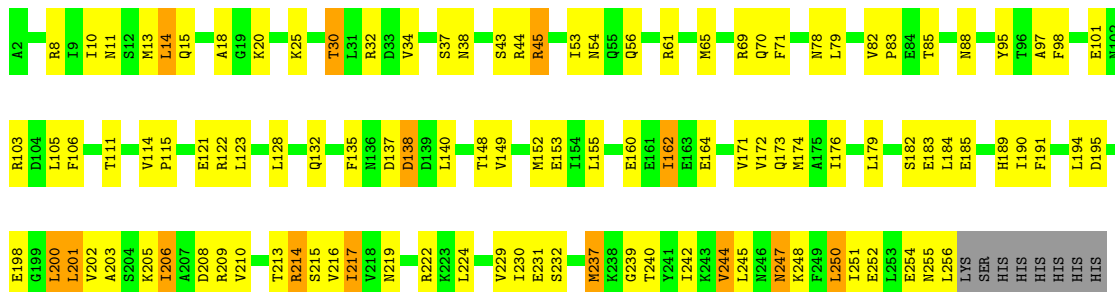
● Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY





- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain K: 56% 36% 5%



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	138.87Å 110.55Å 257.41Å 90.00° 91.30° 90.00°	Depositor
Resolution (Å)	46.68 – 3.00 46.66 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.68-3.00) 72.7 (46.66-3.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.26 (at 3.01Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.230 , 0.279 0.278 , 0.316	Depositor DCC
R_{free} test set	2880 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	94.6	Xtrriage
Anisotropy	0.491	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.018 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	20090	wwPDB-VP
Average B, all atoms (Å ²)	86.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% 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 i

5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: 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	A	0.78	0/2015	0.88	1/2702 (0.0%)
1	B	1.37	21/2015 (1.0%)	1.16	15/2702 (0.6%)
1	C	0.70	4/2015 (0.2%)	0.77	2/2702 (0.1%)
1	D	0.96	10/2015 (0.5%)	0.83	2/2702 (0.1%)
1	E	1.00	5/2015 (0.2%)	0.85	2/2702 (0.1%)
1	F	0.98	15/2015 (0.7%)	0.79	0/2702
1	G	0.90	5/2015 (0.2%)	0.86	3/2702 (0.1%)
1	H	0.66	1/2015 (0.0%)	0.74	0/2702
1	I	0.68	0/2015	0.83	1/2702 (0.0%)
1	K	0.70	3/2015 (0.1%)	0.75	0/2702
All	All	0.90	64/20150 (0.3%)	0.85	26/27020 (0.1%)

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	B	0	2
1	C	0	1
1	D	0	1
1	E	0	1
1	G	0	1
1	H	0	1
1	I	0	1
All	All	0	8

The worst 5 of 64 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	205	LYS	CD-CE	25.69	2.15	1.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	208	ASP	CG-OD1	19.59	1.70	1.25
1	E	173	GLN	CD-NE2	17.84	1.77	1.32
1	D	234	SER	CB-OG	17.26	1.64	1.42
1	E	144	GLU	CD-OE2	16.55	1.43	1.25

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	209	ARG	NE-CZ-NH2	-19.90	110.35	120.30
1	B	214	ARG	NE-CZ-NH2	-19.79	110.41	120.30
1	D	233	ARG	NE-CZ-NH1	-11.18	114.71	120.30
1	B	214	ARG	NE-CZ-NH1	10.51	125.56	120.30
1	B	233	ARG	NE-CZ-NH1	-10.12	115.24	120.30

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	208	ASP	Sidechain
1	B	237	MSE	Peptide
1	C	237	MSE	Peptide
1	D	237	MSE	Peptide
1	E	237	MSE	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	2002	0	2041	103	0
1	B	2002	0	2041	116	0
1	C	2002	0	2041	80	0
1	D	2002	0	2041	106	0
1	E	2002	0	2041	92	0
1	F	2002	0	2041	82	0
1	G	2002	0	2041	90	0
1	H	2002	0	2041	117	0
1	I	2002	0	2041	101	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	K	2002	0	2041	82	0
2	A	10	0	0	0	0
2	B	10	0	0	1	0
2	C	5	0	0	0	0
2	E	10	0	0	1	0
2	F	5	0	0	0	0
2	G	5	0	0	0	0
2	H	5	0	0	0	0
2	I	5	0	0	0	0
2	K	5	0	0	0	0
3	A	4	0	0	0	0
3	B	2	0	0	0	0
3	C	1	0	0	0	0
3	H	1	0	0	0	0
3	I	1	0	0	0	0
3	K	1	0	0	0	0
All	All	20090	0	20410	914	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 23.

The worst 5 of 914 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:F:243:LYS:CE	1:F:243:LYS:CD	1.74	1.63
1:E:20:LYS:CG	1:E:20:LYS:CB	1.77	1.58
1:D:158:LYS:CE	1:D:158:LYS:CD	1.76	1.58
1:F:243:LYS:CE	1:F:243:LYS:NZ	1.69	1.55
1:C:223:LYS:CE	1:C:223:LYS:NZ	1.69	1.54

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	253/263 (96%)	229 (90%)	19 (8%)	5 (2%)	7	34
1	B	253/263 (96%)	229 (90%)	22 (9%)	2 (1%)	19	57
1	C	253/263 (96%)	231 (91%)	18 (7%)	4 (2%)	9	40
1	D	253/263 (96%)	229 (90%)	20 (8%)	4 (2%)	9	40
1	E	253/263 (96%)	229 (90%)	21 (8%)	3 (1%)	13	48
1	F	253/263 (96%)	228 (90%)	22 (9%)	3 (1%)	13	48
1	G	253/263 (96%)	229 (90%)	21 (8%)	3 (1%)	13	48
1	H	253/263 (96%)	235 (93%)	16 (6%)	2 (1%)	19	57
1	I	253/263 (96%)	227 (90%)	23 (9%)	3 (1%)	13	48
1	K	253/263 (96%)	228 (90%)	23 (9%)	2 (1%)	19	57
All	All	2530/2630 (96%)	2294 (91%)	205 (8%)	31 (1%)	13	48

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	18	ALA
1	A	237	MSE
1	B	18	ALA
1	B	237	MSE
1	C	18	ALA

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	221/222 (100%)	187 (85%)	34 (15%)	2	13
1	B	221/222 (100%)	184 (83%)	37 (17%)	2	11
1	C	221/222 (100%)	185 (84%)	36 (16%)	2	11
1	D	221/222 (100%)	192 (87%)	29 (13%)	4	18

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	221/222 (100%)	177 (80%)	44 (20%)	1	7
1	F	221/222 (100%)	191 (86%)	30 (14%)	3	17
1	G	221/222 (100%)	182 (82%)	39 (18%)	2	10
1	H	221/222 (100%)	179 (81%)	42 (19%)	1	8
1	I	221/222 (100%)	178 (80%)	43 (20%)	1	7
1	K	221/222 (100%)	184 (83%)	37 (17%)	2	11
All	All	2210/2220 (100%)	1839 (83%)	371 (17%)	2	11

5 of 371 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	174	MSE
1	H	242	ILE
1	G	214	ARG
1	H	66	LEU
1	I	56	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 59 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	70	GLN
1	K	173	GLN
1	G	134	GLN
1	K	88	ASN
1	I	189	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

12 ligands are modelled in this entry.

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
2	SO4	K	301	-	4,4,4	0.64	0	6,6,6	0.27	0
2	SO4	G	301	-	4,4,4	0.18	0	6,6,6	0.45	0
2	SO4	I	301	-	4,4,4	0.17	0	6,6,6	0.19	0
2	SO4	B	302	-	4,4,4	0.16	0	6,6,6	0.16	0
2	SO4	E	301	-	4,4,4	0.63	0	6,6,6	0.34	0
2	SO4	C	301	-	4,4,4	1.15	0	6,6,6	0.58	0
2	SO4	E	302	-	4,4,4	0.41	0	6,6,6	0.14	0
2	SO4	H	301	-	4,4,4	0.56	0	6,6,6	0.15	0
2	SO4	B	301	-	4,4,4	0.21	0	6,6,6	0.33	0
2	SO4	A	301	-	4,4,4	0.49	0	6,6,6	0.32	0
2	SO4	A	302	-	4,4,4	0.16	0	6,6,6	0.18	0
2	SO4	F	301	-	4,4,4	0.22	0	6,6,6	0.25	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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	302	SO4	1	0
2	E	301	SO4	1	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.