



wwPDB X-ray Structure Validation Summary Report

Dec 16, 2023 – 11:31 PM EST

PDB ID : 2O3O
Title : Crystal Structure of the sensor histidine kinase regulator YycI from *Bacillus subtilis*
Authors : Santelli, E.; Liddington, R.C.
Deposited on : 2006-12-01
Resolution : 2.89 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 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:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

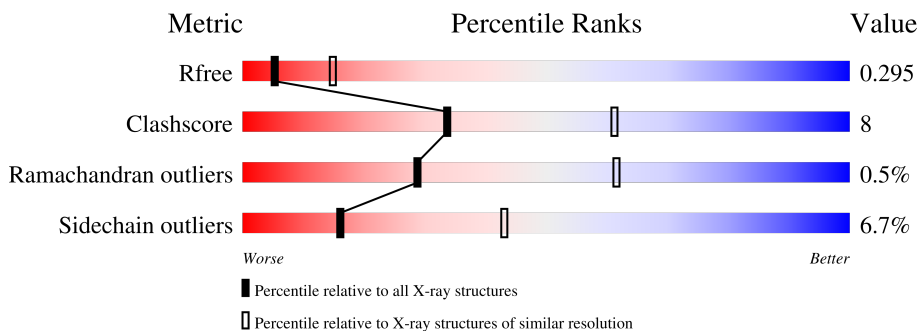
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.89 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)

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	254	
1	B	254	
1	C	254	
1	D	254	
1	E	254	
1	F	254	
1	G	254	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	254	 77% 16% • 5%
1	I	254	 76% 17% • •
1	J	254	 75% 16% • 6%
1	K	254	 75% 16% • 7%
1	L	254	 74% 19% • 6%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 23401 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 YycI protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	241	1965	1251	313	394	1	6	0	0	0
1	B	239	1947	1241	310	389	1	6	0	0	0
1	C	239	1947	1240	310	390	1	6	0	0	0
1	D	241	1965	1249	313	396	1	6	0	0	0
1	E	237	1932	1232	308	385	1	6	0	0	0
1	F	240	1958	1248	313	390	1	6	0	0	0
1	G	237	1929	1228	306	388	1	6	0	0	0
1	H	241	1965	1249	313	396	1	6	0	0	0
1	I	244	1986	1264	318	397	1	6	0	0	0
1	J	238	1941	1238	310	386	1	6	0	0	0
1	K	235	1913	1219	304	383	1	6	0	0	0
1	L	239	1947	1237	309	394	1	6	0	0	0

There are 120 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	GLY	-	cloning artifact	UNP Q45612
A	28	SER	-	cloning artifact	UNP Q45612
A	29	HIS	-	cloning artifact	UNP Q45612
A	30	MSE	-	cloning artifact	UNP Q45612
A	47	MSE	MET	modified residue	UNP Q45612

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	89	MSE	MET	modified residue	UNP Q45612
A	91	MSE	MET	modified residue	UNP Q45612
A	103	MSE	MET	modified residue	UNP Q45612
A	167	MSE	MET	modified residue	UNP Q45612
A	204	MSE	MET	modified residue	UNP Q45612
B	27	GLY	-	cloning artifact	UNP Q45612
B	28	SER	-	cloning artifact	UNP Q45612
B	29	HIS	-	cloning artifact	UNP Q45612
B	30	MSE	-	cloning artifact	UNP Q45612
B	47	MSE	MET	modified residue	UNP Q45612
B	89	MSE	MET	modified residue	UNP Q45612
B	91	MSE	MET	modified residue	UNP Q45612
B	103	MSE	MET	modified residue	UNP Q45612
B	167	MSE	MET	modified residue	UNP Q45612
B	204	MSE	MET	modified residue	UNP Q45612
C	27	GLY	-	cloning artifact	UNP Q45612
C	28	SER	-	cloning artifact	UNP Q45612
C	29	HIS	-	cloning artifact	UNP Q45612
C	30	MSE	-	cloning artifact	UNP Q45612
C	47	MSE	MET	modified residue	UNP Q45612
C	89	MSE	MET	modified residue	UNP Q45612
C	91	MSE	MET	modified residue	UNP Q45612
C	103	MSE	MET	modified residue	UNP Q45612
C	167	MSE	MET	modified residue	UNP Q45612
C	204	MSE	MET	modified residue	UNP Q45612
D	27	GLY	-	cloning artifact	UNP Q45612
D	28	SER	-	cloning artifact	UNP Q45612
D	29	HIS	-	cloning artifact	UNP Q45612
D	30	MSE	-	cloning artifact	UNP Q45612
D	47	MSE	MET	modified residue	UNP Q45612
D	89	MSE	MET	modified residue	UNP Q45612
D	91	MSE	MET	modified residue	UNP Q45612
D	103	MSE	MET	modified residue	UNP Q45612
D	167	MSE	MET	modified residue	UNP Q45612
D	204	MSE	MET	modified residue	UNP Q45612
E	27	GLY	-	cloning artifact	UNP Q45612
E	28	SER	-	cloning artifact	UNP Q45612
E	29	HIS	-	cloning artifact	UNP Q45612
E	30	MSE	-	cloning artifact	UNP Q45612
E	47	MSE	MET	modified residue	UNP Q45612
E	89	MSE	MET	modified residue	UNP Q45612
E	91	MSE	MET	modified residue	UNP Q45612

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	103	MSE	MET	modified residue	UNP Q45612
E	167	MSE	MET	modified residue	UNP Q45612
E	204	MSE	MET	modified residue	UNP Q45612
F	27	GLY	-	cloning artifact	UNP Q45612
F	28	SER	-	cloning artifact	UNP Q45612
F	29	HIS	-	cloning artifact	UNP Q45612
F	30	MSE	-	cloning artifact	UNP Q45612
F	47	MSE	MET	modified residue	UNP Q45612
F	89	MSE	MET	modified residue	UNP Q45612
F	91	MSE	MET	modified residue	UNP Q45612
F	103	MSE	MET	modified residue	UNP Q45612
F	167	MSE	MET	modified residue	UNP Q45612
F	204	MSE	MET	modified residue	UNP Q45612
G	27	GLY	-	cloning artifact	UNP Q45612
G	28	SER	-	cloning artifact	UNP Q45612
G	29	HIS	-	cloning artifact	UNP Q45612
G	30	MSE	-	cloning artifact	UNP Q45612
G	47	MSE	MET	modified residue	UNP Q45612
G	89	MSE	MET	modified residue	UNP Q45612
G	91	MSE	MET	modified residue	UNP Q45612
G	103	MSE	MET	modified residue	UNP Q45612
G	167	MSE	MET	modified residue	UNP Q45612
G	204	MSE	MET	modified residue	UNP Q45612
H	27	GLY	-	cloning artifact	UNP Q45612
H	28	SER	-	cloning artifact	UNP Q45612
H	29	HIS	-	cloning artifact	UNP Q45612
H	30	MSE	-	cloning artifact	UNP Q45612
H	47	MSE	MET	modified residue	UNP Q45612
H	89	MSE	MET	modified residue	UNP Q45612
H	91	MSE	MET	modified residue	UNP Q45612
H	103	MSE	MET	modified residue	UNP Q45612
H	167	MSE	MET	modified residue	UNP Q45612
H	204	MSE	MET	modified residue	UNP Q45612
I	27	GLY	-	cloning artifact	UNP Q45612
I	28	SER	-	cloning artifact	UNP Q45612
I	29	HIS	-	cloning artifact	UNP Q45612
I	30	MSE	-	cloning artifact	UNP Q45612
I	47	MSE	MET	modified residue	UNP Q45612
I	89	MSE	MET	modified residue	UNP Q45612
I	91	MSE	MET	modified residue	UNP Q45612
I	103	MSE	MET	modified residue	UNP Q45612
I	167	MSE	MET	modified residue	UNP Q45612

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
I	204	MSE	MET	modified residue	UNP Q45612
J	27	GLY	-	cloning artifact	UNP Q45612
J	28	SER	-	cloning artifact	UNP Q45612
J	29	HIS	-	cloning artifact	UNP Q45612
J	30	MSE	-	cloning artifact	UNP Q45612
J	47	MSE	MET	modified residue	UNP Q45612
J	89	MSE	MET	modified residue	UNP Q45612
J	91	MSE	MET	modified residue	UNP Q45612
J	103	MSE	MET	modified residue	UNP Q45612
J	167	MSE	MET	modified residue	UNP Q45612
J	204	MSE	MET	modified residue	UNP Q45612
K	27	GLY	-	cloning artifact	UNP Q45612
K	28	SER	-	cloning artifact	UNP Q45612
K	29	HIS	-	cloning artifact	UNP Q45612
K	30	MSE	-	cloning artifact	UNP Q45612
K	47	MSE	MET	modified residue	UNP Q45612
K	89	MSE	MET	modified residue	UNP Q45612
K	91	MSE	MET	modified residue	UNP Q45612
K	103	MSE	MET	modified residue	UNP Q45612
K	167	MSE	MET	modified residue	UNP Q45612
K	204	MSE	MET	modified residue	UNP Q45612
L	27	GLY	-	cloning artifact	UNP Q45612
L	28	SER	-	cloning artifact	UNP Q45612
L	29	HIS	-	cloning artifact	UNP Q45612
L	30	MSE	-	cloning artifact	UNP Q45612
L	47	MSE	MET	modified residue	UNP Q45612
L	89	MSE	MET	modified residue	UNP Q45612
L	91	MSE	MET	modified residue	UNP Q45612
L	103	MSE	MET	modified residue	UNP Q45612
L	167	MSE	MET	modified residue	UNP Q45612
L	204	MSE	MET	modified residue	UNP Q45612

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Cl 1 1	0	0
2	D	1	Total Cl 1 1	0	0
2	E	1	Total Cl 1 1	0	0
2	F	1	Total Cl 1 1	0	0

Continued on next page...

Continued from previous page...

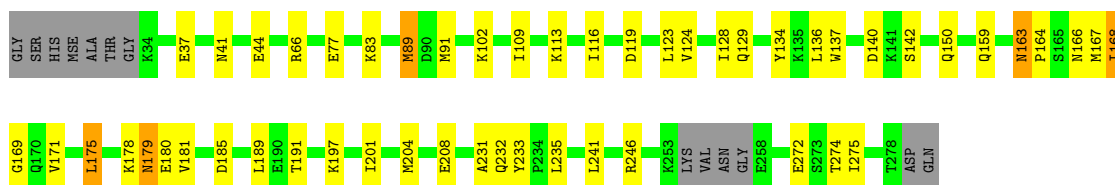
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	K	1	Total 1	Cl 1	0	0
2	L	1	Total 1	Cl 1	0	0

3 Residue-property plots [i](#)

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.

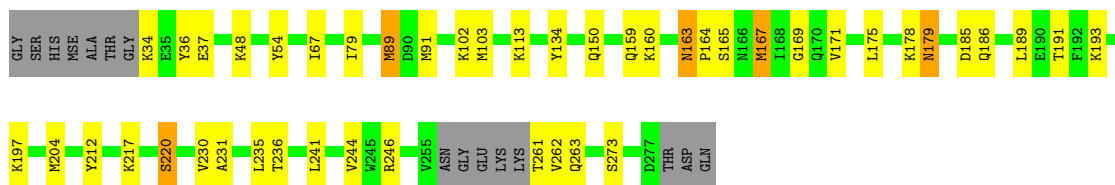
- Molecule 1: YycI protein

Chain A: 




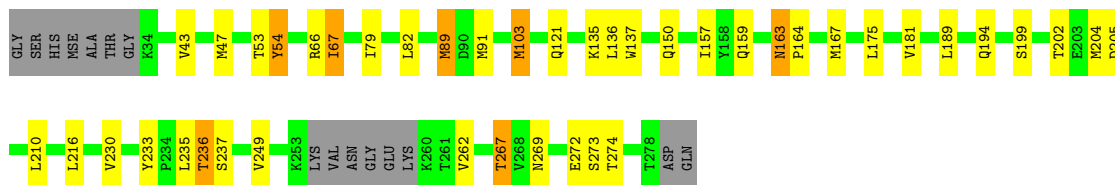
- Molecule 1: YycI protein

Chain B: 



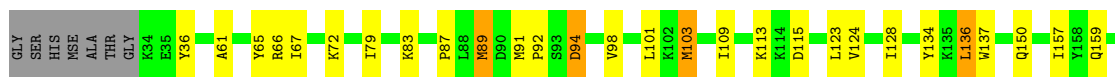
- Molecule 1: YycI protein

Chain C: 



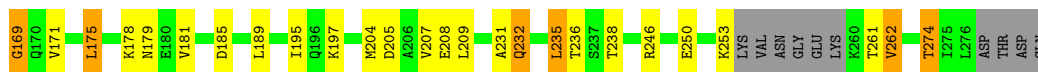
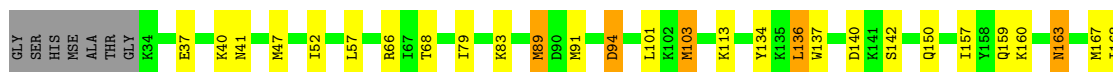
- Molecule 1: YycI protein

Chain D: 

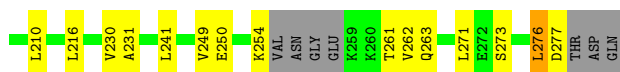
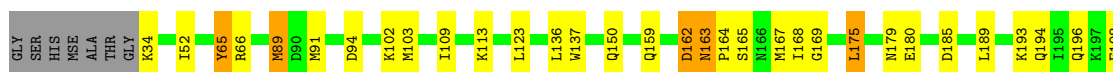
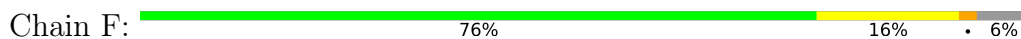




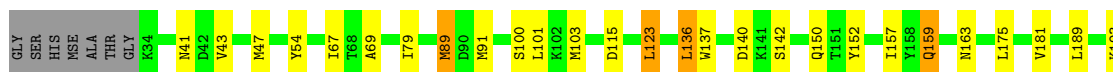
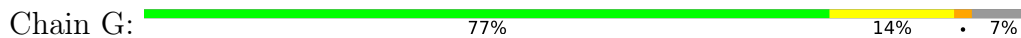
• Molecule 1: YycI protein



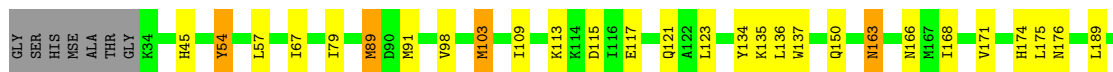
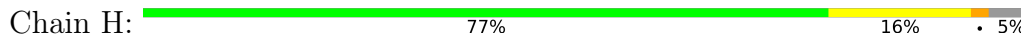
• Molecule 1: YycI protein



• Molecule 1: YycI protein

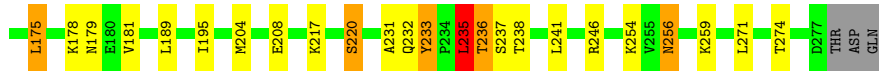


• Molecule 1: YycI protein

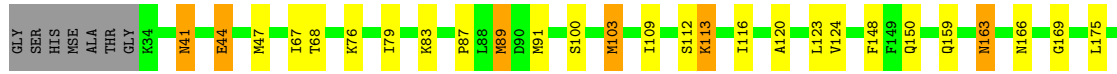


• Molecule 1: YycI protein

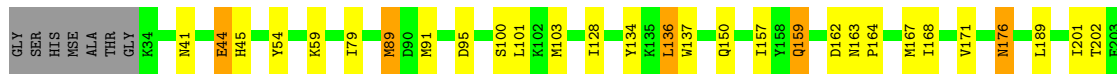
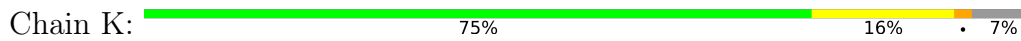




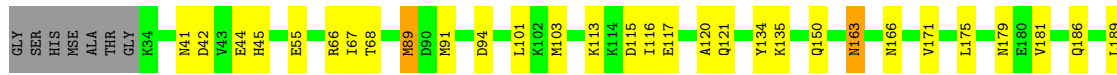
• Molecule 1: YycI protein



• Molecule 1: YycI protein



• Molecule 1: YycI protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	60.79Å 161.92Å 180.16Å 90.00° 90.90° 90.00°	Depositor
Resolution (Å)	90.00 – 2.89 48.43 – 2.89	Depositor EDS
% Data completeness (in resolution range)	92.2 (90.00-2.89) 92.2 (48.43-2.89)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.20 (at 2.91Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.212 , 0.262 0.254 , 0.295	Depositor DCC
R_{free} test set	3605 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	53.2	Xtrriage
Anisotropy	0.245	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , -10.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.037 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	23401	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 19.67% 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: CL

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	0.51	0/1994	0.59	0/2679
1	B	0.50	0/1976	0.61	0/2656
1	C	0.50	0/1976	0.59	0/2656
1	D	0.50	0/1994	0.62	0/2679
1	E	0.50	0/1961	0.60	0/2635
1	F	0.50	0/1987	0.62	0/2668
1	G	0.49	0/1958	0.58	0/2634
1	H	0.47	0/1994	0.57	0/2679
1	I	0.53	0/2016	0.62	0/2709
1	J	0.51	0/1970	0.61	0/2646
1	K	0.48	0/1942	0.58	0/2612
1	L	0.46	0/1976	0.57	0/2657
All	All	0.49	0/23744	0.60	0/31910

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	1965	0	1943	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1947	0	1926	30	0
1	C	1947	0	1924	34	0
1	D	1965	0	1936	47	0
1	E	1932	0	1913	44	0
1	F	1958	0	1943	31	0
1	G	1929	0	1898	32	0
1	H	1965	0	1936	33	0
1	I	1986	0	1968	37	0
1	J	1941	0	1926	36	0
1	K	1913	0	1885	36	0
1	L	1947	0	1910	29	0
2	A	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	K	1	0	0	0	0
2	L	1	0	0	0	0
All	All	23401	0	23108	393	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 8.

The worst 5 of 393 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:C:89:MSE:HE2	1:C:91:MSE:HE1	1.22	1.12
1:C:79:ILE:HG23	1:C:103:MSE:HE1	1.33	1.06
1:A:89:MSE:HE2	1:A:91:MSE:HE1	1.34	1.05
1:A:89:MSE:HE2	1:A:91:MSE:CE	1.93	0.97
1:E:235:LEU:HD13	1:G:235:LEU:HD13	1.49	0.91

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	237/254 (93%)	224 (94%)	12 (5%)	1 (0%)	34	66
1	B	235/254 (92%)	219 (93%)	15 (6%)	1 (0%)	34	66
1	C	235/254 (92%)	220 (94%)	13 (6%)	2 (1%)	17	48
1	D	237/254 (93%)	221 (93%)	15 (6%)	1 (0%)	34	66
1	E	233/254 (92%)	215 (92%)	16 (7%)	2 (1%)	17	48
1	F	236/254 (93%)	219 (93%)	16 (7%)	1 (0%)	34	66
1	G	233/254 (92%)	217 (93%)	16 (7%)	0	100	100
1	H	237/254 (93%)	223 (94%)	12 (5%)	2 (1%)	19	51
1	I	242/254 (95%)	223 (92%)	15 (6%)	4 (2%)	9	31
1	J	234/254 (92%)	221 (94%)	13 (6%)	0	100	100
1	K	231/254 (91%)	218 (94%)	13 (6%)	0	100	100
1	L	235/254 (92%)	222 (94%)	12 (5%)	1 (0%)	34	66
All	All	2825/3048 (93%)	2642 (94%)	168 (6%)	15 (0%)	29	61

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	179	ASN
1	C	236	THR
1	C	273	SER
1	I	233	TYR
1	B	179	ASN

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	221/224 (99%)	203 (92%)	18 (8%)	11	33
1	B	219/224 (98%)	206 (94%)	13 (6%)	19	49

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	219/224 (98%)	207 (94%)	12 (6%)	21	53
1	D	221/224 (99%)	207 (94%)	14 (6%)	18	46
1	E	217/224 (97%)	198 (91%)	19 (9%)	10	30
1	F	220/224 (98%)	204 (93%)	16 (7%)	14	38
1	G	217/224 (97%)	202 (93%)	15 (7%)	15	41
1	H	221/224 (99%)	209 (95%)	12 (5%)	22	54
1	I	223/224 (100%)	207 (93%)	16 (7%)	14	39
1	J	218/224 (97%)	205 (94%)	13 (6%)	19	49
1	K	215/224 (96%)	202 (94%)	13 (6%)	19	49
1	L	219/224 (98%)	204 (93%)	15 (7%)	16	42
All	All	2630/2688 (98%)	2454 (93%)	176 (7%)	16	43

5 of 176 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	204	MSE
1	J	116	ILE
1	I	44	GLU
1	I	220	SER
1	K	54	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 82 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	71	GLN
1	K	71	GLN
1	I	150	GLN
1	J	159	GLN
1	K	232	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 6 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.

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.