



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:13 AM GMT

PDB ID : 3DL8
Title : Structure of the complex of aquifex aeolicus SecYEG and bacillus subtilis SecA
Authors : Nam, Y.; Zimmer, J.; Rapoport, T.A.
Deposited on : 2008-06-26
Resolution : 7.50 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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

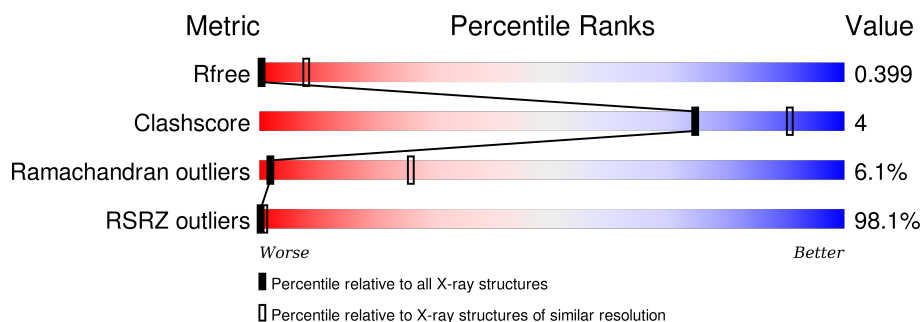
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 7.50 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1014 (9.50-3.66)
Clashscore	102246	1063 (10.00-3.70)
Ramachandran outliers	100387	1035 (9.50-3.66)
RSRZ outliers	91569	1013 (9.50-3.66)

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	A	779	<div> <div>96%</div> <div> <div></div> <div>91%</div> <div>8%</div> <div>.</div> </div> </div>
1	B	779	<div> <div>98%</div> <div> <div></div> <div>91%</div> <div>7%</div> <div>.</div> </div> </div>
2	G	429	<div> <div>93%</div> <div> <div></div> <div>85%</div> <div>9%</div> <div>5%</div> </div> </div>
2	H	429	<div> <div>93%</div> <div> <div></div> <div>85%</div> <div>9%</div> <div>5%</div> </div> </div>
3	C	65	<div> <div>54%</div> <div> <div></div> <div>54%</div> <div>46%</div> </div> </div>
3	D	65	<div> <div>54%</div> <div> <div></div> <div>54%</div> <div>46%</div> </div> </div>
4	E	107	<div> <div>61%</div> <div> <div></div> <div>58%</div> <div>39%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	107	<p>61% 39%</p>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10232 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 Protein translocase subunit secA.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	773	Total	C	N	O	0	0	0
			3092	1546	773	773			
1	B	773	Total	C	N	O	0	0	0
			3092	1546	773	773			

- Molecule 2 is a protein called Preprotein translocase subunit secY.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	406	Total	C	N	O	0	0	0
			1624	812	406	406			
2	H	406	Total	C	N	O	0	0	0
			1624	812	406	406			

- Molecule 3 is a protein called SecE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	35	Total	C	N	O	0	0	0
			140	70	35	35			
3	D	35	Total	C	N	O	0	0	0
			140	70	35	35			

- Molecule 4 is a protein called Protein-export membrane protein secG.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	E	65	Total	C	N	O	0	0	0
			260	130	65	65			
4	F	65	Total	C	N	O	0	0	0
			260	130	65	65			

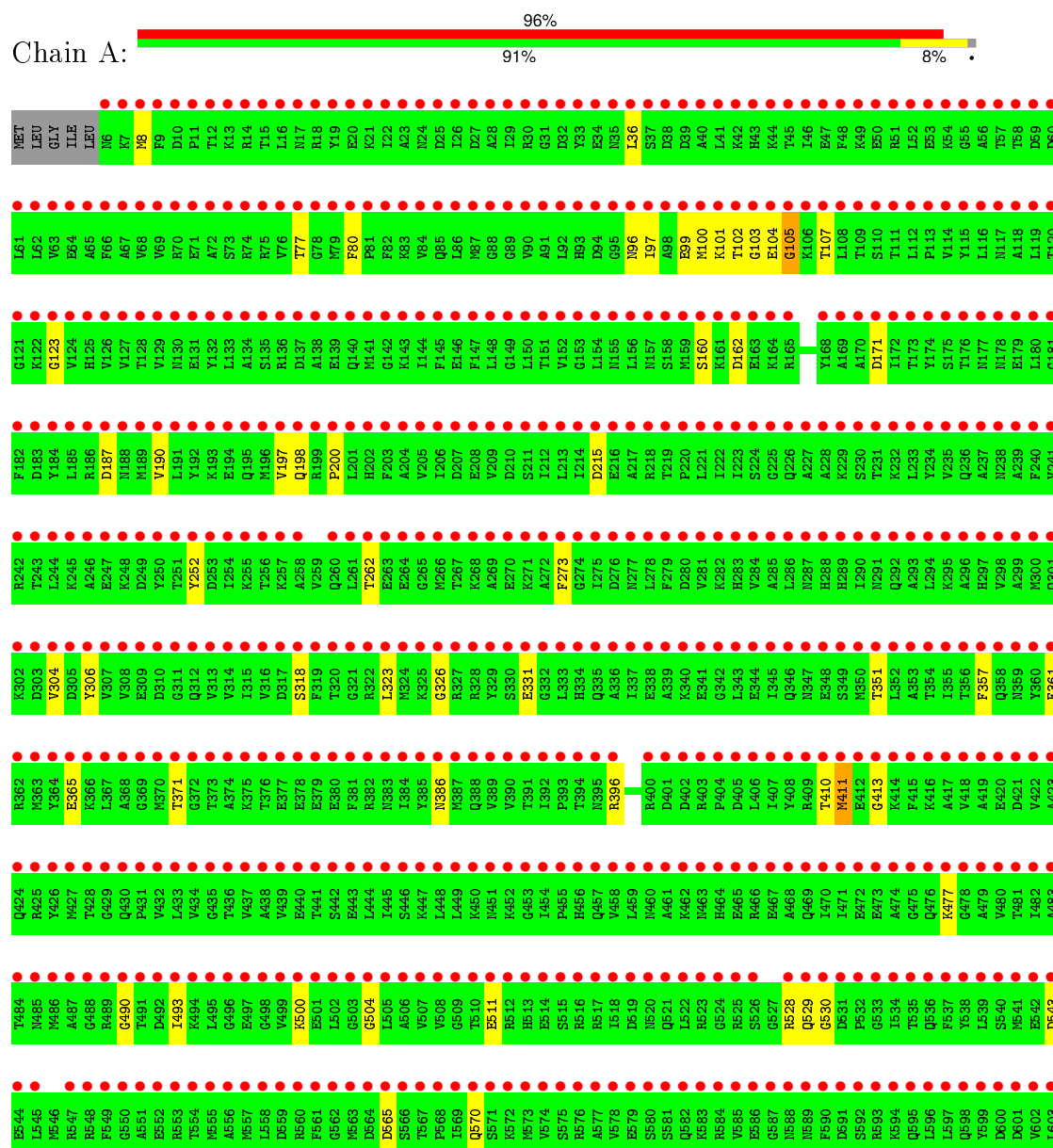
There are 14 discrepancies between the modelled and reference sequences:

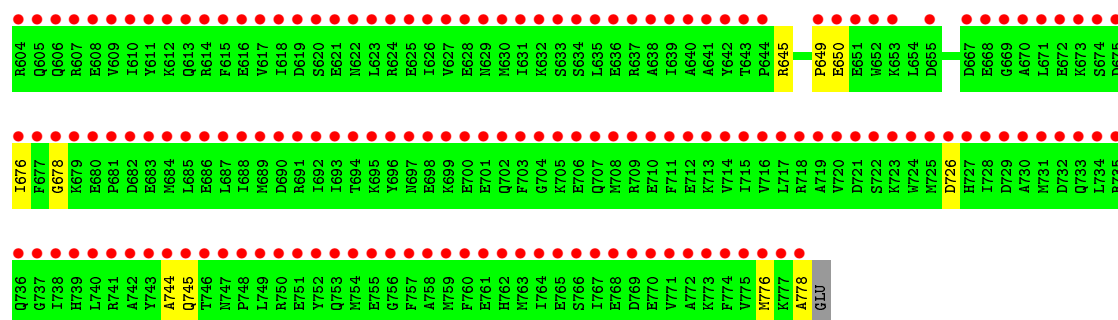
Chain	Residue	Modelled	Actual	Comment	Reference
E	101	GLY	-	EXPRESSION TAG	UNP O66505
E	102	HIS	-	EXPRESSION TAG	UNP O66505
E	103	HIS	-	INSERTION	UNP O66505
E	104	HIS	-	EXPRESSION TAG	UNP O66505
E	105	HIS	-	EXPRESSION TAG	UNP O66505
E	106	HIS	-	EXPRESSION TAG	UNP O66505
E	107	HIS	-	EXPRESSION TAG	UNP O66505
F	101	GLY	-	EXPRESSION TAG	UNP O66505
F	102	HIS	-	EXPRESSION TAG	UNP O66505
F	103	HIS	-	EXPRESSION TAG	UNP O66505
F	104	HIS	-	EXPRESSION TAG	UNP O66505
F	105	HIS	-	EXPRESSION TAG	UNP O66505
F	106	HIS	-	EXPRESSION TAG	UNP O66505
F	107	HIS	-	EXPRESSION TAG	UNP O66505

3 Residue-property plots

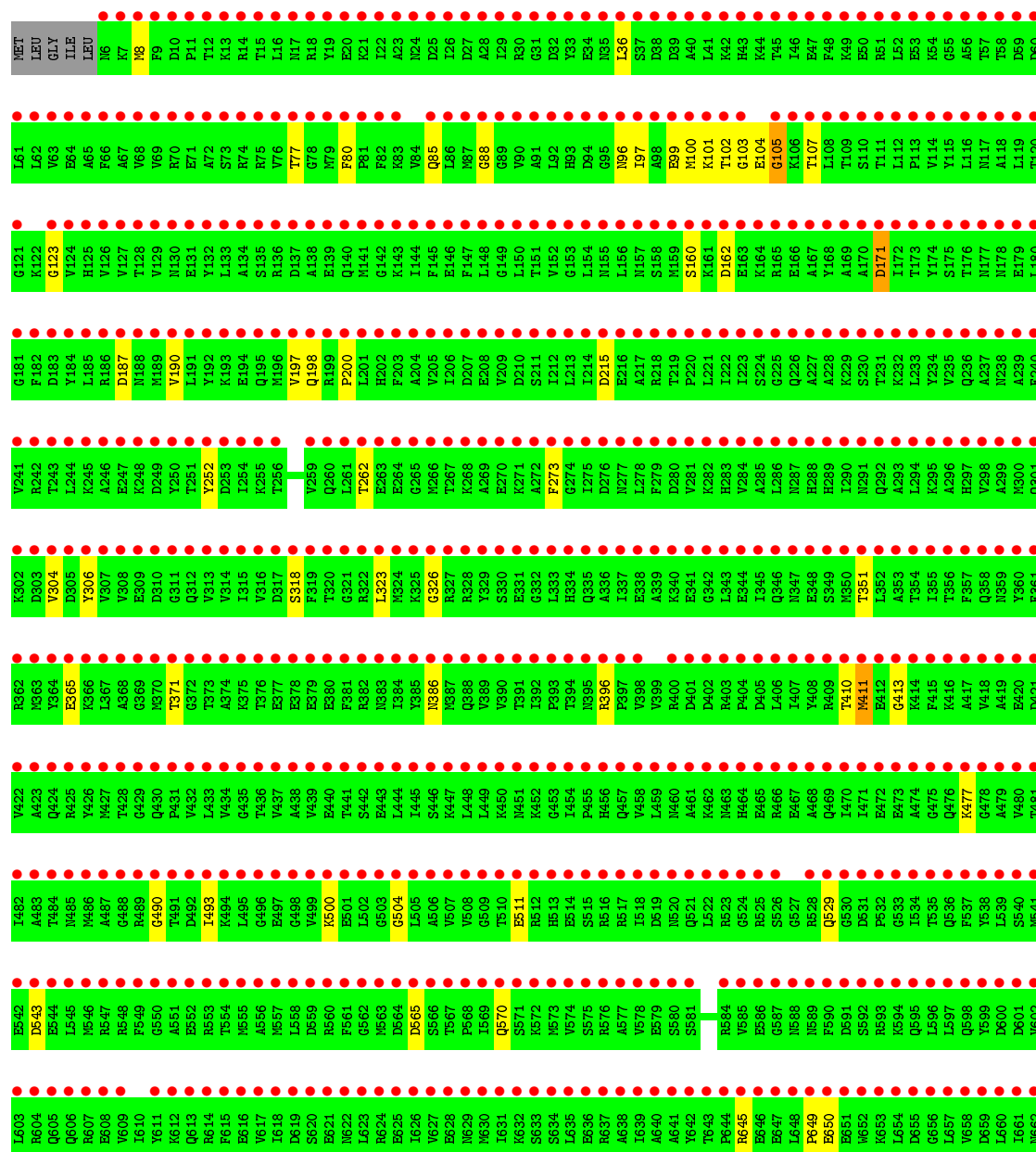
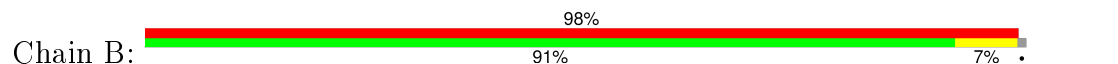
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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. A red dot above a residue indicates a poor fit to the electron density ($\text{RSRZ} > 2$). 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: Protein translocase subunit secA





• Molecule 1: Protein translocase subunit secA

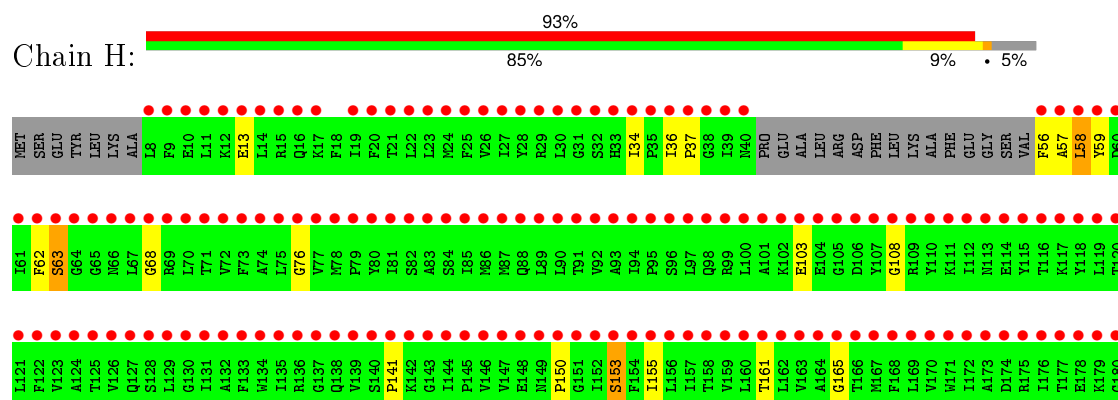


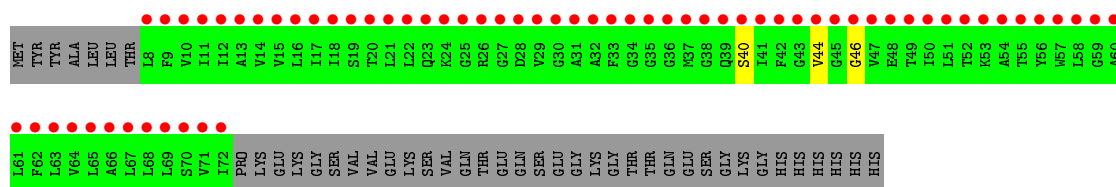


• Molecule 2: Preprotein translocase subunit secY



• Molecule 2: Preprotein translocase subunit secY





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	146.36Å 167.97Å 187.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.99 – 7.50 48.09 – 7.49	Depositor EDS
% Data completeness (in resolution range)	99.5 (14.99-7.50) 99.0 (48.09-7.49)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.28 (at 7.37Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.365 , 0.390 0.397 , 0.399	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	530.1	Xtriage
Anisotropy	0.140	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.59 , -10.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	1 of 6217 reflections (0.016%)	Xtriage
F_o, F_c correlation	0.60	EDS
Total number of atoms	10232	wwPDB-VP
Average B, all atoms (Å ²)	452.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.91% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.88	0/3091	0.90	2/3862 (0.1%)
1	B	0.88	0/3091	0.90	2/3862 (0.1%)
2	G	0.98	0/1622	0.96	2/2024 (0.1%)
2	H	0.97	0/1622	0.96	2/2024 (0.1%)
3	C	0.94	0/139	0.75	0/172
3	D	0.94	0/139	0.75	0/172
4	E	1.02	0/259	0.93	0/322
4	F	1.02	0/259	0.93	0/322
All	All	0.92	0/10222	0.92	8/12760 (0.1%)

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	160	SER	N-CA-C	-7.19	91.59	111.00
1	B	160	SER	N-CA-C	-7.17	91.63	111.00
1	A	304	VAL	C-N-CA	6.29	137.41	121.70
1	B	304	VAL	C-N-CA	6.25	137.31	121.70
2	H	302	ASP	C-N-CA	-5.36	108.31	121.70

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	3092	0	856	18	0
1	B	3092	0	856	12	0
2	G	1624	0	446	17	0
2	H	1624	0	446	13	0
3	C	140	0	36	0	0
3	D	140	0	36	0	0
4	E	260	0	86	1	0
4	F	260	0	86	1	0
All	All	10232	0	2848	52	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 4.

The worst 5 of 52 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:A:776:MET:O	1:B:778:ALA:O	1.54	1.25
1:A:776:MET:O	1:B:778:ALA:C	1.98	1.02
1:A:778:ALA:C	1:B:776:MET:O	1.99	1.00
1:A:778:ALA:O	1:B:776:MET:O	1.83	0.95
2:H:103:GLU:H	2:H:108:GLY:HA3	1.39	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 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	771/779 (99%)	638 (83%)	86 (11%)	47 (6%)	2	26
1	B	771/779 (99%)	638 (83%)	85 (11%)	48 (6%)	2	26
2	G	402/429 (94%)	321 (80%)	52 (13%)	29 (7%)	1	22
2	H	402/429 (94%)	319 (79%)	54 (13%)	29 (7%)	1	22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	33/65 (51%)	33 (100%)	0	0	100	100
3	D	33/65 (51%)	33 (100%)	0	0	100	100
4	E	63/107 (59%)	57 (90%)	5 (8%)	1 (2%)	12	56
4	F	63/107 (59%)	57 (90%)	5 (8%)	1 (2%)	12	56
All	All	2538/2760 (92%)	2096 (83%)	287 (11%)	155 (6%)	2	26

5 of 155 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	36	LEU
1	A	80	PHE
1	A	99	GLU
1	A	100	MET
1	A	101	LYS

5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

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](#)

There are no ligands in this entry.

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	A	773/779 (99%)	11.54	749 (96%) 0 1	321, 444, 550, 648	0
1	B	773/779 (99%)	12.42	762 (98%) 0 0	340, 470, 540, 612	0
2	G	406/429 (94%)	13.16	400 (98%) 0 0	321, 409, 488, 512	0
2	H	406/429 (94%)	13.15	398 (98%) 0 1	372, 466, 536, 558	0
3	C	35/65 (53%)	19.86	35 (100%) 0 0	328, 363, 402, 405	0
3	D	35/65 (53%)	27.20	35 (100%) 0 0	383, 396, 463, 466	0
4	E	65/107 (60%)	15.91	65 (100%) 0 0	386, 446, 468, 473	0
4	F	65/107 (60%)	8.92	65 (100%) 0 0	477, 523, 550, 551	0
All	All	2558/2760 (92%)	12.69	2509 (98%) 0 1	321, 450, 541, 648	0

The worst 5 of 2509 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	153	SER	61.1
1	B	472	GLU	56.9
1	A	694	THR	56.8
2	G	356	THR	54.8
2	H	158	THR	52.8

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.