



# wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Aug 20, 2017 – 08:05 AM EDT

PDB ID : 2Y9K  
EMDB ID: : EMD-1871  
Title : Three-dimensional model of Salmonella's needle complex at subnanometer resolution  
Authors : Schraidt, O.; Marlovits, T.C.  
Deposited on : unknown  
Resolution : 8.30 Å(reported)  
Based on PDB ID : 3GR5

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

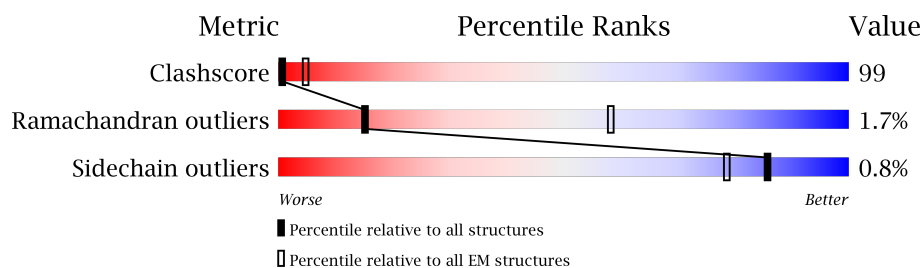
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 8.30 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.







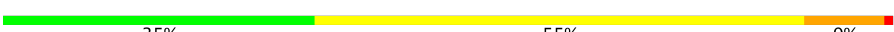

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	137	36% 54% 9% .
1	B	137	36% 54% 9% .
1	C	137	37% 53% 9% .
1	D	137	36% 54% 9% .
1	E	137	36% 54% 9% .
1	F	137	35% 55% 9% .
1	G	137	36% 53% 9% .
1	H	137	36% 54% 9% .
1	I	137	36% 54% 9% .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	J	137	 36%54%9% .
1	K	137	 36%54%9% .
1	L	137	 36%54%9% .
1	M	137	 36%54%9% .
1	N	137	 35%55%9% .
1	O	137	 36%53%9% .

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 16485 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.

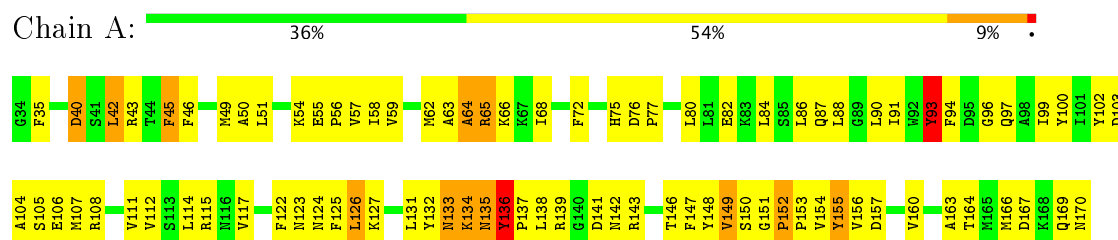
- Molecule 1 is a protein called PROTEIN INVG.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	B	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	C	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	D	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	E	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	F	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	G	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	H	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	I	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	J	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	K	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	L	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	M	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	N	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		
1	O	137	Total	C	N	O	S	0	0
			1099	705	187	201	6		

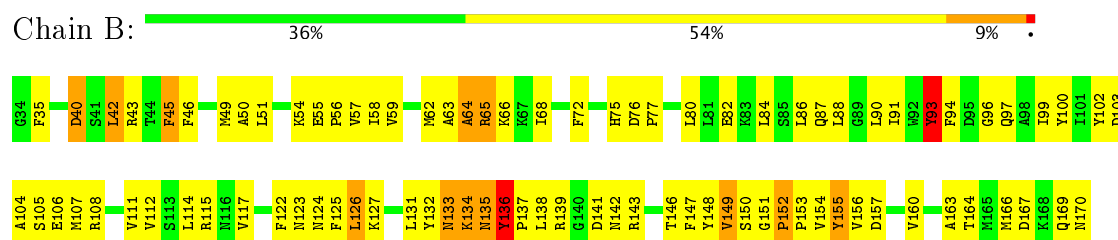
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

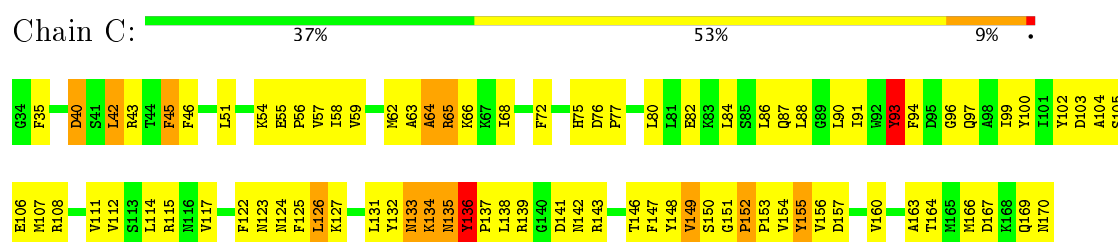
#### • Molecule 1: PROTEIN INVG



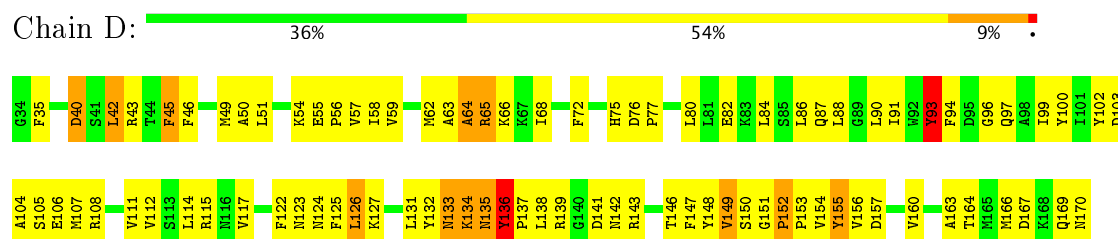
#### • Molecule 1: PROTEIN INVG



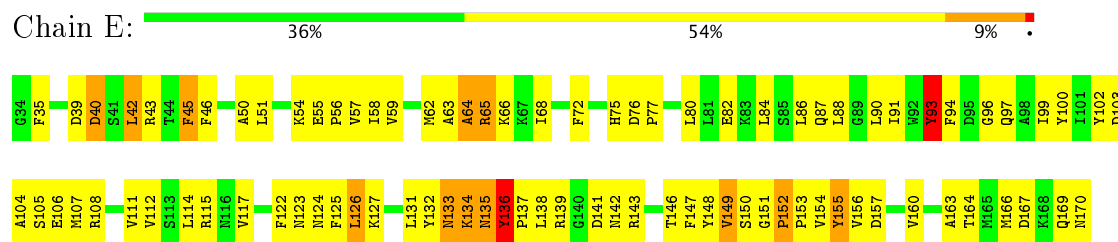
#### • Molecule 1: PROTEIN INVG



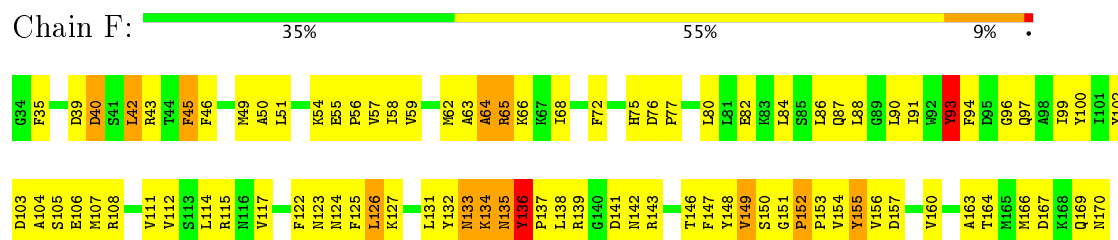
#### • Molecule 1: PROTEIN INVG



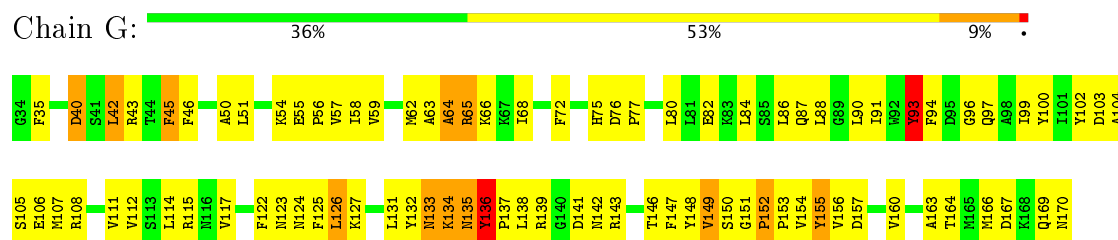
#### • Molecule 1: PROTEIN INVG



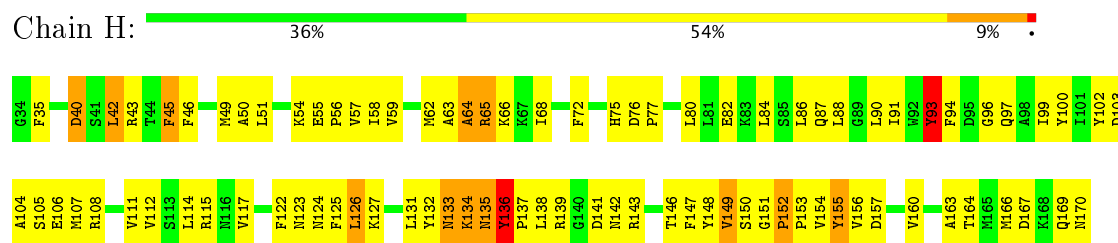
- Molecule 1: PROTEIN INVG



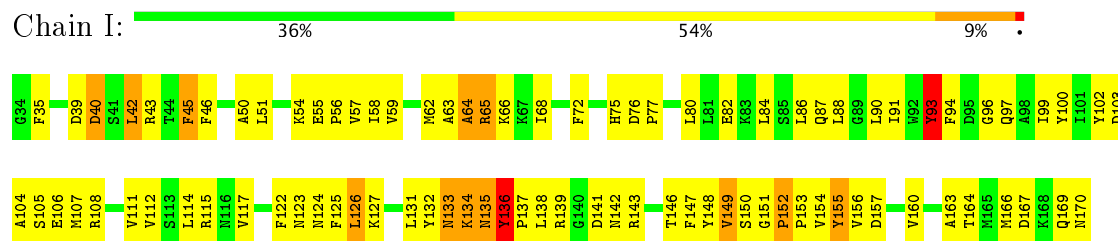
- Molecule 1: PROTEIN INVG



- Molecule 1: PROTEIN INVG

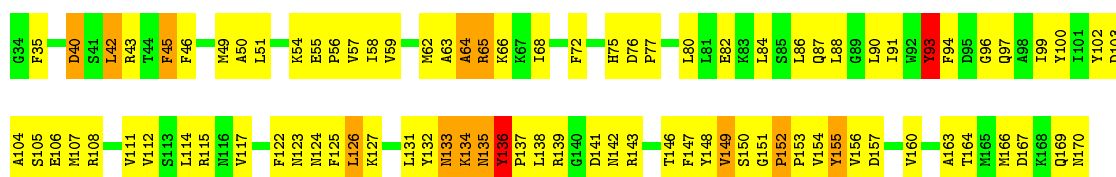


- Molecule 1: PROTEIN INVG



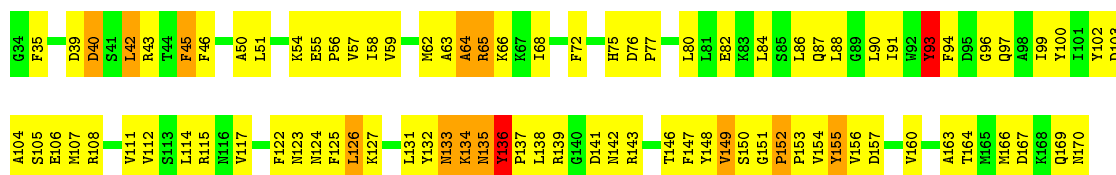
- Molecule 1: PROTEIN INVG





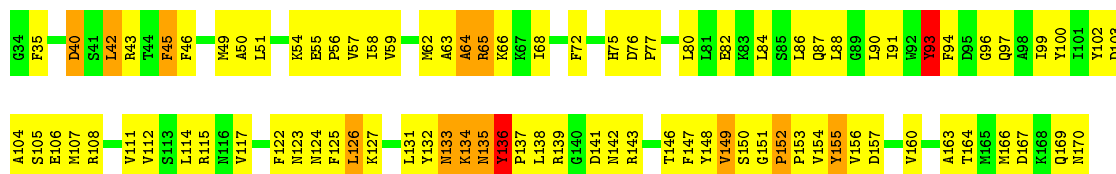
• Molecule 1: PROTEIN INVG

Chain K: 36% 54% 9%



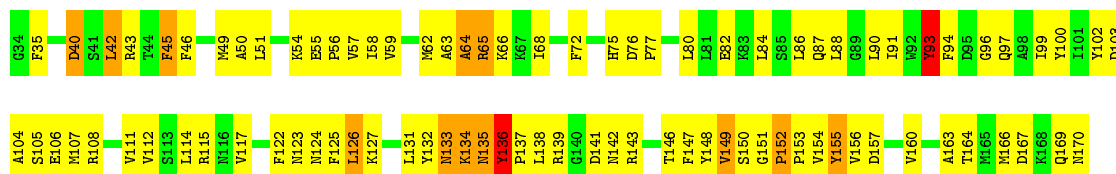
• Molecule 1: PROTEIN INVG

Chain L: 36% 54% 9%



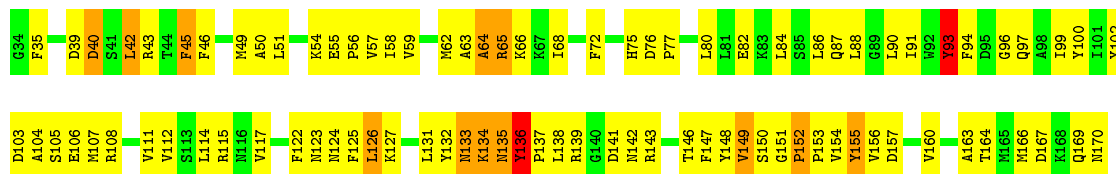
• Molecule 1: PROTEIN INVG

Chain M: 36% 54% 9%



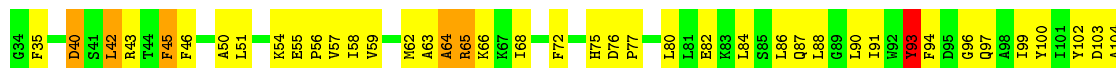
• Molecule 1: PROTEIN INVG

Chain N: 35% 55% 9%



• Molecule 1: PROTEIN INVG

Chain O: 36% 53% 9%



S105	E106	M107	R108		V111	V112	S113	L114	R115	M116	V117		F122	M123	M124	F125	L126	K127		L131	Y132	M133	K134	M135	Y136	P137	L138	R139		D141	M142	R143		T146	F147	Y148	Y149	S150	G151	P152	P153	V154	Y155	V156	D157		V160		A163	T164	M165	M166	D167	K168	Q169	N170
------	------	------	------	--	------	------	------	------	------	------	------	--	------	------	------	------	------	------	--	------	------	------	------	------	------	------	------	------	--	------	------	------	--	------	------	------	------	------	------	------	------	------	------	------	------	--	------	--	------	------	------	------	------	------	------	------



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C15	Depositor
Number of particles used	Not provided	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	EACH CCD FRAME	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	93000	Depositor
Image detector	GENERIC GATAN (4k x 4k)	Depositor

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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	$\# Z  > 2$	RMSZ	$\# Z  > 2$
1	A	1.06	8/1122 (0.7%)	1.94	19/1512 (1.3%)
1	B	1.06	8/1122 (0.7%)	1.94	19/1512 (1.3%)
1	C	1.06	8/1122 (0.7%)	1.94	19/1512 (1.3%)
1	D	1.06	7/1122 (0.6%)	1.94	19/1512 (1.3%)
1	E	1.06	8/1122 (0.7%)	1.94	19/1512 (1.3%)
1	F	1.06	8/1122 (0.7%)	1.94	19/1512 (1.3%)
1	G	1.06	8/1122 (0.7%)	1.94	19/1512 (1.3%)
1	H	1.06	8/1122 (0.7%)	1.94	19/1512 (1.3%)
1	I	1.06	8/1122 (0.7%)	1.94	19/1512 (1.3%)
1	J	1.06	7/1122 (0.6%)	1.94	19/1512 (1.3%)
1	K	1.05	8/1122 (0.7%)	1.94	19/1512 (1.3%)
1	L	1.06	8/1122 (0.7%)	1.94	19/1512 (1.3%)
1	M	1.06	8/1122 (0.7%)	1.93	19/1512 (1.3%)
1	N	1.06	7/1122 (0.6%)	1.94	19/1512 (1.3%)
1	O	1.06	7/1122 (0.6%)	1.94	19/1512 (1.3%)
All	All	1.06	116/16830 (0.7%)	1.94	285/22680 (1.3%)

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	A	0	4
1	B	0	4
1	C	0	4
1	D	0	4
1	E	0	4
1	F	0	4
1	G	0	4
1	H	0	4
1	I	0	4
1	J	0	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
1	K	0	4
1	L	0	4
1	M	0	4
1	N	0	4
1	O	0	4
All	All	0	60

The worst 5 of 116 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	O	136	TYR	CD2-CE2	-14.13	1.18	1.39
1	E	136	TYR	CD2-CE2	-14.11	1.18	1.39
1	L	136	TYR	CD2-CE2	-14.08	1.18	1.39
1	C	136	TYR	CD2-CE2	-14.07	1.18	1.39
1	N	136	TYR	CD2-CE2	-14.07	1.18	1.39

The worst 5 of 285 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	136	TYR	CB-CG-CD2	28.96	138.38	121.00
1	A	136	TYR	CB-CG-CD2	28.88	138.33	121.00
1	G	136	TYR	CB-CG-CD2	28.88	138.33	121.00
1	I	136	TYR	CB-CG-CD2	28.84	138.31	121.00
1	D	136	TYR	CB-CG-CD2	28.83	138.30	121.00

There are no chirality outliers.

5 of 60 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	133	ASN	Peptide
1	A	136	TYR	Sidechain
1	A	64	ALA	Peptide
1	A	93	TYR	Sidechain
1	B	64	ALA	Peptide

## 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	1099	0	1094	278	0
1	B	1099	0	1094	285	0
1	C	1099	0	1094	277	0
1	D	1099	0	1094	280	0
1	E	1099	0	1094	284	0
1	F	1099	0	1094	277	0
1	G	1099	0	1094	278	0
1	H	1099	0	1094	277	0
1	I	1099	0	1094	278	0
1	J	1099	0	1094	283	0
1	K	1099	0	1094	274	0
1	L	1099	0	1094	278	0
1	M	1099	0	1094	284	0
1	N	1099	0	1094	279	0
1	O	1099	0	1094	280	0
All	All	16485	0	16410	3269	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 99.

The worst 5 of 3269 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:97:GLN:HG3	1:H:105:SER:CB	1.23	1.66
1:M:143:ARG:HH22	1:N:166:MET:CG	1.06	1.65
1:I:166:MET:CG	1:O:143:ARG:HH22	1.08	1.63
1:B:166:MET:CG	1:I:143:ARG:HH22	1.06	1.63
1:A:97:GLN:HG3	1:D:105:SER:CB	1.24	1.63

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	135/137 (98%)	129 (96%)	4 (3%)	2 (2%)	12	53
1	B	135/137 (98%)	129 (96%)	4 (3%)	2 (2%)	12	53
1	C	135/137 (98%)	129 (96%)	4 (3%)	2 (2%)	12	53
1	D	135/137 (98%)	129 (96%)	4 (3%)	2 (2%)	12	53
1	E	135/137 (98%)	129 (96%)	3 (2%)	3 (2%)	8	44
1	F	135/137 (98%)	129 (96%)	3 (2%)	3 (2%)	8	44
1	G	135/137 (98%)	129 (96%)	4 (3%)	2 (2%)	12	53
1	H	135/137 (98%)	129 (96%)	4 (3%)	2 (2%)	12	53
1	I	135/137 (98%)	129 (96%)	3 (2%)	3 (2%)	8	44
1	J	135/137 (98%)	129 (96%)	4 (3%)	2 (2%)	12	53
1	K	135/137 (98%)	129 (96%)	3 (2%)	3 (2%)	8	44
1	L	135/137 (98%)	129 (96%)	4 (3%)	2 (2%)	12	53
1	M	135/137 (98%)	129 (96%)	4 (3%)	2 (2%)	12	53
1	N	135/137 (98%)	129 (96%)	3 (2%)	3 (2%)	8	44
1	O	135/137 (98%)	129 (96%)	4 (3%)	2 (2%)	12	53
All	All	2025/2055 (98%)	1935 (96%)	55 (3%)	35 (2%)	15	50

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	65	ARG
1	A	134	LYS
1	B	65	ARG
1	B	134	LYS
1	C	65	ARG

### 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 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	118/118 (100%)	117 (99%)	1 (1%)	85	92

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	C	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	D	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	E	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	F	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	G	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	H	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	I	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	J	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	K	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	L	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	M	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	N	118/118 (100%)	117 (99%)	1 (1%)	85	92
1	O	118/118 (100%)	117 (99%)	1 (1%)	85	92
All	All	1770/1770 (100%)	1755 (99%)	15 (1%)	86	92

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	152	PRO
1	H	152	PRO
1	M	152	PRO
1	F	152	PRO
1	L	152	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 [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.