



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Dec 9, 2019 – 01:29 AM EST

PDB ID : 6R6B
EMDB ID: : EMD-4734
Title : Structure of the core *Shigella flexneri* type III secretion system export gate complex SctRST (Spa24/Spa9/Spa29).
Authors : Johnson, S.; Kuhlen, L.; Deme, J.C.; Abrusci, P.; Lea, S.M.
Deposited on : 2019-03-26
Resolution : 3.50 Å(reported)

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : 2.4

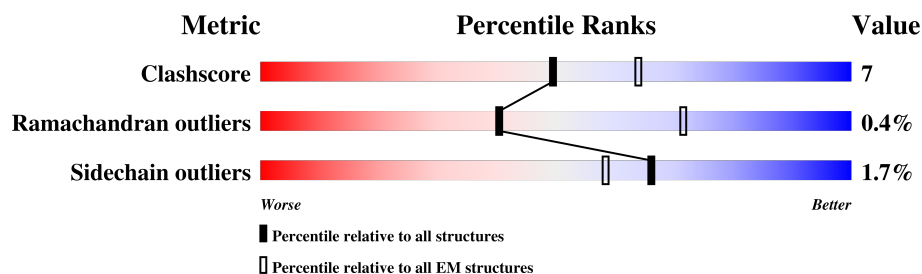
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.50 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)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

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 ≥ 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	216	
1	B	216	
1	C	216	
1	D	216	
1	E	216	
2	F	295	
3	G	86	
3	H	86	
3	I	86	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	J	86	 73%26%.

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11416 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 Surface presentation of antigens protein SpaP.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	189	Total	C	N	O	S	0	0
			1502	1007	226	259	10		
1	B	171	Total	C	N	O	S	0	0
			1347	910	198	229	10		
1	C	169	Total	C	N	O	S	0	0
			1340	905	199	226	10		
1	D	168	Total	C	N	O	S	0	0
			1332	901	198	223	10		
1	E	170	Total	C	N	O	S	0	0
			1348	911	200	227	10		

- Molecule 2 is a protein called Surface presentation of antigens protein SpaR.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	F	245	Total	C	N	O	S	0	0
			1923	1301	291	323	8		

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	257	GLU	-	expression tag	UNP P0A1M6
F	258	ASN	-	expression tag	UNP P0A1M6
F	259	LEU	-	expression tag	UNP P0A1M6
F	260	TYR	-	expression tag	UNP P0A1M6
F	261	PHE	-	expression tag	UNP P0A1M6
F	262	GLN	-	expression tag	UNP P0A1M6
F	263	GLY	-	expression tag	UNP P0A1M6
F	264	GLN	-	expression tag	UNP P0A1M6
F	265	PHE	-	expression tag	UNP P0A1M6
F	266	GLY	-	expression tag	UNP P0A1M6
F	267	SER	-	expression tag	UNP P0A1M6
F	268	TRP	-	expression tag	UNP P0A1M6
F	269	SER	-	expression tag	UNP P0A1M6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	270	HIS	-	expression tag	UNP P0A1M6
F	271	PRO	-	expression tag	UNP P0A1M6
F	272	GLN	-	expression tag	UNP P0A1M6
F	273	PHE	-	expression tag	UNP P0A1M6
F	274	GLU	-	expression tag	UNP P0A1M6
F	275	LYS	-	expression tag	UNP P0A1M6
F	276	GLY	-	expression tag	UNP P0A1M6
F	277	GLY	-	expression tag	UNP P0A1M6
F	278	GLY	-	expression tag	UNP P0A1M6
F	279	SER	-	expression tag	UNP P0A1M6
F	280	GLY	-	expression tag	UNP P0A1M6
F	281	GLY	-	expression tag	UNP P0A1M6
F	282	GLY	-	expression tag	UNP P0A1M6
F	283	SER	-	expression tag	UNP P0A1M6
F	284	GLY	-	expression tag	UNP P0A1M6
F	285	GLY	-	expression tag	UNP P0A1M6
F	286	GLY	-	expression tag	UNP P0A1M6
F	287	SER	-	expression tag	UNP P0A1M6
F	288	TRP	-	expression tag	UNP P0A1M6
F	289	SER	-	expression tag	UNP P0A1M6
F	290	HIS	-	expression tag	UNP P0A1M6
F	291	PRO	-	expression tag	UNP P0A1M6
F	292	GLN	-	expression tag	UNP P0A1M6
F	293	PHE	-	expression tag	UNP P0A1M6
F	294	GLU	-	expression tag	UNP P0A1M6
F	295	LYS	-	expression tag	UNP P0A1M6

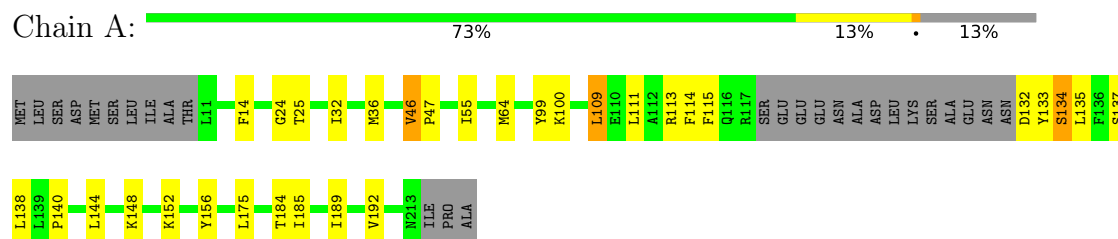
- Molecule 3 is a protein called Surface presentation of antigens protein SpaQ.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	I	85	Total	C	N	O	S	0	0
			656	442	97	113	4		
3	G	85	Total	C	N	O	S	0	0
			656	442	97	113	4		
3	H	85	Total	C	N	O	S	0	0
			656	442	97	113	4		
3	J	85	Total	C	N	O	S	0	0
			656	442	97	113	4		

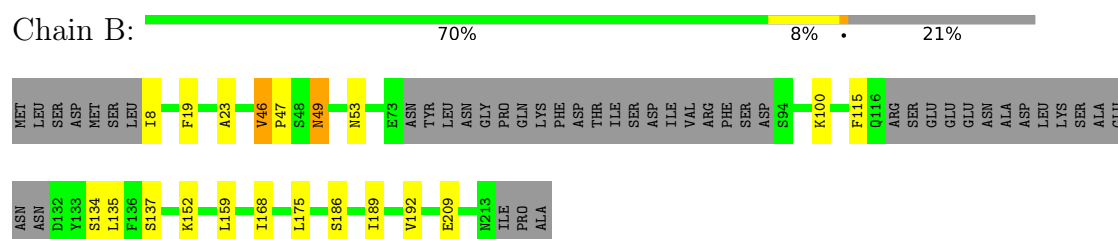
3 Residue-property plots

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.

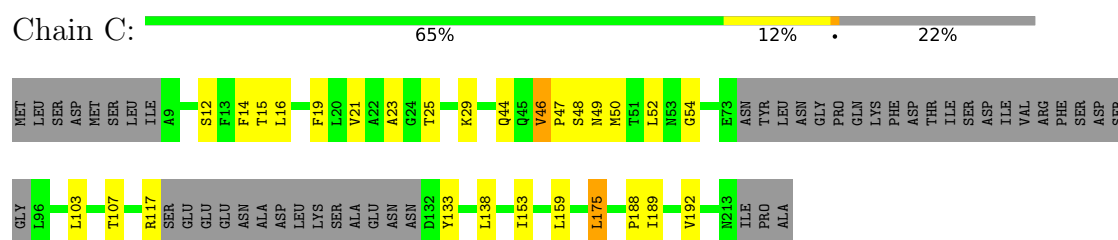
• Molecule 1: Surface presentation of antigens protein SpaP



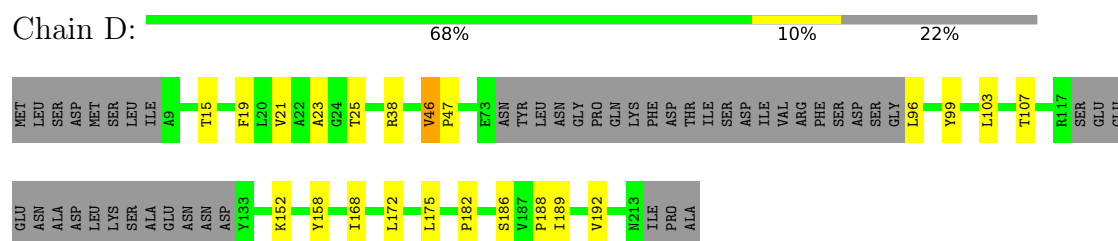
• Molecule 1: Surface presentation of antigens protein SpaP



• Molecule 1: Surface presentation of antigens protein SpaP

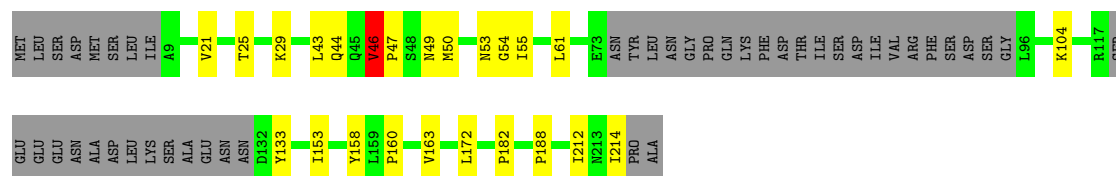


• Molecule 1: Surface presentation of antigens protein SpaP



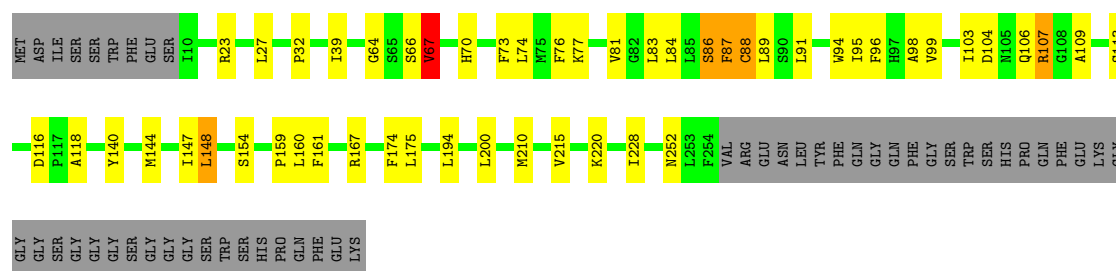
• Molecule 1: Surface presentation of antigens protein SpaP

Chain E:  68% 11% 21%




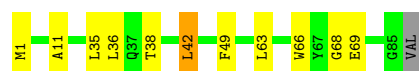
- Molecule 2: Surface presentation of antigens protein SpaR

Chain F:  66% 15% 17%




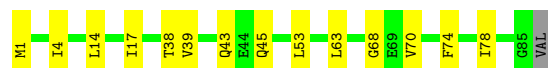
- Molecule 3: Surface presentation of antigens protein SpaQ

Chain I:  86% 12% ..




- Molecule 3: Surface presentation of antigens protein SpaQ

Chain G:  83% 16% .



- Molecule 3: Surface presentation of antigens protein SpaQ

Chain H:  84% 13% ..



- Molecule 3: Surface presentation of antigens protein SpaQ

Chain J:  73% 26% .



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	212561	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 ($e^-/\text{\AA}^2$)	48	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (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	0.52	1/1536 (0.1%)	0.72	0/2080
1	B	0.50	0/1376	0.72	0/1863
1	C	0.45	0/1369	0.73	1/1853 (0.1%)
1	D	0.45	0/1361	0.68	0/1842
1	E	0.45	0/1377	0.68	0/1864
2	F	0.45	0/1973	0.78	3/2680 (0.1%)
3	G	0.42	0/669	0.74	1/908 (0.1%)
3	H	0.38	0/669	0.71	0/908
3	I	0.35	0/669	0.70	2/908 (0.2%)
3	J	0.34	0/669	0.62	1/908 (0.1%)
All	All	0.45	1/11668 (0.0%)	0.72	8/15814 (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	A	0	1
1	E	0	1
2	F	0	1
3	G	0	2
3	I	0	1
All	All	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	109	LEU	C-N	5.70	1.47	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	67	VAL	C-N-CA	12.99	154.18	121.70
2	F	86	SER	CB-CA-C	9.04	127.27	110.10
3	I	63	LEU	CA-CB-CG	6.31	129.81	115.30
3	G	63	LEU	CA-CB-CG	6.18	129.51	115.30
1	C	175	LEU	CA-CB-CG	5.96	129.00	115.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	109	LEU	Mainchain
1	E	46	VAL	Peptide
2	F	32	PRO	Peptide
3	G	38	THR	Peptide
3	I	68	GLY	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	1502	0	1566	21	0
1	B	1347	0	1421	24	0
1	C	1340	0	1415	18	0
1	D	1332	0	1411	18	0
1	E	1348	0	1426	18	0
2	F	1923	0	1998	55	0
3	G	656	0	704	7	0
3	H	656	0	704	20	0
3	I	656	0	704	5	0
3	J	656	0	704	11	0
All	All	11416	0	12053	154	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 7.

The worst 5 of 154 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:B:209:GLU:OE1	3:H:83:LYS:CE	1.81	1.29
1:B:209:GLU:OE1	3:H:83:LYS:CD	1.93	1.16
1:A:115:PHE:O	1:A:137:SER:OG	1.77	1.00
2:F:83:LEU:HD23	2:F:87:PHE:HD1	1.26	0.99
1:B:209:GLU:CD	3:H:83:LYS:CE	2.31	0.97

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	185/216 (86%)	169 (91%)	16 (9%)	0	100	100
1	B	165/216 (76%)	151 (92%)	13 (8%)	1 (1%)	27	68
1	C	163/216 (76%)	148 (91%)	14 (9%)	1 (1%)	27	68
1	D	162/216 (75%)	151 (93%)	11 (7%)	0	100	100
1	E	164/216 (76%)	152 (93%)	12 (7%)	0	100	100
2	F	243/295 (82%)	219 (90%)	22 (9%)	2 (1%)	21	62
3	G	83/86 (96%)	74 (89%)	8 (10%)	1 (1%)	14	53
3	H	83/86 (96%)	76 (92%)	7 (8%)	0	100	100
3	I	83/86 (96%)	75 (90%)	8 (10%)	0	100	100
3	J	83/86 (96%)	73 (88%)	10 (12%)	0	100	100
All	All	1414/1719 (82%)	1288 (91%)	121 (9%)	5 (0%)	40	75

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	G	39	VAL
2	F	159	PRO
1	B	49	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	49	ASN
2	F	67	VAL

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	166/189 (88%)	162 (98%)	4 (2%)	52	79
1	B	148/189 (78%)	147 (99%)	1 (1%)	85	94
1	C	147/189 (78%)	145 (99%)	2 (1%)	69	87
1	D	146/189 (77%)	145 (99%)	1 (1%)	85	94
1	E	148/189 (78%)	146 (99%)	2 (1%)	69	87
2	F	219/258 (85%)	213 (97%)	6 (3%)	48	76
3	G	74/75 (99%)	73 (99%)	1 (1%)	69	87
3	H	74/75 (99%)	71 (96%)	3 (4%)	33	67
3	I	74/75 (99%)	73 (99%)	1 (1%)	69	87
3	J	74/75 (99%)	73 (99%)	1 (1%)	69	87
All	All	1270/1503 (84%)	1248 (98%)	22 (2%)	66	84

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	50	MET
2	F	88	CYS
3	H	83	LYS
2	F	67	VAL
2	F	87	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	53	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	44	GLN
2	F	70	HIS
2	F	106	GLN
3	H	45	GLN

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.