



Full wwPDB X-ray Structure Validation Report

Mar 1, 2014 – 02:37 AM GMT

PDB ID : 4B9G
Title : Structure of CssB subunit complemented with donor strand from CssA subunit of enterotoxigenic Escherichia coli colonization factor CS6
Authors : Roy, S.P.; Rahman, M.M.; Yu, X.D.; Tuittila, M.; Knight, S.D.; Zavialov, A.V.
Deposited on : 2012-09-04
Resolution : 1.04 Å(reported)

This is a full wwPDB validation 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/ValidationPDFNotes.html>

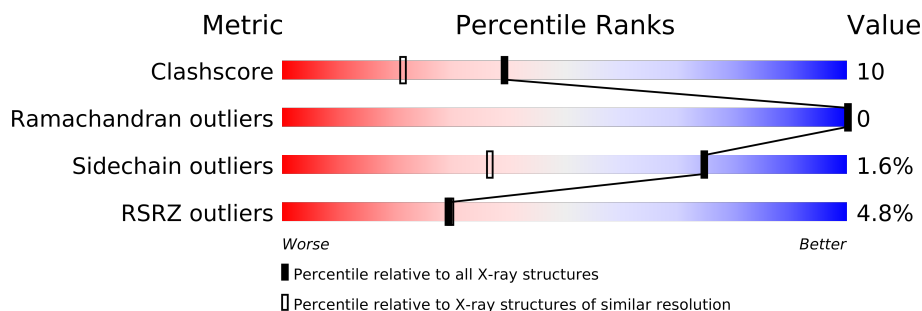
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 1.04 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))
Clashscore	79885	1011 (1.12-0.96)
Ramachandran outliers	78287	1151 (1.14-0.94)
Sidechain outliers	78261	1150 (1.14-0.94)
RSRZ outliers	66119	1106 (1.14-0.94)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	161	
1	B	161	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4844 atoms, of which 2152 are hydrogens and 0 are deuterium.

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 CS6 FIMBRIAL SUBUNIT B, CS6 FIMBRIAL SUBUNIT A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	146	Total	C	H	N	O	S	0	0	0
			2189	695	1076	184	231	3			
1	B	146	Total	C	H	N	O	S	0	0	0
			2189	695	1076	184	231	3			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	EXPRESSION TAG	UNP P53510
A	2	ASN	-	EXPRESSION TAG	UNP P53510
A	3	TRP	-	EXPRESSION TAG	UNP P53510
A	142	ASP	-	LINKER	UNP P53510
A	143	ASN	-	LINKER	UNP P53510
A	144	LYS	-	LINKER	UNP P53510
A	145	GLN	-	LINKER	UNP P53510
B	1	GLY	-	EXPRESSION TAG	UNP P53510
B	2	ASN	-	EXPRESSION TAG	UNP P53510
B	3	TRP	-	EXPRESSION TAG	UNP P53510
B	142	ASP	-	LINKER	UNP P53510
B	143	ASN	-	LINKER	UNP P53510
B	144	LYS	-	LINKER	UNP P53510
B	145	GLN	-	LINKER	UNP P53510

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	254	Total	O	0	0
			254	254		
2	B	212	Total	O	0	0
			212	212		

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 ($RSRZ > 2$). 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.

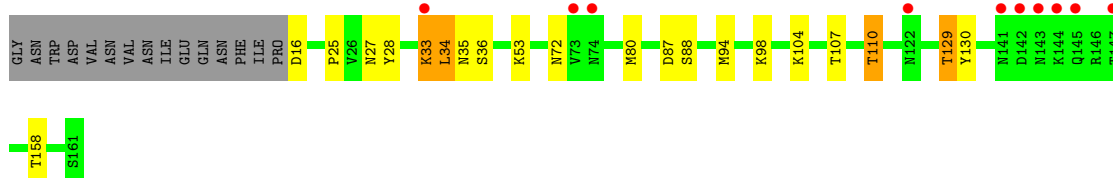
- Molecule 1: CS6 FIMBRIAL SUBUNIT B, CS6 FIMBRIAL SUBUNIT A

Chain A: 



- Molecule 1: CS6 FIMBRIAL SUBUNIT B, CS6 FIMBRIAL SUBUNIT A

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	51.96Å 69.22Å 72.36Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.29 – 1.04 27.29 – 1.05	Depositor EDS
% Data completeness (in resolution range)	98.9 (27.29-1.04) 98.0 (27.29-1.05)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.23 (at 1.05Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.155 , 0.167 0.153 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	8.1	Xtriage
Anisotropy	0.385	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.47 , 43.7	EDS
Estimated twinning fraction	0.014 for -h,l,k	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 119741 reflections	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	4844	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.11% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.76	2/1135 (0.2%)	0.78	0/1540
1	B	0.80	2/1135 (0.2%)	0.83	3/1540 (0.2%)
All	All	0.78	4/2270 (0.2%)	0.80	3/3080 (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	2	0
1	B	3	0
All	All	5	0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	112	SER	CB-OG	-7.25	1.32	1.42
1	B	110	THR	CB-CG2	-6.67	1.30	1.52
1	B	129	THR	CB-CG2	-6.47	1.30	1.52
1	A	152	LYS	CE-NZ	-6.06	1.33	1.49

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	34	LEU	CB-CG-CD1	-6.20	100.46	111.00
1	B	129	THR	CA-CB-CG2	5.99	120.79	112.40
1	B	129	THR	OG1-CB-CG2	5.67	123.05	110.00

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	40	THR	CB
1	A	159	THR	CB
1	B	107	THR	CB
1	B	110	THR	CB
1	B	129	THR	CB

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1113	1076	0	22	1
1	B	1113	1076	0	20	1
2	A	254	0	0	16	3
2	B	212	0	0	8	4
All	All	2692	2152	0	42	5

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 10.

All (42) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:80:MET:HE2	1:B:158:THR:CG2	1.87	1.03
1:A:75:ASN:OD1	2:A:2168:HOH:O	1.84	0.95
1:A:44:THR:HB	2:A:2012:HOH:O	1.70	0.91
1:A:75:ASN:HB2	2:A:2164:HOH:O	1.72	0.89
1:A:75:ASN:HB2	2:A:2160:HOH:O	1.69	0.89
1:B:80:MET:HE2	1:B:158:THR:HG22	1.57	0.85
1:A:29:ASP:OD1	2:A:2043:HOH:O	1.94	0.85
1:A:25:PRO:HG3	2:A:2252:HOH:O	1.79	0.82
1:B:104:LYS:O	1:B:107:THR:HG23	1.82	0.80
1:A:25:PRO:CG	2:A:2252:HOH:O	2.29	0.79
1:B:80:MET:HE2	1:B:158:THR:HG23	1.64	0.78
1:B:87:ASP:HB2	1:B:110:THR:HG21	1.69	0.75
1:A:31:ASP:OD1	2:A:2048:HOH:O	2.09	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:110:THR:HG22	2:B:2143:HOH:O	1.93	0.69
1:B:88:SER:OG	1:B:110:THR:HG23	1.93	0.68
1:A:40:THR:HG22	2:A:2084:HOH:O	1.98	0.63
1:B:27:ASN:HB3	2:B:2028:HOH:O	2.00	0.61
1:B:94:MET:HE1	2:B:2102:HOH:O	2.02	0.59
1:B:53:LYS:HE3	1:B:98:LYS:O	2.02	0.58
1:A:44:THR:CA	2:A:2012:HOH:O	2.51	0.57
1:A:44:THR:O	2:A:2012:HOH:O	2.17	0.56
1:B:16:ASP:N	2:B:2001:HOH:O	2.39	0.55
1:B:33:LYS:HE3	2:B:2044:HOH:O	2.07	0.54
1:B:25:PRO:HG3	2:B:2018:HOH:O	2.08	0.53
1:A:27:ASN:OD1	1:A:159:THR:HG23	2.08	0.53
1:B:33:LYS:H	1:B:36:SER:HG	1.58	0.52
1:A:75:ASN:HD21	1:A:75:ASN:C	2.13	0.51
1:A:159:THR:HG22	2:B:2080:HOH:O	2.11	0.50
1:A:44:THR:CB	2:A:2012:HOH:O	2.41	0.49
1:A:24:ILE:HB	1:A:40:THR:HG23	1.96	0.47
1:A:75:ASN:CB	2:A:2160:HOH:O	2.46	0.45
1:B:33:LYS:HE2	2:B:2036:HOH:O	2.16	0.45
1:B:34:LEU:O	1:B:35:ASN:HB2	2.16	0.45
1:B:28:TYR:HH	1:B:36:SER:HB3	1.80	0.45
1:A:25:PRO:HG2	2:A:2252:HOH:O	2.07	0.45
1:A:75:ASN:CB	2:A:2164:HOH:O	2.45	0.45
1:A:53:LYS:HE3	1:A:98:LYS:O	2.16	0.45
1:A:125:TYR:HB3	1:A:130:TYR:HH	1.82	0.43
1:B:72:ASN:HB2	1:B:130:TYR:CE2	2.53	0.43
1:B:80:MET:HE1	1:B:80:MET:HB2	1.65	0.42
1:B:80:MET:HE3	1:B:130:TYR:HB3	2.02	0.41
1:A:44:THR:C	2:A:2012:HOH:O	2.58	0.41

All (5) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:94:MET:CE	2:B:2211:HOH:O[3_544]	1.82	0.38
2:A:2193:HOH:O	2:A:2219:HOH:O[3_555]	1.85	0.35
2:A:2103:HOH:O	2:B:2053:HOH:O[3_554]	1.87	0.33
2:A:2243:HOH:O	2:B:2143:HOH:O[4_445]	1.90	0.30
1:A:104:LYS:CD	2:B:2199:HOH:O[4_545]	1.98	0.22

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	144/161 (89%)	143 (99%)	1 (1%)	0	100	100
1	B	144/161 (89%)	141 (98%)	3 (2%)	0	100	100
All	All	288/322 (89%)	284 (99%)	4 (1%)	0	100	100

There are no Ramachandran outliers to report.

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 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	126/140 (90%)	124 (98%)	2 (2%)	75	37
1	B	126/140 (90%)	124 (98%)	2 (2%)	75	37
All	All	252/280 (90%)	248 (98%)	4 (2%)	75	37

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	34	LEU
1	A	75	ASN
1	B	33	LYS
1	B	129	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	37	GLN
1	A	74	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	79	ASN
1	A	106	ASN
1	B	65	GLN
1	B	74	ASN
1	B	141	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

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 ⓘ

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	146/161 (90%)	0.15	4 (2%) 52 51	5, 11, 23, 30	0
1	B	146/161 (90%)	0.18	10 (6%) 17 17	6, 11, 25, 32	0
All	All	292/322 (90%)	0.17	14 (4%) 29 30	5, 11, 24, 32	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	144	LYS	5.6
1	B	143	ASN	4.9
1	A	141	ASN	4.6
1	B	142	ASP	4.5
1	A	142	ASP	3.6
1	B	145	GLN	3.5
1	A	145	GLN	3.2
1	B	122	ASN	2.9
1	B	141	ASN	2.7
1	B	73	VAL	2.6
1	B	33	LYS	2.5
1	A	146	ARG	2.5
1	B	147	THR	2.1
1	B	74	ASN	2.0

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.