Full wwPDB X-ray Structure Validation Report (i)

Oct 11, 2021 – 08:45 AM EDT

PDB ID	:	2QWU
Title	:	Crystal structure of F. tularensis pathogenicity island protein C
Authors	:	Sun, P.; Austin, B.P.; Schubot, F.D.; Waugh, D.S.
Deposited on	:	2007-08-10
Resolution	:	1.65 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

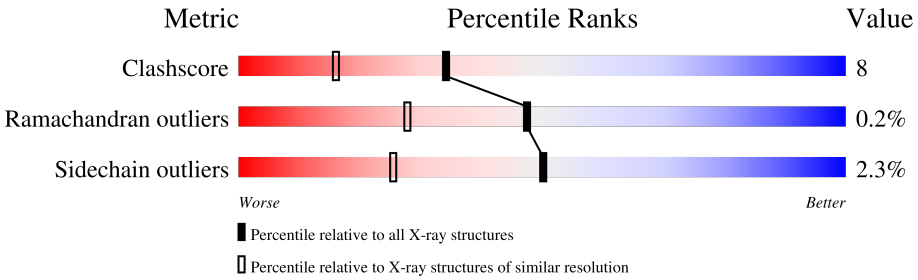
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.65 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	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)

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 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	211	84% 13% • •
1	B	211	80% 16% • •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3534 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 Intracellular growth locus, subunit C.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	207	Total	C	N	O	S	Se	0	4	0
			1547	962	259	314	4	8			
1	B	207	Total	C	N	O	S	Se	0	6	0
			1557	974	259	312	4	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	36	MSE	LEU	engineered mutation	UNP Q5NEC5
A	76	MSE	LEU	engineered mutation	UNP Q5NEC5
B	36	MSE	LEU	engineered mutation	UNP Q5NEC5
B	76	MSE	LEU	engineered mutation	UNP Q5NEC5

- Molecule 2 is water.

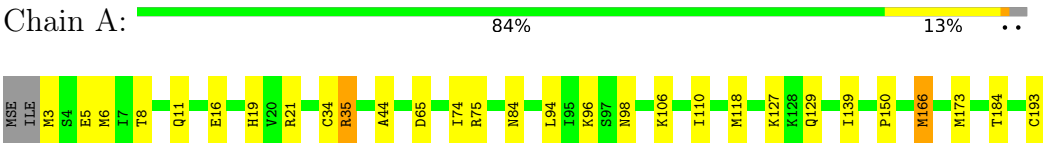
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	239	Total	O	0	0
			239	239		
2	B	191	Total	O	0	0
			191	191		

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.

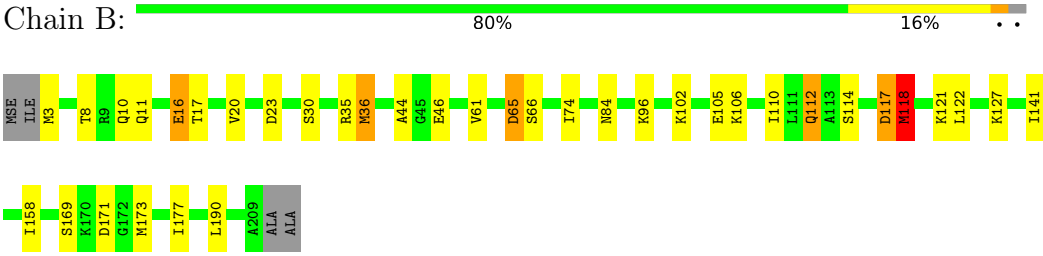
- Molecule 1: Intracellular growth locus, subunit C

Chain A:



- Molecule 1: Intracellular growth locus, subunit C

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	49.03Å 83.99Å 89.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.00 – 1.65 33.12 – 1.65	Depositor EDS
% Data completeness (in resolution range)	90.7 (35.00-1.65) 94.7 (33.12-1.65)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.95 (at 1.65Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.184 , 0.220 0.195 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	18.3	Xtriage
Anisotropy	0.556	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 35.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3534	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.86% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	0.84	2/1573 (0.1%)	0.89	1/2114 (0.0%)
1	B	0.84	2/1591 (0.1%)	0.89	0/2140
All	All	0.84	4/3164 (0.1%)	0.89	1/4254 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	166	MSE	SE-CE	-7.48	1.51	1.95
1	A	6	MSE	SE-CE	-6.69	1.55	1.95
1	B	36	MSE	SE-CE	-5.95	1.60	1.95
1	B	118	MSE	CG-SE	-5.41	1.77	1.95

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	35	ARG	NE-CZ-NH1	5.57	123.09	120.30

There are no chirality outliers.

There are no planarity outliers.

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	1547	0	1588	21	0
1	B	1557	0	1610	33	0
2	A	239	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	191	0	0	4	0
All	All	3534	0	3198	53	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.

All (53) 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:141[B]:ILE:HD11	1:B:190:LEU:HD13	1.57	0.86
1:B:46:GLU:HG2	1:B:106:LYS:HD3	1.64	0.78
1:A:44:ALA:HB1	1:A:110:ILE:HD12	1.69	0.74
1:B:96:LYS:HE2	1:B:171:ASP:O	1.91	0.70
1:A:5:GLU:OE1	2:A:378:HOH:O	2.11	0.68
1:B:3:MSE:O	1:B:3:MSE:HG3	1.94	0.68
1:A:3:MSE:HA	1:A:5:GLU:OE2	1.95	0.66
1:A:94:LEU:HD13	1:A:166:MSE:HE2	1.79	0.65
1:B:46:GLU:CG	1:B:106:LYS:HD3	2.27	0.65
1:B:121:LYS:HD2	1:B:147:GLU:HG2	1.80	0.64
1:B:11:GLN:HG3	1:B:16:GLU:HG2	1.80	0.63
1:A:106:LYS:HE2	1:A:110:ILE:HD11	1.81	0.63
1:B:65:ASP:CG	1:B:66:SER:N	2.51	0.60
1:B:44:ALA:HB2	1:B:118:MSE:HE3	1.82	0.59
1:A:16:GLU:OE1	2:A:286:HOH:O	2.16	0.59
1:B:169:SER:OG	1:B:171:ASP:OD1	2.17	0.58
1:B:30:SER:HB2	1:B:61[B]:VAL:HG11	1.86	0.57
1:B:96:LYS:HG3	1:B:173:MSE:HG2	1.87	0.55
1:A:94:LEU:HD13	1:A:166:MSE:CE	2.37	0.55
1:B:16:GLU:HG3	1:B:17:THR:N	2.21	0.55
1:B:65:ASP:CG	1:B:66:SER:H	2.09	0.54
1:B:141[B]:ILE:HD11	1:B:190:LEU:CD1	2.32	0.54
1:B:84:ASN:ND2	2:B:381:HOH:O	2.41	0.53
1:B:110:ILE:HG22	1:B:110:ILE:O	2.09	0.52
1:A:8:THR:H	1:A:11:GLN:NE2	2.08	0.52
1:A:5:GLU:HB3	1:A:195:ALA:HA	1.91	0.51
1:A:8:THR:H	1:A:11:GLN:HE21	1.59	0.51
1:B:122:LEU:O	1:B:145:ASN:HA	2.11	0.51
1:B:44:ALA:HA	2:B:265:HOH:O	2.12	0.49
1:B:102:LYS:O	1:B:106:LYS:HG3	2.12	0.49
1:A:127:LYS:HE2	1:B:127:LYS:HE3	1.95	0.49
1:B:20[B]:VAL:HG11	1:B:36:MSE:SE	2.63	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:121:LYS:CD	1:B:147:GLU:HG2	2.44	0.47
1:A:5:GLU:O	1:A:5:GLU:HG2	2.15	0.46
1:A:65:ASP:N	1:A:65:ASP:OD1	2.47	0.46
1:B:66:SER:HB3	2:B:380:HOH:O	2.14	0.46
1:A:34:CYS:HA	1:A:129:GLN:O	2.16	0.46
1:B:112:GLN:H	1:B:112:GLN:HG2	1.58	0.46
1:B:3:MSE:O	1:B:3:MSE:CG	2.62	0.45
1:B:74:ILE:CD1	1:B:177[B]:ILE:HD12	2.46	0.45
1:B:158:ILE:O	1:B:158:ILE:HG23	2.17	0.44
1:A:19:HIS:CE1	1:A:21:ARG:HH21	2.36	0.44
1:A:118:MSE:O	1:A:150:PRO:HD3	2.17	0.44
1:A:139:ILE:O	1:A:193:CYS:HB3	2.17	0.44
1:A:96:LYS:HG3	1:A:173:MSE:HG2	1.99	0.44
1:A:74:ILE:HG22	1:A:75:ARG:HG3	1.99	0.43
1:A:84:ASN:HB3	1:A:184:THR:OG1	2.19	0.43
1:A:5:GLU:O	1:A:5:GLU:CG	2.67	0.42
1:B:10:GLN:HG3	2:B:329:HOH:O	2.18	0.42
1:B:106:LYS:HB3	1:B:106:LYS:HE2	1.86	0.42
1:B:114:SER:HB2	1:B:117:ASP:HB2	2.01	0.41
1:B:8:THR:O	1:B:11:GLN:HB3	2.22	0.40

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	209/211 (99%)	202 (97%)	7 (3%)	0	100	100
1	B	211/211 (100%)	205 (97%)	5 (2%)	1 (0%)	29	11
All	All	420/422 (100%)	407 (97%)	12 (3%)	1 (0%)	47	28

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	65	ASP

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	179/168 (106%)	177 (99%)	2 (1%)	73 57
1	B	181/168 (108%)	175 (97%)	6 (3%)	38 12
All	All	360/336 (107%)	352 (98%)	8 (2%)	50 27

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	35	ARG
1	A	98	ASN
1	B	16	GLU
1	B	23	ASP
1	B	35	ARG
1	B	112	GLN
1	B	117	ASP
1	B	118	MSE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	11	GLN
1	A	19	HIS
1	A	112	GLN
1	B	84	ASN

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 monosaccharides 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

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.