wwPDB X-ray Structure Validation Summary Report (i)

Sep 23, 2023 – 05:07 PM EDT

PDB ID	:	5TS9
Title	:	Crystal structure of Chorismate mutase from Burkholderia phymatum
Authors	:	Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on	:	2016-10-28
Resolution	:	1.95 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

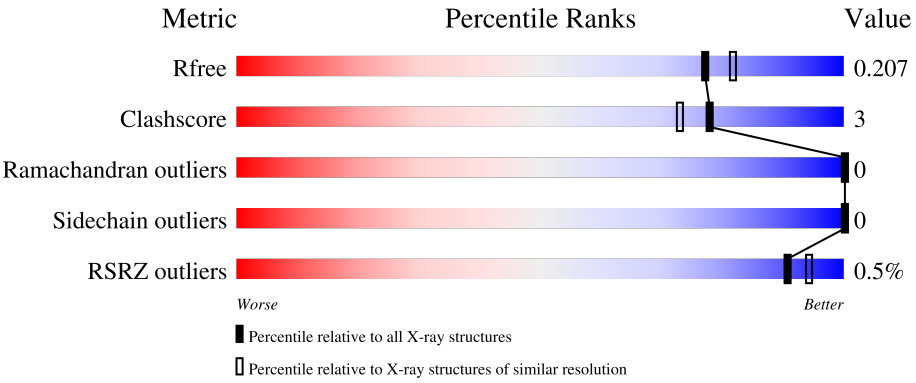
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.95 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.



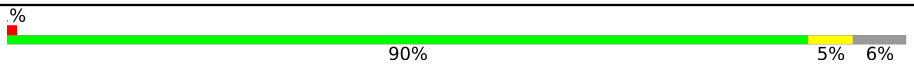
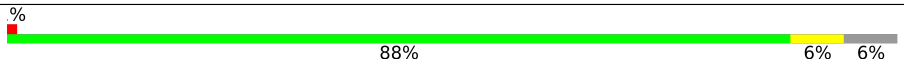
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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% 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	173	2%, 91%, 5%
1	B	173	86%, 8%, 7%
1	C	173	89%, 7%
1	D	173	87%, 6%, 7%
1	E	173	86%, 8%, 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	173	
1	G	173	
1	H	173	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11348 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 Chorismate mutase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	164	Total	C	N	O	S	0	11	0
			1258	782	224	247	5			
1	B	161	Total	C	N	O	S	0	12	0
			1259	783	222	249	5			
1	C	161	Total	C	N	O	S	0	12	0
			1265	786	227	247	5			
1	D	161	Total	C	N	O	S	0	9	0
			1245	774	222	244	5			
1	E	161	Total	C	N	O	S	0	8	0
			1242	771	225	241	5			
1	F	163	Total	C	N	O	S	0	12	0
			1277	793	231	248	5			
1	G	161	Total	C	N	O	S	0	13	0
			1271	791	223	251	6			
1	H	163	Total	C	N	O	S	0	6	0
			1236	766	225	240	5			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	10	MET	-	initiating methionine	UNP B2JYH9
A	11	ALA	-	expression tag	UNP B2JYH9
A	12	HIS	-	expression tag	UNP B2JYH9
A	13	HIS	-	expression tag	UNP B2JYH9
A	14	HIS	-	expression tag	UNP B2JYH9
A	15	HIS	-	expression tag	UNP B2JYH9
A	16	HIS	-	expression tag	UNP B2JYH9
A	17	HIS	-	expression tag	UNP B2JYH9
B	10	MET	-	initiating methionine	UNP B2JYH9
B	11	ALA	-	expression tag	UNP B2JYH9
B	12	HIS	-	expression tag	UNP B2JYH9
B	13	HIS	-	expression tag	UNP B2JYH9
B	14	HIS	-	expression tag	UNP B2JYH9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	15	HIS	-	expression tag	UNP B2JYH9
B	16	HIS	-	expression tag	UNP B2JYH9
B	17	HIS	-	expression tag	UNP B2JYH9
C	10	MET	-	initiating methionine	UNP B2JYH9
C	11	ALA	-	expression tag	UNP B2JYH9
C	12	HIS	-	expression tag	UNP B2JYH9
C	13	HIS	-	expression tag	UNP B2JYH9
C	14	HIS	-	expression tag	UNP B2JYH9
C	15	HIS	-	expression tag	UNP B2JYH9
C	16	HIS	-	expression tag	UNP B2JYH9
C	17	HIS	-	expression tag	UNP B2JYH9
D	10	MET	-	initiating methionine	UNP B2JYH9
D	11	ALA	-	expression tag	UNP B2JYH9
D	12	HIS	-	expression tag	UNP B2JYH9
D	13	HIS	-	expression tag	UNP B2JYH9
D	14	HIS	-	expression tag	UNP B2JYH9
D	15	HIS	-	expression tag	UNP B2JYH9
D	16	HIS	-	expression tag	UNP B2JYH9
D	17	HIS	-	expression tag	UNP B2JYH9
E	10	MET	-	initiating methionine	UNP B2JYH9
E	11	ALA	-	expression tag	UNP B2JYH9
E	12	HIS	-	expression tag	UNP B2JYH9
E	13	HIS	-	expression tag	UNP B2JYH9
E	14	HIS	-	expression tag	UNP B2JYH9
E	15	HIS	-	expression tag	UNP B2JYH9
E	16	HIS	-	expression tag	UNP B2JYH9
E	17	HIS	-	expression tag	UNP B2JYH9
F	10	MET	-	initiating methionine	UNP B2JYH9
F	11	ALA	-	expression tag	UNP B2JYH9
F	12	HIS	-	expression tag	UNP B2JYH9
F	13	HIS	-	expression tag	UNP B2JYH9
F	14	HIS	-	expression tag	UNP B2JYH9
F	15	HIS	-	expression tag	UNP B2JYH9
F	16	HIS	-	expression tag	UNP B2JYH9
F	17	HIS	-	expression tag	UNP B2JYH9
G	10	MET	-	initiating methionine	UNP B2JYH9
G	11	ALA	-	expression tag	UNP B2JYH9
G	12	HIS	-	expression tag	UNP B2JYH9
G	13	HIS	-	expression tag	UNP B2JYH9
G	14	HIS	-	expression tag	UNP B2JYH9
G	15	HIS	-	expression tag	UNP B2JYH9
G	16	HIS	-	expression tag	UNP B2JYH9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	17	HIS	-	expression tag	UNP B2JYH9
H	10	MET	-	initiating methionine	UNP B2JYH9
H	11	ALA	-	expression tag	UNP B2JYH9
H	12	HIS	-	expression tag	UNP B2JYH9
H	13	HIS	-	expression tag	UNP B2JYH9
H	14	HIS	-	expression tag	UNP B2JYH9
H	15	HIS	-	expression tag	UNP B2JYH9
H	16	HIS	-	expression tag	UNP B2JYH9
H	17	HIS	-	expression tag	UNP B2JYH9

- Molecule 2 is water.

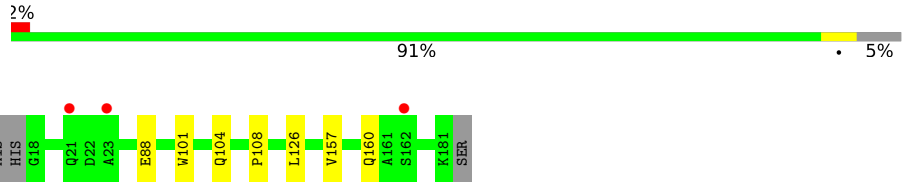
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	113	Total O 114 114	0	1
2	B	158	Total O 160 160	0	3
2	C	171	Total O 171 171	0	0
2	D	175	Total O 177 177	0	2
2	E	183	Total O 184 184	0	1
2	F	170	Total O 171 171	0	1
2	G	182	Total O 183 183	0	1
2	H	132	Total O 135 135	0	3

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 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.

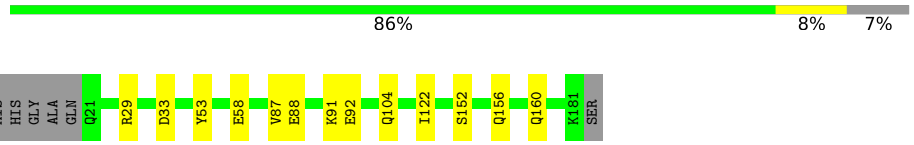
• Molecule 1: Chorismate mutase

Chain A:



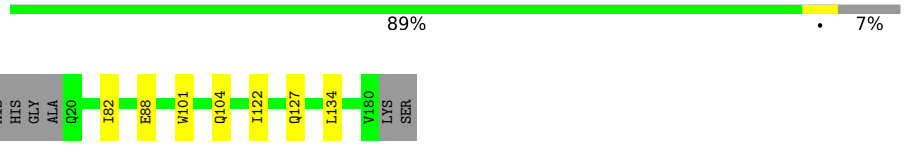
• Molecule 1: Chorismate mutase

Chain B:



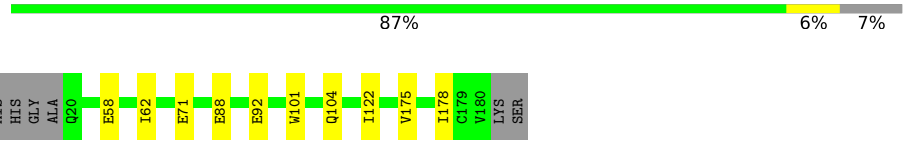
• Molecule 1: Chorismate mutase

Chain C:



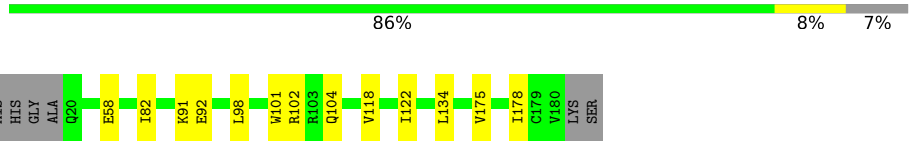
• Molecule 1: Chorismate mutase

Chain D:



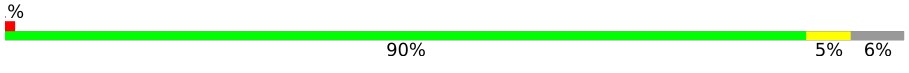
• Molecule 1: Chorismate mutase

Chain E:



• Molecule 1: Chorismate mutase

Chain F:

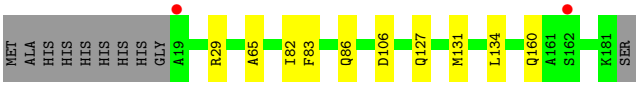


• Molecule 1: Chorismate mutase

Chain G:

• Molecule 1: Chorismate mutase

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	62.59Å 151.12Å 73.08Å 90.00° 90.84° 90.00°	Depositor
Resolution (Å)	47.88 – 1.95 47.88 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.7 (47.88-1.95) 99.7 (47.88-1.95)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.93 (at 1.95Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.156 , 0.206 0.156 , 0.207	Depositor DCC
R_{free} test set	2020 reflections (2.06%)	wwPDB-VP
Wilson B-factor (Å^2)	17.7	Xtriage
Anisotropy	0.279	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 57.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.170 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11348	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.02% 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	0.39	0/1307	0.53	0/1783
1	B	0.45	0/1308	0.56	0/1782
1	C	0.42	0/1320	0.62	0/1798
1	D	0.44	0/1288	0.57	0/1756
1	E	0.47	0/1285	0.58	0/1751
1	F	0.46	0/1326	0.59	0/1806
1	G	0.48	0/1320	0.56	0/1800
1	H	0.42	0/1270	0.55	0/1732
All	All	0.44	0/10424	0.57	0/14208

There are no bond length outliers.

There are no bond angle outliers.

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	1258	0	1244	5	0
1	B	1259	0	1256	12	0
1	C	1265	0	1271	7	0
1	D	1245	0	1245	9	0
1	E	1242	0	1245	10	0
1	F	1277	0	1283	8	0
1	G	1271	0	1267	6	0
1	H	1236	0	1220	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	114	0	0	2	0
2	B	160	0	0	2	0
2	C	171	0	0	3	0
2	D	177	0	0	1	0
2	E	184	0	0	1	0
2	F	171	0	0	3	0
2	G	183	0	0	1	0
2	H	135	0	0	2	0
All	All	11348	0	10031	59	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 3.

The worst 5 of 59 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:29:ARG:HE	1:B:160[B]:GLN:HE22	1.15	0.95
1:H:29:ARG:HE	1:H:160[B]:GLN:HE22	1.20	0.86
1:D:122:ILE:HD11	1:E:122:ILE:HD11	1.56	0.86
1:G:29:ARG:HE	1:G:160[B]:GLN:HE22	1.22	0.83
1:B:122:ILE:HD11	1:C:122:ILE:HD11	1.67	0.75

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	173/173 (100%)	170 (98%)	3 (2%)	0	100	100
1	B	170/173 (98%)	168 (99%)	2 (1%)	0	100	100
1	C	172/173 (99%)	170 (99%)	2 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	169/173 (98%)	168 (99%)	1 (1%)	0	100	100
1	E	168/173 (97%)	165 (98%)	3 (2%)	0	100	100
1	F	173/173 (100%)	171 (99%)	2 (1%)	0	100	100
1	G	171/173 (99%)	170 (99%)	1 (1%)	0	100	100
1	H	167/173 (96%)	167 (100%)	0	0	100	100
All	All	1363/1384 (98%)	1349 (99%)	14 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	131/136 (96%)	131 (100%)	0	100	100
1	B	135/136 (99%)	135 (100%)	0	100	100
1	C	136/136 (100%)	136 (100%)	0	100	100
1	D	133/136 (98%)	133 (100%)	0	100	100
1	E	132/136 (97%)	132 (100%)	0	100	100
1	F	136/136 (100%)	136 (100%)	0	100	100
1	G	136/136 (100%)	136 (100%)	0	100	100
1	H	127/136 (93%)	127 (100%)	0	100	100
All	All	1066/1088 (98%)	1066 (100%)	0	100	100

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

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	64	ASN
1	F	133	ASN
1	H	133	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	156	GLN
1	E	156	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 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

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, 95^{th} 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	164/173 (94%)	-0.03	3 (1%) 68 76	12, 26, 42, 62	0
1	B	161/173 (93%)	-0.33	0 100 100	11, 17, 31, 44	0
1	C	161/173 (93%)	-0.38	0 100 100	9, 16, 29, 51	0
1	D	161/173 (93%)	-0.39	0 100 100	10, 17, 32, 48	0
1	E	161/173 (93%)	-0.41	0 100 100	9, 16, 27, 45	0
1	F	163/173 (94%)	-0.31	1 (0%) 89 93	10, 17, 31, 57	0
1	G	161/173 (93%)	-0.31	0 100 100	10, 16, 29, 57	0
1	H	163/173 (94%)	-0.14	2 (1%) 79 84	11, 24, 40, 54	0
All	All	1295/1384 (93%)	-0.29	6 (0%) 91 94	9, 18, 36, 62	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	19	ALA	3.1
1	A	23	ALA	2.8
1	H	19	ALA	2.4
1	A	162	SER	2.3
1	H	162	SER	2.3

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 monosaccharides 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.