Full wwPDB X-ray Structure Validation Report ⓘ

Aug 19, 2023 – 10:21 PM EDT

PDB ID	:	2FP2
Title	:	Secreted Chorismate Mutase from Mycobacterium tuberculosis
Authors	:	Okvist, M.; Dey, R.; Sasso, S.; Grahn, E.; Kast, P.; Krengel, U.
Deposited on	:	2006-01-15
Resolution	:	1.64 Å(reported)

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

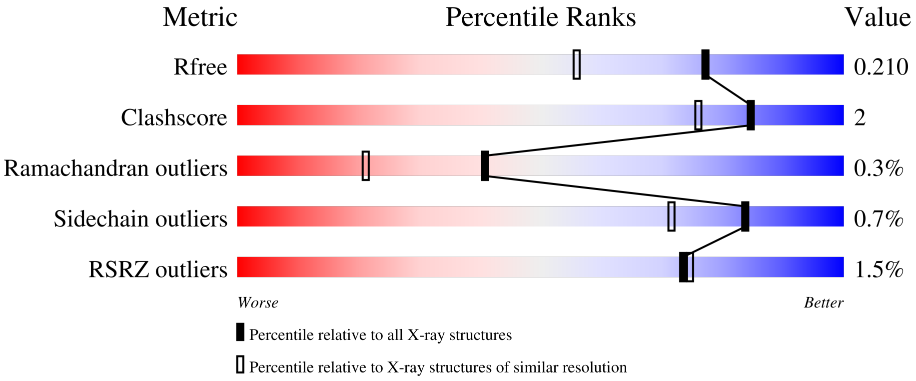
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance

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

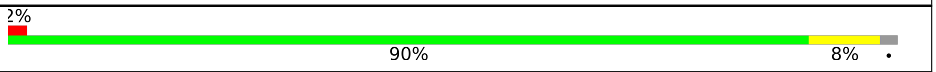
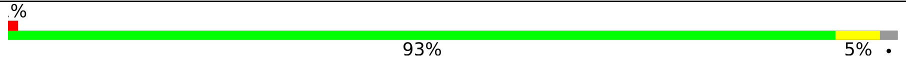
The reported resolution of this entry is 1.64 Å.

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	3122 (1.66-1.62)
Clashscore	141614	3268 (1.66-1.62)
Ramachandran outliers	138981	3215 (1.66-1.62)
Sidechain outliers	138945	3215 (1.66-1.62)
RSRZ outliers	127900	3079 (1.66-1.62)

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	166	
1	B	166	

2 Entry composition ⓘ

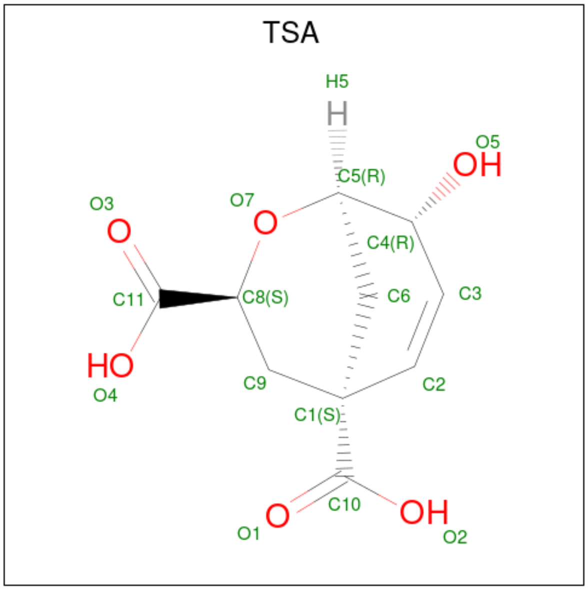
There are 3 unique types of molecules in this entry. The entry contains 3092 atoms, of which 0 are hydrogens and 0 are deuteriums.

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	162	Total 1340	C 836	N 243	O 258	S 3	0	12	0
1	B	163	Total 1320	C 817	N 237	O 263	S 3	0	7	0

- Molecule 2 is 8-HYDROXY-2-OXA-BICYCLO[3.3.1]NON-6-ENE-3,5-DICARBOXYLIC ACID (three-letter code: TSA) (formula: $C_{10}H_{12}O_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total 16	C 10	O 6	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	241	Total 241	O 241	0	4

Continued on next page...

Continued from previous page...

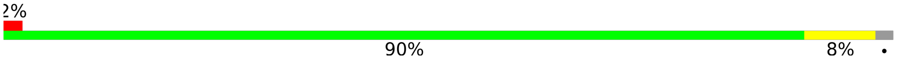
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	B	175	175	175	0	0

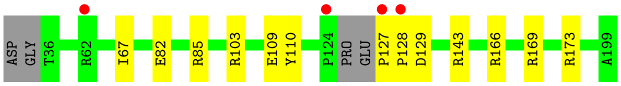
3 Residue-property plots ⓘ

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

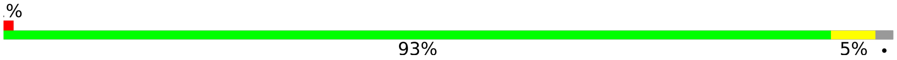
Chain A:

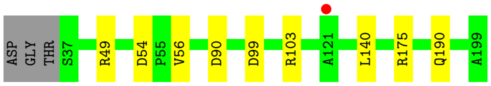




- Molecule 1: Chorismate mutase

Chain B:





4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	42.91Å 72.80Å 61.72Å 90.00° 103.99° 90.00°	Depositor
Resolution (Å)	38.92 – 1.64 38.87 – 1.64	Depositor EDS
% Data completeness (in resolution range)	99.5 (38.92-1.64) 99.5 (38.87-1.64)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.40 (at 1.63Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.170 , 0.203 0.176 , 0.210	Depositor DCC
R_{free} test set	2284 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	15.5	Xtriage
Anisotropy	0.343	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 41.7	EDS
L-test for twinning[2]	$< \|L\| >$ = 0.49, $< L^2 >$ = 0.32	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3092	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.92% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: TSA

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	0.72	0/1405	0.81	1/1906 (0.1%)
1	B	0.71	0/1367	0.77	3/1861 (0.2%)
All	All	0.71	0/2772	0.79	4/3767 (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

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	175	ARG	NE-CZ-NH1	7.32	123.96	120.30
1	B	175	ARG	NE-CZ-NH2	-6.02	117.29	120.30
1	B	90	ASP	CB-CG-OD1	5.60	123.34	118.30
1	A	143	ARG	NE-CZ-NH2	-5.02	117.79	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	127	PRO	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	1340	0	1341	8	0
1	B	1320	0	1272	4	0
2	B	16	0	10	1	0
3	A	241	0	0	4	0
3	B	175	0	0	0	0
All	All	3092	0	2623	12	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 2.

All (12) 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:56:VAL:HG11	2:B:500:TSA:H8	1.85	0.58
1:B:103:ARG:HH22	1:B:190:GLN:HE22	1.50	0.58
1:B:103:ARG:HH22	1:B:190:GLN:NE2	2.04	0.55
1:A:129:ASP:HB3	3:A:433:HOH:O	2.07	0.55
1:A:103[A]:ARG:NH2	3:A:416:HOH:O	2.33	0.52
1:A:166:ARG:NH1	1:A:169:ARG:HH12	2.07	0.51
1:A:82:GLU:OE1	1:A:85:ARG:NH1	2.45	0.50
1:A:67[B]:ILE:HD12	1:A:110:TYR:CD1	2.47	0.50
1:A:67[A]:ILE:HD12	1:A:109:GLU:HB3	1.95	0.47
1:A:173[A]:ARG:HD2	3:A:321:HOH:O	2.20	0.41
1:A:129:ASP:N	3:A:433:HOH:O	2.54	0.40
1:B:49:ARG:HA	1:B:140:LEU:HD13	2.02	0.40

There are no symmetry-related clashes.

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	171/166 (103%)	169 (99%)	1 (1%)	1 (1%)	25	8
1	B	168/166 (101%)	168 (100%)	0	0	100	100
All	All	339/332 (102%)	337 (99%)	1 (0%)	1 (0%)	41	21

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	128	PRO

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	149/139 (107%)	149 (100%)	0	100	100
1	B	144/139 (104%)	141 (98%)	3 (2%)	53	26
All	All	293/278 (105%)	290 (99%)	3 (1%)	84	59

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	54	ASP
1	B	99[A]	ASP
1	B	99[B]	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	190	GLN

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 ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	TSA	B	500	-	16,17,17	1.13	1 (6%)	16,26,26	1.53	5 (31%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	TSA	B	500	-	-	1/10/34/34	1/3/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	500	TSA	C9-C1	-2.19	1.52	1.55

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	500	TSA	O7-C8-C11	2.91	113.41	107.70
2	B	500	TSA	C9-C1-C10	-2.76	103.26	109.85
2	B	500	TSA	O4-C11-C8	2.34	119.70	113.03
2	B	500	TSA	O3-C11-C8	-2.23	117.30	122.57
2	B	500	TSA	C5-C4-C3	2.02	111.92	109.05

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	500	TSA	C9-C1-C10-O2

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	500	TSA	C1-C2-C3-C4-C5-C8-C9-O7

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	500	TSA	1	0

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 (i)

6.1 Protein, DNA and RNA chains (i)

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	162/166 (97%)	-0.16	4 (2%)	57	56	7, 12, 24, 39	0
1	B	163/166 (98%)	-0.17	1 (0%)	89	89	9, 16, 28, 32	0
All	All	325/332 (97%)	-0.17	5 (1%)	73	74	7, 13, 28, 39	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	128	PRO	4.1
1	A	127	PRO	3.0
1	B	121	ALA	2.9
1	A	124	PRO	2.2
1	A	62	ARG	2.1

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\AA^2$)	Q<0.9
2	TSA	B	500	16/16	0.94	0.08	11,13,16,17	0

6.5 Other polymers ⓘ

There are no such residues in this entry.