Full wwPDB X-ray Structure Validation Report (i)

Dec 9, 2020 – 11:03 AM GMT

PDB ID	:	7AGR
Title	:	Structure of the M624V-S726F mutant of AcylTransferase domain of Mycocerosic Acid Synthase from Mycobacterium tuberculosis soaked with Methyl-Malonyl Coenzyme A
Authors	:	Brison, Y.; Mourey, L.; Maveyraud, L.
Deposited on	:	2020-09-23
Resolution	:	2.80 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:

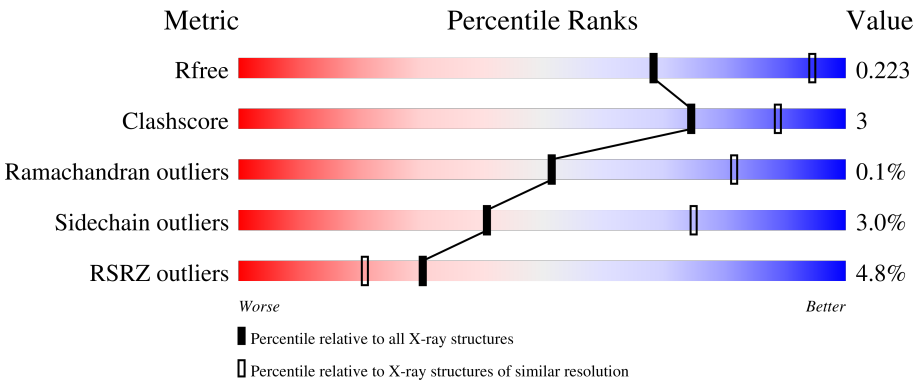
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.15.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.15.1

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 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	439	4% 89% 9% •
1	B	439	6% 90% 7% • •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6314 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 Mycocerosic acid synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	428	3137	1970	561	592	14	0	2	0
1	B	430	3160	1985	557	604	14	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

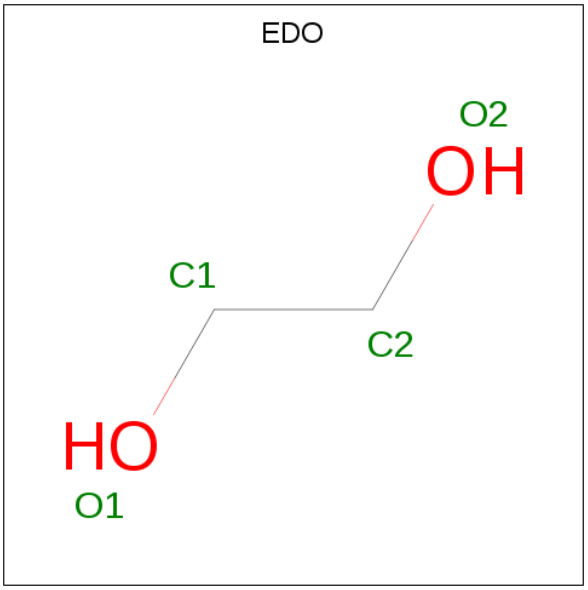
Chain	Residue	Modelled	Actual	Comment	Reference
A	435	MET	-	initiating methionine	UNP Q02251
A	436	GLY	-	expression tag	UNP Q02251
A	437	SER	-	expression tag	UNP Q02251
A	438	SER	-	expression tag	UNP Q02251
A	439	HIS	-	expression tag	UNP Q02251
A	440	HIS	-	expression tag	UNP Q02251
A	441	HIS	-	expression tag	UNP Q02251
A	442	HIS	-	expression tag	UNP Q02251
A	443	HIS	-	expression tag	UNP Q02251
A	444	HIS	-	expression tag	UNP Q02251
A	445	SER	-	expression tag	UNP Q02251
A	446	SER	-	expression tag	UNP Q02251
A	624	VAL	MET	engineered mutation	UNP Q02251
A	726	PHE	SER	engineered mutation	UNP Q02251
A	872	GLY	-	cloning artifact	UNP Q02251
A	873	SER	-	cloning artifact	UNP Q02251
B	435	MET	-	initiating methionine	UNP Q02251
B	436	GLY	-	expression tag	UNP Q02251
B	437	SER	-	expression tag	UNP Q02251
B	438	SER	-	expression tag	UNP Q02251
B	439	HIS	-	expression tag	UNP Q02251
B	440	HIS	-	expression tag	UNP Q02251
B	441	HIS	-	expression tag	UNP Q02251
B	442	HIS	-	expression tag	UNP Q02251
B	443	HIS	-	expression tag	UNP Q02251

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	444	HIS	-	expression tag	UNP Q02251
B	445	SER	-	expression tag	UNP Q02251
B	446	SER	-	expression tag	UNP Q02251
B	624	VAL	MET	engineered mutation	UNP Q02251
B	726	PHE	SER	engineered mutation	UNP Q02251
B	872	GLY	-	cloning artifact	UNP Q02251
B	873	SER	-	cloning artifact	UNP Q02251

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	B	1	4	2	2	0	0

- Molecule 3 is water.

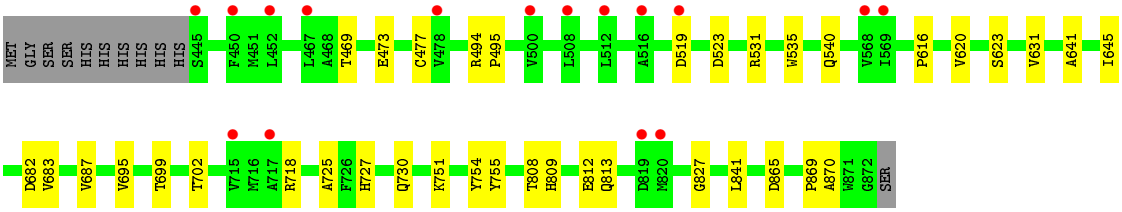
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	4	4	4	0	0
			Total	O		
3	B	9	9	9	0	0

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.

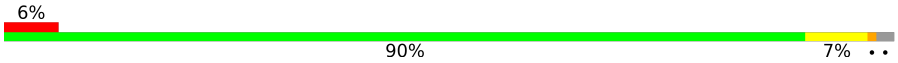
- Molecule 1: Mycocerosic acid synthase

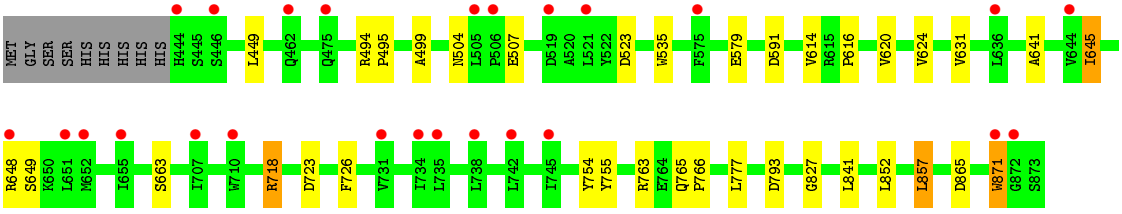
Chain A:



- Molecule 1: Mycocerosic acid synthase

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	151.80Å 152.23Å 115.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.70 – 2.80 48.70 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.70-2.80) 99.9 (48.70-2.80)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	1.09 (at 2.81Å)	Xtriage
Refinement program	BUSTER 2.10.3, REFMAC 5.8.0258	Depositor
R, R_{free}	0.196 , 0.219 0.203 , 0.223	Depositor DCC
R_{free} test set	1737 reflections (5.24%)	wwPDB-VP
Wilson B-factor (Å^2)	105.1	Xtriage
Anisotropy	0.341	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 78.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.012 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6314	wwPDB-VP
Average B, all atoms (Å^2)	115.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.54% 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 (i)

5.1 Standard geometry (i)

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

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	0.51	0/3201	0.64	0/4377
1	B	0.55	0/3222	0.66	0/4403
All	All	0.53	0/6423	0.65	0/8780

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 (i)

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	3137	0	3097	23	0
1	B	3160	0	3119	18	0
2	B	4	0	6	0	0
3	A	4	0	0	0	0
3	B	9	0	0	0	0
All	All	6314	0	6222	40	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.

All (40) 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:A:683:VAL:CG2	1:A:702:THR:HG23	2.11	0.81
1:A:469:THR:O	1:A:473:GLU:HG2	1.85	0.77
1:B:648:ARG:NH1	1:B:777:LEU:HB2	2.03	0.74
1:A:683:VAL:HG21	1:A:702:THR:HG23	1.70	0.73
1:B:857:LEU:H	1:B:857:LEU:HD12	1.53	0.72
1:A:687:VAL:HB	1:A:695:VAL:HB	1.74	0.70
1:A:623:SER:HB2	1:A:727:HIS:NE2	2.07	0.69
1:B:763:ARG:NH1	1:B:793:ASP:OD1	2.28	0.66
1:B:504:ASN:OD1	1:B:507:GLU:HB2	2.01	0.61
1:A:683:VAL:HG23	1:A:702:THR:CG2	2.31	0.61
1:B:620:VAL:HG22	1:B:755:TYR:HB2	1.83	0.60
1:A:683:VAL:CG2	1:A:702:THR:CG2	2.81	0.59
1:B:648:ARG:HH12	1:B:777:LEU:HB2	1.67	0.59
1:A:494:ARG:HG2	1:B:871:TRP:NE1	2.17	0.58
1:A:725:ALA:HB3	1:A:730:GLN:HG3	1.86	0.56
1:B:624:VAL:HG13	1:B:645:ILE:HD11	1.87	0.56
1:A:682:ASP:HB2	1:A:699:THR:HG23	1.89	0.55
1:A:699:THR:OG1	1:A:702:THR:HG22	2.07	0.55
1:B:591:ASP:HA	1:B:649:SER:HB2	1.90	0.54
1:B:449:LEU:HD22	1:B:499:ALA:HB1	1.90	0.54
1:A:620:VAL:HG22	1:A:755:TYR:HB2	1.91	0.53
1:A:865:ASP:O	1:A:865:ASP:OD2	2.26	0.53
1:A:808:THR:O	1:A:812:GLU:HG3	2.08	0.53
1:A:809[A]:HIS:CE1	1:A:813:GLN:HE21	2.27	0.52
1:A:494:ARG:HH22	1:A:869:PRO:HB2	1.76	0.51
1:A:494:ARG:HB3	1:A:495:PRO:HD2	1.94	0.50
1:A:683:VAL:HG23	1:A:702:THR:HG21	1.92	0.50
1:B:614:VAL:HG22	1:B:852:LEU:HD22	1.95	0.48
1:A:535:TRP:CD1	1:A:616:PRO:HB3	2.50	0.46
1:B:857:LEU:HD12	1:B:857:LEU:N	2.24	0.46
1:B:827:GLY:HA2	1:B:841:LEU:HD22	1.97	0.45
1:B:535:TRP:CD1	1:B:616:PRO:HB3	2.51	0.45
1:A:631:VAL:CG2	1:A:641:ALA:CB	2.95	0.45
1:A:682:ASP:OD2	1:A:702:THR:HG21	2.19	0.43
1:B:494:ARG:HB3	1:B:495:PRO:HD2	2.00	0.42
1:B:631:VAL:CG2	1:B:641:ALA:CB	2.97	0.42
1:B:765:GLN:HA	1:B:766:PRO:HD3	1.90	0.42
1:B:663:SER:HB3	1:B:718:ARG:HD3	2.04	0.40
1:A:827:GLY:HA2	1:A:841:LEU:HD22	2.03	0.40
1:A:631:VAL:CG2	1:A:641:ALA:HB3	2.52	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	428/439 (98%)	420 (98%)	7 (2%)	1 (0%)	47	78
1	B	428/439 (98%)	423 (99%)	5 (1%)	0	100	100
All	All	856/878 (98%)	843 (98%)	12 (1%)	1 (0%)	51	81

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	870	ALA

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	313/335 (93%)	304 (97%)	9 (3%)	42	76
1	B	318/335 (95%)	308 (97%)	10 (3%)	40	74
All	All	631/670 (94%)	612 (97%)	19 (3%)	41	75

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	477	CYS
1	A	519	ASP
1	A	523	ASP
1	A	531	ARG
1	A	540	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	645	ILE
1	A	718	ARG
1	A	751	LYS
1	A	754	TYR
1	B	523	ASP
1	B	579	GLU
1	B	645	ILE
1	B	718	ARG
1	B	723	ASP
1	B	726	PHE
1	B	754	TYR
1	B	857	LEU
1	B	865	ASP
1	B	871	TRP

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

Mol	Chain	Res	Type
1	A	475	GLN
1	A	528	HIS
1	A	730	GLN
1	B	528	HIS
1	B	730	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)

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	EDO	B	901	-	3,3,3	0.82	0	2,2,2	0.23	0

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	EDO	B	901	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	901	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

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 (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	428/439 (97%)	0.24	16 (3%)	41	31	91, 115, 155, 175	0
1	B	430/439 (97%)	0.43	25 (5%)	23	15	83, 109, 139, 181	0
All	All	858/878 (97%)	0.34	41 (4%)	30	21	83, 113, 147, 181	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	652	MET	4.0
1	B	738	LEU	4.0
1	A	450	PHE	3.6
1	B	731	VAL	3.6
1	B	444	HIS	3.5
1	B	871	TRP	3.4
1	B	707	ILE	3.3
1	B	636	LEU	3.1
1	A	715	VAL	3.1
1	B	742	LEU	3.1
1	B	644	VAL	3.1
1	B	651	LEU	2.9
1	B	655	ILE	2.8
1	B	505	LEU	2.8
1	B	648	ARG	2.8
1	A	452	LEU	2.6
1	B	710	TRP	2.6
1	A	519	ASP	2.5
1	A	500	VAL	2.5
1	B	506	PRO	2.4
1	B	475	GLN	2.4
1	A	569	ILE	2.4
1	A	820	MET	2.4
1	A	516	ALA	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	735	LEU	2.2
1	B	745	ILE	2.2
1	A	717	ALA	2.2
1	A	512	LEU	2.2
1	B	519	ASP	2.2
1	A	568	VAL	2.2
1	B	462	GLN	2.1
1	B	734	ILE	2.1
1	A	445	SER	2.1
1	A	478	VAL	2.1
1	B	446	SER	2.1
1	A	819	ASP	2.1
1	B	872	GLY	2.1
1	B	575	PHE	2.1
1	B	521	LEU	2.0
1	A	467	LEU	2.0
1	A	508	LEU	2.0

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(Å^2)	Q<0.9
2	EDO	B	901	4/4	0.77	0.25	111,112,113,113	0

6.5 Other polymers (i)

There are no such residues in this entry.