



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 06:07 pm BST

PDB ID : 5FJT
Title : N-acyl amino acid racemase from *Amycolatopsis* sp. Ts-1-60: G291D F323 mutant in complex with N-acetyl phenylalanine
Authors : Sanchez Carron, G.; Campopiano, D.; Grogan, G.
Deposited on : 2015-10-12
Resolution : 2.11 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 ⓘ symbol.

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 : 2.11
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) : 2.11

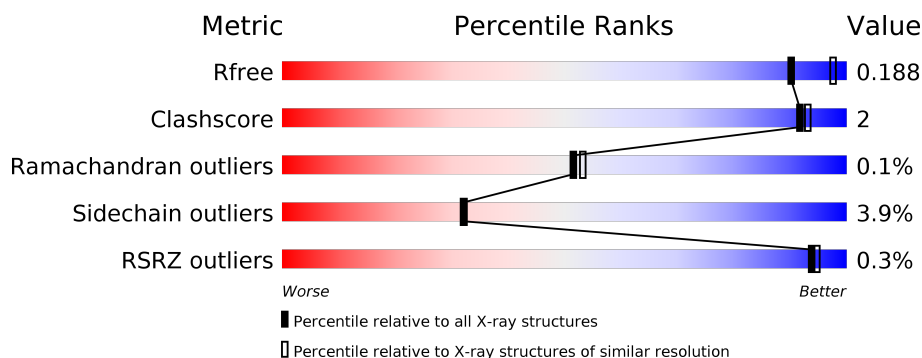
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.11 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	6241 (2.14-2.10)
Clashscore	141614	6778 (2.14-2.10)
Ramachandran outliers	138981	6705 (2.14-2.10)
Sidechain outliers	138945	6706 (2.14-2.10)
RSRZ outliers	127900	6112 (2.14-2.10)

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 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 $\leq 5\%$. 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	368	<div> <div style="width: 91%;"></div> <div style="width: 6%;"></div> <div style="width: 3%;"></div> <div style="width: 0%;"></div> <div style="width: 0%;"></div> </div> <div>91% 6% .</div>
1	B	368	<div> <div style="width: 91%;"></div> <div style="width: 7%;"></div> <div style="width: 2%;"></div> <div style="width: 0%;"></div> <div style="width: 0%;"></div> </div> <div>91% 7% .</div>
1	C	368	<div> <div style="width: 93%;"></div> <div style="width: 6%;"></div> <div style="width: 1%;"></div> <div style="width: 0%;"></div> <div style="width: 0%;"></div> </div> <div>93% 6% .</div>
1	D	368	<div> <div style="width: 92%;"></div> <div style="width: 6%;"></div> <div style="width: 2%;"></div> <div style="width: 0%;"></div> <div style="width: 0%;"></div> </div> <div>92% 6% ..</div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11624 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.

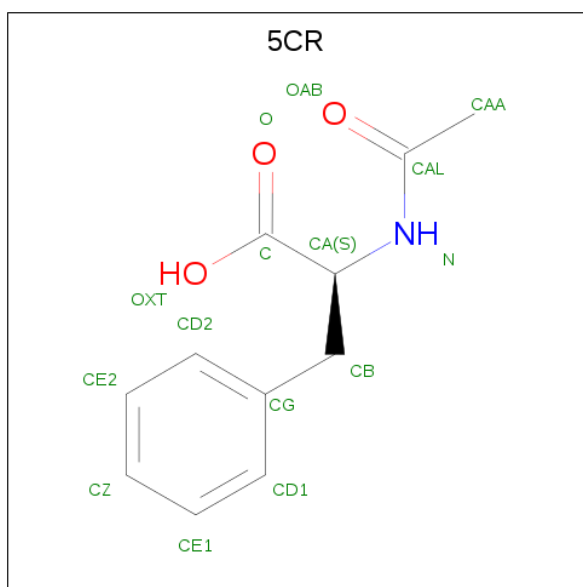
- Molecule 1 is a protein called O-SUCCINYLBENZOATE SYNTHASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	367	Total	C	N	O	S	0	0	0
			2740	1745	476	508	11			
1	B	367	Total	C	N	O	S	0	0	0
			2753	1752	479	510	12			
1	C	367	Total	C	N	O	S	0	0	0
			2741	1745	478	507	11			
1	D	368	Total	C	N	O	S	0	0	0
			2742	1744	475	511	12			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	291	ASP	GLY	engineered mutation	UNP Q44244
A	323	TYR	PHE	engineered mutation	UNP Q44244
B	291	ASP	GLY	engineered mutation	UNP Q44244
B	323	TYR	PHE	engineered mutation	UNP Q44244
C	291	ASP	GLY	engineered mutation	UNP Q44244
C	323	TYR	PHE	engineered mutation	UNP Q44244
D	291	ASP	GLY	engineered mutation	UNP Q44244
D	323	TYR	PHE	engineered mutation	UNP Q44244

- Molecule 2 is N-acetyl-L-phenylalanine (three-letter code: 5CR) (formula: C₁₁H₁₃NO₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	11	1	3		
2	B	1	Total	C	N	O	0	0
			15	11	1	3		
2	C	1	Total	C	N	O	0	0
			15	11	1	3		
2	D	1	Total	C	N	O	0	0
			15	11	1	3		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		
3	A	1	Total	Mg	0	0
			1	1		
3	D	1	Total	Mg	0	0
			1	1		
3	C	1	Total	Mg	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	170	Total	O	0	0
			170	170		

Continued on next page...

Continued from previous page...

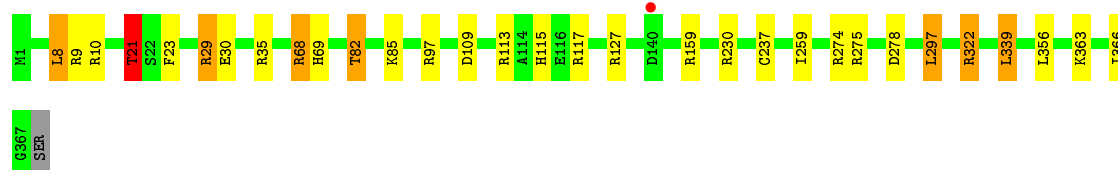
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	176	Total 176	O 176	0	0
4	C	114	Total 114	O 114	0	0
4	D	124	Total 124	O 124	0	0

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

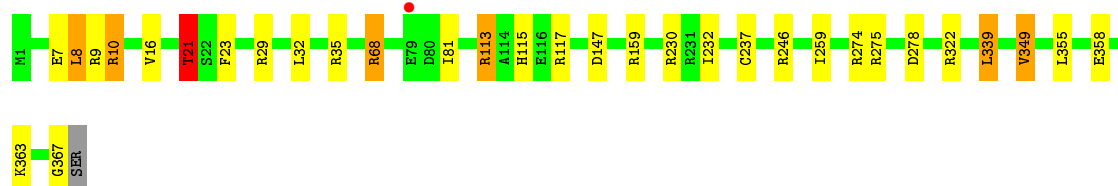
• Molecule 1: O-SUCCINYLBENZOATE SYNTHASE

Chain A: 



• Molecule 1: O-SUCCINYLBENZOATE SYNTHASE

Chain B: 



• Molecule 1: O-SUCCINYLBENZOATE SYNTHASE

Chain C: 



• Molecule 1: O-SUCCINYLBENZOATE SYNTHASE

Chain D: 



4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	216.20Å 216.20Å 258.69Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	151.68 – 2.11 75.84 – 2.11	Depositor EDS
% Data completeness (in resolution range)	99.9 (151.68-2.11) 99.9 (75.84-2.11)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.69 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.160 , 0.181 0.169 , 0.188	Depositor DCC
R_{free} test set	6737 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	31.9	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 28.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.019 for -2/3*h-1/3*k+2/3*l,-1/3*h-2/3*k-2/3*l,2/3*h-2/3*k+1/3*l 0.015 for -h,1/3*h-1/3*k+2/3*l,2/3*h+4/3*k+1/3*l 0.011 for -1/3*h+1/3*k-2/3*l,-k,-4/3*h-2/3*k+1/3*l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	11624	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.51% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: MG, 5CR

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	1.04	1/2796 (0.0%)	1.09	24/3814 (0.6%)
1	B	1.00	1/2809 (0.0%)	1.08	25/3829 (0.7%)
1	C	0.93	0/2797	1.05	20/3815 (0.5%)
1	D	0.99	0/2798	1.10	22/3818 (0.6%)
All	All	0.99	2/11200 (0.0%)	1.08	91/15276 (0.6%)

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	D	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	30	GLU	CD-OE1	7.32	1.33	1.25
1	B	21	THR	CB-CG2	-5.18	1.35	1.52

The worst 5 of 91 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	10	ARG	NE-CZ-NH2	-14.13	113.23	120.30
1	B	10	ARG	NE-CZ-NH2	-13.91	113.34	120.30
1	C	127	ARG	NE-CZ-NH1	12.53	126.56	120.30
1	D	127	ARG	NE-CZ-NH1	11.92	126.26	120.30
1	D	127	ARG	NE-CZ-NH2	-9.81	115.39	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	322	ARG	Sidechain

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	2740	0	2763	13	0
1	B	2753	0	2785	9	0
1	C	2741	0	2766	7	0
1	D	2742	0	2756	8	0
2	A	15	0	0	0	0
2	B	15	0	0	0	0
2	C	15	0	0	0	0
2	D	15	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	170	0	0	1	0
4	B	176	0	0	2	0
4	C	114	0	0	1	0
4	D	124	0	0	3	0
All	All	11624	0	11070	36	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.

The worst 5 of 36 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:113:ARG:NH1	4:B:2073:HOH:O	1.86	0.96
1:D:113:ARG:NH1	4:D:2053:HOH:O	1.88	0.89
1:D:274:ARG:NH1	4:D:2108:HOH:O	2.18	0.75
1:B:7:GLU:OE1	1:B:35:ARG:NH1	2.24	0.70
1:B:232:ILE:O	4:B:2140:HOH:O	2.12	0.67

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	365/368 (99%)	358 (98%)	7 (2%)	0	100	100
1	B	365/368 (99%)	360 (99%)	5 (1%)	0	100	100
1	C	365/368 (99%)	358 (98%)	6 (2%)	1 (0%)	41	40
1	D	366/368 (100%)	360 (98%)	5 (1%)	1 (0%)	41	40
All	All	1461/1472 (99%)	1436 (98%)	23 (2%)	2 (0%)	51	53

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	79	GLU
1	D	79	GLU

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	283/292 (97%)	273 (96%)	10 (4%)	36	37
1	B	286/292 (98%)	273 (96%)	13 (4%)	27	26
1	C	283/292 (97%)	275 (97%)	8 (3%)	43	46
1	D	284/292 (97%)	271 (95%)	13 (5%)	27	25
All	All	1136/1168 (97%)	1092 (96%)	44 (4%)	32	32

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	339	LEU
1	C	16	VAL
1	D	322	ARG
1	B	349	VAL
1	B	363	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	338	HIS
1	D	258	GLN
1	C	338	HIS
1	B	12	GLN
1	C	12	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 for Mogul analysis.

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	5CR	A	1368	3	12,15,15	1.99	3 (25%)	14,19,19	2.33	4 (28%)
2	5CR	C	1368	3	12,15,15	1.67	2 (16%)	14,19,19	1.95	5 (35%)
2	5CR	D	1369	3	12,15,15	1.53	2 (16%)	14,19,19	2.28	5 (35%)
2	5CR	B	1368	3	12,15,15	1.63	3 (25%)	14,19,19	1.83	3 (21%)

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	5CR	A	1368	3	-	1/8/12/12	0/1/1/1
2	5CR	C	1368	3	-	4/8/12/12	0/1/1/1
2	5CR	D	1369	3	-	3/8/12/12	0/1/1/1
2	5CR	B	1368	3	-	4/8/12/12	0/1/1/1

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1368	5CR	CB-CG	-4.19	1.41	1.51
2	C	1368	5CR	CA-N	3.91	1.51	1.46
2	D	1369	5CR	CB-CG	-3.88	1.42	1.51
2	A	1368	5CR	CA-N	3.57	1.51	1.46
2	C	1368	5CR	CB-CG	-3.48	1.42	1.51

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1368	5CR	CA-N-CAL	5.35	128.31	122.44
2	B	1368	5CR	CA-N-CAL	4.96	127.89	122.44
2	D	1369	5CR	CA-N-CAL	4.96	127.88	122.44
2	A	1368	5CR	CB-CA-N	4.81	117.74	109.01
2	C	1368	5CR	CA-N-CAL	4.36	127.22	122.44

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1368	5CR	C-CA-N-CAL
2	D	1369	5CR	C-CA-CB-CG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	C	1368	5CR	C-CA-N-CAL
2	D	1369	5CR	CB-CA-N-CAL
2	D	1369	5CR	C-CA-N-CAL

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	367/368 (99%)	-0.40	1 (0%) 94 95	20, 30, 50, 65	0
1	B	367/368 (99%)	-0.44	1 (0%) 94 95	21, 30, 48, 82	0
1	C	367/368 (99%)	-0.25	3 (0%) 86 88	22, 35, 57, 84	0
1	D	368/368 (100%)	-0.34	0 100 100	22, 35, 53, 65	0
All	All	1469/1472 (99%)	-0.36	5 (0%) 94 95	20, 32, 53, 84	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	39	PRO	2.8
1	C	40	ALA	2.6
1	B	79	GLU	2.3
1	A	140	ASP	2.3
1	C	18	PRO	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 carbohydrates 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	5CR	A	1368	15/15	0.92	0.21	31,50,98,101	0
2	5CR	D	1369	15/15	0.93	0.19	30,48,84,88	0
2	5CR	C	1368	15/15	0.94	0.19	31,51,80,85	0
3	MG	D	1370	1/1	0.95	0.12	38,38,38,38	0
2	5CR	B	1368	15/15	0.96	0.13	27,37,65,65	0
3	MG	A	1369	1/1	0.97	0.15	41,41,41,41	0
3	MG	C	1369	1/1	0.97	0.05	36,36,36,36	0
3	MG	B	1369	1/1	0.98	0.08	28,28,28,28	0

6.5 Other polymers [i](#)

There are no such residues in this entry.