



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 21, 2020 – 07:20 AM BST

PDB ID : 4B5W
Title : Crystal structures of divalent metal dependent pyruvate aldolase R70A mutant, HpaI, in complex with pyruvate
Authors : Coincon, M.; Wang, W.; Seah, S.Y.K.; Sygusch, J.
Deposited on : 2012-08-07
Resolution : 1.79 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 : 2.13.1
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) : 2.13.1

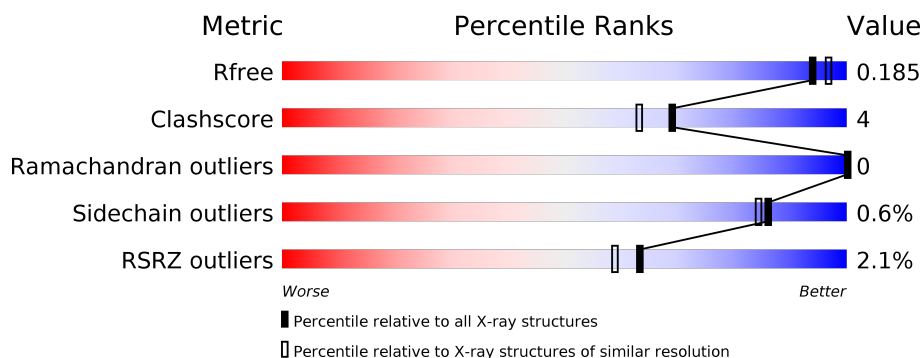
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.79 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 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	256	<div> <div>2%</div> <div> <div></div> <div>93%</div> <div>7%</div> </div> </div>
1	B	256	<div> <div>4%</div> <div> <div></div> <div>91%</div> <div>7%</div> <div>.</div> </div> </div>
1	C	256	<div> <div></div> <div> <div></div> <div>94%</div> <div>.</div> <div>.</div> </div> </div>
1	D	256	<div> <div>%</div> <div> <div></div> <div>92%</div> <div>6%</div> <div>.</div> </div> </div>
1	E	256	<div> <div>2%</div> <div> <div></div> <div>94%</div> <div>6%</div> </div> </div>
1	F	256	<div> <div>4%</div> <div> <div></div> <div>93%</div> <div>6%</div> <div>.</div> </div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 24975 atoms, of which 11597 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 4-HYDROXY-2-OXO-HEPTANE-1,7-DIOATE ALDOLASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	256	Total	C	H	N	O	S	0	0	0
			3868	1217	1943	339	363	6			
1	B	253	Total	C	H	N	O	S	0	0	0
			3834	1207	1926	336	359	6			
1	C	252	Total	C	H	N	O	S	0	0	0
			3817	1202	1918	334	357	6			
1	D	253	Total	C	H	N	O	S	0	0	0
			3831	1207	1923	336	359	6			
1	E	255	Total	C	H	N	O	S	0	0	0
			3858	1214	1938	338	362	6			
1	F	254	Total	C	H	N	O	S	0	0	0
			3844	1210	1931	337	360	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	70	ALA	ARG	engineered mutation	UNP B1IS70
B	70	ALA	ARG	engineered mutation	UNP B1IS70
C	70	ALA	ARG	engineered mutation	UNP B1IS70
D	70	ALA	ARG	engineered mutation	UNP B1IS70
E	70	ALA	ARG	engineered mutation	UNP B1IS70
F	70	ALA	ARG	engineered mutation	UNP B1IS70

- Molecule 2 is COBALT (II) ION (three-letter code: CO) (formula: Co).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total	Co	0	0
			1	1		
2	E	1	Total	Co	0	0
			1	1		
2	B	1	Total	Co	0	0
			1	1		

Continued on next page...

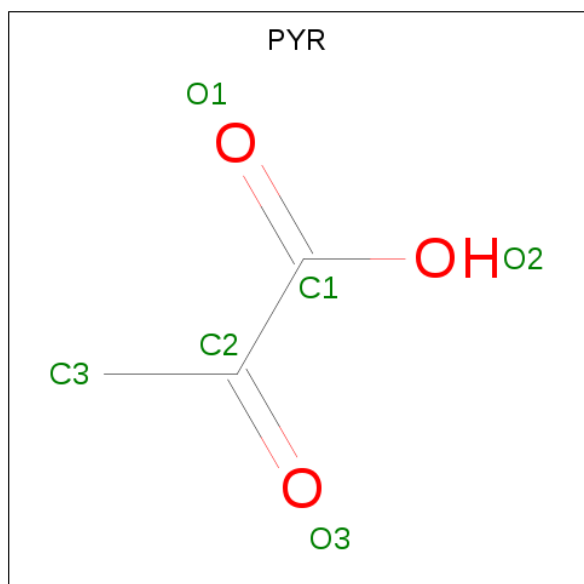
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	1	Total	Co	0	0
			1	1		
2	A	1	Total	Co	0	0
			1	1		
2	F	1	Total	Co	0	0
			1	1		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		
3	D	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		
3	F	1	Total	Ca	0	0
			1	1		
3	E	1	Total	Ca	0	0
			1	1		

- Molecule 4 is PYRUVIC ACID (three-letter code: PYR) (formula: C₃H₄O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			9	3	3	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total 9	C 3	H 3	O 3	0	0
4	C	1	Total 9	C 3	H 3	O 3	0	0
4	D	1	Total 9	C 3	H 3	O 3	0	0
4	E	1	Total 9	C 3	H 3	O 3	0	0
4	F	1	Total 9	C 3	H 3	O 3	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	365	Total 365	O 365	0	0
5	B	322	Total 322	O 322	0	0
5	C	290	Total 290	O 290	0	0
5	D	306	Total 306	O 306	0	0
5	E	306	Total 306	O 306	0	0
5	F	269	Total 269	O 269	0	0

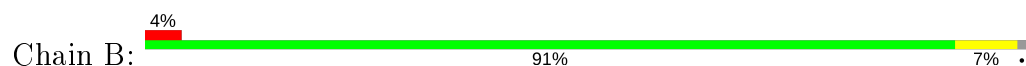
3 Residue-property plots [i](#)

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: 4-HYDROXY-2-OXO-HEPTANE-1,7-DIOATE ALDOLASE



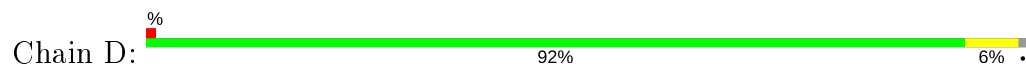
- Molecule 1: 4-HYDROXY-2-OXO-HEPTANE-1,7-DIOATE ALDOLASE



- Molecule 1: 4-HYDROXY-2-OXO-HEPTANE-1,7-DIOATE ALDOLASE



- Molecule 1: 4-HYDROXY-2-OXO-HEPTANE-1,7-DIOATE ALDOLASE



- Molecule 1: 4-HYDROXY-2-OXO-HEPTANE-1,7-DIOATE ALDOLASE



- Molecule 1: 4-HYDROXY-2-OXO-HEPTANE-1,7-DIOATE ALDOLASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	81.35Å 119.14Å 140.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.49 – 1.79 39.67 – 1.79	Depositor EDS
% Data completeness (in resolution range)	99.3 (38.49-1.79) 97.6 (39.67-1.79)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.71 (at 1.79Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.144 , 0.184 0.146 , 0.185	Depositor DCC
R_{free} test set	12762 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å ²)	17.4	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 48.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	24975	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.05% 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: PYR, CO, CA

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	0.47	0/1961	0.60	0/2674
1	B	0.47	0/1944	0.59	0/2650
1	C	0.47	0/1935	0.60	0/2638
1	D	0.45	0/1944	0.63	2/2650 (0.1%)
1	E	0.48	0/1956	0.62	0/2667
1	F	0.45	0/1949	0.60	0/2657
All	All	0.46	0/11689	0.61	2/15936 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	40	LEU	CA-CB-CG	-7.19	98.76	115.30
1	D	90	LEU	CA-CB-CG	5.24	127.35	115.30

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	1925	1943	1941	15	1
1	B	1908	1926	1924	20	1
1	C	1899	1918	1916	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1908	1923	1924	15	0
1	E	1920	1938	1936	19	1
1	F	1913	1931	1929	12	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	6	3	3	0	0
4	B	6	3	3	0	0
4	C	6	3	3	0	0
4	D	6	3	3	0	0
4	E	6	3	3	0	0
4	F	6	3	3	0	0
5	A	365	0	0	8	2
5	B	322	0	0	7	3
5	C	290	0	0	7	3
5	D	306	0	0	4	1
5	E	306	0	0	8	1
5	F	269	0	0	4	1
All	All	13378	11597	11588	85	7

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 4.

The worst 5 of 85 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:C:164:GLU:OE2	5:C:2167:HOH:O	1.88	0.91
1:E:204:GLU:OE2	5:E:2266:HOH:O	1.96	0.82
1:A:1:MET:N	5:A:2001:HOH:O	2.13	0.82
1:A:106:ARG:NH2	5:A:2209:HOH:O	2.22	0.73
1:C:1:MET:N	5:C:2001:HOH:O	2.23	0.71

The worst 5 of 7 symmetry-related close contacts are listed below. The label for Atom-2 includes

the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:C:2015:HOH:O	5:D:2234:HOH:O[3_555]	1.85	0.35
5:A:2291:HOH:O	5:F:2214:HOH:O[3_545]	2.01	0.19
5:B:2244:HOH:O	5:C:2234:HOH:O[3_645]	2.08	0.12
1:B:190:GLU:OE1	5:E:2192:HOH:O[2_564]	2.10	0.10
5:A:2001:HOH:O	5:B:2299:HOH:O[2_565]	2.12	0.08

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	254/256 (99%)	252 (99%)	2 (1%)	0	100	100
1	B	251/256 (98%)	246 (98%)	5 (2%)	0	100	100
1	C	250/256 (98%)	248 (99%)	2 (1%)	0	100	100
1	D	251/256 (98%)	247 (98%)	4 (2%)	0	100	100
1	E	253/256 (99%)	251 (99%)	2 (1%)	0	100	100
1	F	252/256 (98%)	246 (98%)	6 (2%)	0	100	100
All	All	1511/1536 (98%)	1490 (99%)	21 (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	196/196 (100%)	195 (100%)	1 (0%)	88	87
1	B	195/196 (100%)	194 (100%)	1 (0%)	88	87
1	C	194/196 (99%)	194 (100%)	0	100	100
1	D	195/196 (100%)	194 (100%)	1 (0%)	88	87
1	E	196/196 (100%)	196 (100%)	0	100	100
1	F	195/196 (100%)	191 (98%)	4 (2%)	53	42
All	All	1171/1176 (100%)	1164 (99%)	7 (1%)	86	84

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	51	GLN
1	F	253	GLN
1	F	122	LEU
1	B	1	MET
1	F	159	GLN

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

Mol	Chain	Res	Type
1	A	188	HIS

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 ⓘ

Of 17 ligands modelled in this entry, 11 are monoatomic - leaving 6 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$
4	PYR	F	1257	2	2,5,5	2.25	1 (50%)	2,6,6	1.24	0
4	PYR	E	1258	2	2,5,5	2.57	1 (50%)	2,6,6	0.74	0
4	PYR	C	1255	2	2,5,5	2.29	1 (50%)	2,6,6	0.56	0
4	PYR	D	1256	2	2,5,5	2.29	1 (50%)	2,6,6	0.66	0
4	PYR	B	1255	2	2,5,5	2.35	1 (50%)	2,6,6	1.09	0
4	PYR	A	1259	2	2,5,5	2.28	1 (50%)	2,6,6	1.13	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
4	PYR	F	1257	2	-	0/0/4/4	-
4	PYR	E	1258	2	-	0/0/4/4	-
4	PYR	C	1255	2	-	0/0/4/4	-
4	PYR	D	1256	2	-	0/0/4/4	-
4	PYR	B	1255	2	-	0/0/4/4	-
4	PYR	A	1259	2	-	0/0/4/4	-

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	1258	PYR	O3-C2	3.63	1.34	1.22
4	B	1255	PYR	O3-C2	3.33	1.33	1.22
4	D	1256	PYR	O3-C2	3.24	1.32	1.22
4	C	1255	PYR	O3-C2	3.23	1.32	1.22
4	A	1259	PYR	O3-C2	3.21	1.32	1.22

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	256/256 (100%)	-0.21	4 (1%) 72 68	8, 15, 32, 49	0
1	B	253/256 (98%)	-0.05	9 (3%) 42 37	9, 17, 38, 54	0
1	C	252/256 (98%)	-0.26	1 (0%) 92 90	11, 16, 31, 41	0
1	D	253/256 (98%)	-0.15	3 (1%) 79 76	10, 17, 34, 55	0
1	E	255/256 (99%)	-0.19	4 (1%) 72 68	9, 15, 33, 51	0
1	F	254/256 (99%)	0.03	11 (4%) 35 29	11, 19, 42, 61	0
All	All	1523/1536 (99%)	-0.14	32 (2%) 63 59	8, 17, 35, 61	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	254	ALA	7.1
1	E	255	THR	7.1
1	D	252	ALA	6.3
1	E	252	ALA	5.7
1	B	252	ALA	5.6

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
3	CA	A	1258	1/1	0.86	0.07	71,71,71,71	0
3	CA	E	1257	1/1	0.91	0.13	69,69,69,69	0
4	PYR	B	1255	6/6	0.93	0.10	15,19,24,24	0
4	PYR	F	1257	6/6	0.94	0.10	18,24,29,29	0
4	PYR	C	1255	6/6	0.95	0.10	12,17,22,23	0
3	CA	C	1254	1/1	0.95	0.11	41,41,41,41	0
4	PYR	A	1259	6/6	0.96	0.08	14,16,22,22	0
4	PYR	E	1258	6/6	0.97	0.06	15,19,27,27	0
4	PYR	D	1256	6/6	0.98	0.07	14,18,24,24	0
3	CA	F	1256	1/1	0.98	0.03	25,25,25,25	0
3	CA	D	1255	1/1	0.99	0.03	22,22,22,22	0
2	CO	F	1255	1/1	0.99	0.04	18,18,18,18	0
2	CO	B	1254	1/1	1.00	0.04	15,15,15,15	0
2	CO	E	1256	1/1	1.00	0.04	14,14,14,14	0
2	CO	A	1257	1/1	1.00	0.05	13,13,13,13	0
2	CO	C	1253	1/1	1.00	0.04	13,13,13,13	0
2	CO	D	1254	1/1	1.00	0.04	14,14,14,14	0

6.5 Other polymers [i](#)

There are no such residues in this entry.