Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 12:23 pm BST

PDB ID	:	5VIP
Title	:	Crystal structure of Pseudomonas malonate decarboxylase MdcD-MdcE hetero-dimer
Authors	:	Maderbocus, R.; Tong, L.
Deposited on	:	2017-04-17
Resolution	:	1.86 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 ⓘ symbol.

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

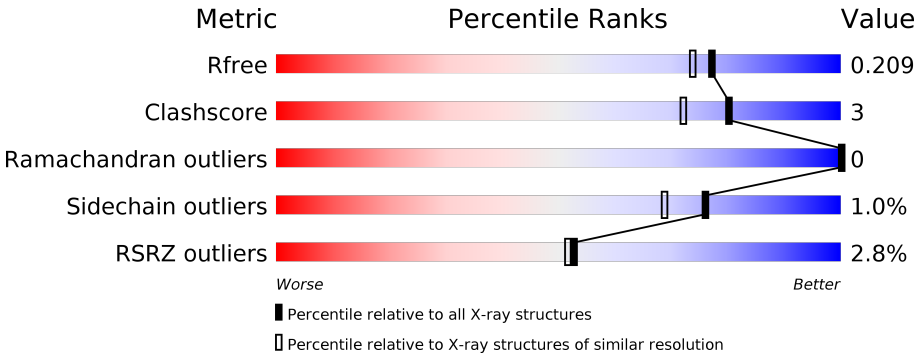
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 1.86 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 <=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	284	4% 84% 5% • 10%
1	C	284	2% 82% 7% 11%
2	B	287	2% 76% • 21%
2	D	287	2% 87% 7% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7693 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 MdcE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	255	Total	C	N	O	S	Se	0	0	0
			1905	1200	358	342	1	4			
1	C	254	Total	C	N	O	S	Se	0	0	0
			1887	1191	352	339	1	4			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-15	MSE	-	initiating methionine	UNP A0A0C6EV56
A	-14	GLY	-	expression tag	UNP A0A0C6EV56
A	-13	SER	-	expression tag	UNP A0A0C6EV56
A	-12	SER	-	expression tag	UNP A0A0C6EV56
A	-11	HIS	-	expression tag	UNP A0A0C6EV56
A	-10	HIS	-	expression tag	UNP A0A0C6EV56
A	-9	HIS	-	expression tag	UNP A0A0C6EV56
A	-8	HIS	-	expression tag	UNP A0A0C6EV56
A	-7	HIS	-	expression tag	UNP A0A0C6EV56
A	-6	HIS	-	expression tag	UNP A0A0C6EV56
A	-5	SER	-	expression tag	UNP A0A0C6EV56
A	-4	GLN	-	expression tag	UNP A0A0C6EV56
A	-3	ASP	-	expression tag	UNP A0A0C6EV56
A	-2	PRO	-	expression tag	UNP A0A0C6EV56
A	-1	ASN	-	expression tag	UNP A0A0C6EV56
A	0	SER	-	expression tag	UNP A0A0C6EV56
C	-15	MSE	-	initiating methionine	UNP A0A0C6EV56
C	-14	GLY	-	expression tag	UNP A0A0C6EV56
C	-13	SER	-	expression tag	UNP A0A0C6EV56
C	-12	SER	-	expression tag	UNP A0A0C6EV56
C	-11	HIS	-	expression tag	UNP A0A0C6EV56
C	-10	HIS	-	expression tag	UNP A0A0C6EV56
C	-9	HIS	-	expression tag	UNP A0A0C6EV56
C	-8	HIS	-	expression tag	UNP A0A0C6EV56
C	-7	HIS	-	expression tag	UNP A0A0C6EV56

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-6	HIS	-	expression tag	UNP A0A0C6EV56
C	-5	SER	-	expression tag	UNP A0A0C6EV56
C	-4	GLN	-	expression tag	UNP A0A0C6EV56
C	-3	ASP	-	expression tag	UNP A0A0C6EV56
C	-2	PRO	-	expression tag	UNP A0A0C6EV56
C	-1	ASN	-	expression tag	UNP A0A0C6EV56
C	0	SER	-	expression tag	UNP A0A0C6EV56

- Molecule 2 is a protein called MdcD.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	228	Total	C	N	O	S	Se	0	0	0
			1661	1041	298	318	3	1			
2	D	272	Total	C	N	O	S	Se	0	0	0
			2024	1266	370	384	3	1			

- Molecule 3 is water.

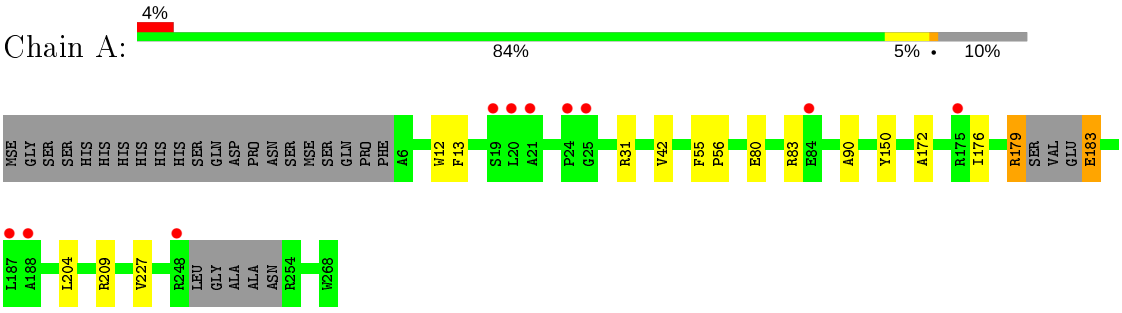
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	38	Total	O	0	0
			38	38		
3	B	51	Total	O	0	0
			51	51		
3	C	51	Total	O	0	0
			51	51		
3	D	76	Total	O	0	0
			76	76		

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.

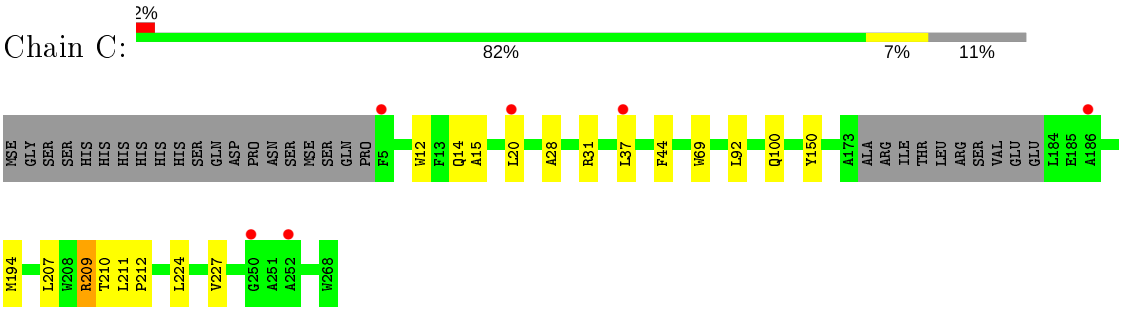
- Molecule 1: MdcE

Chain A:



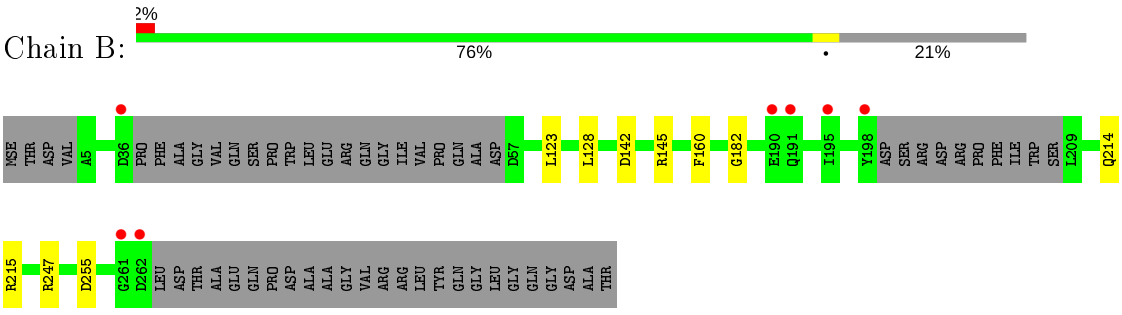
- Molecule 1: MdcE

Chain C:



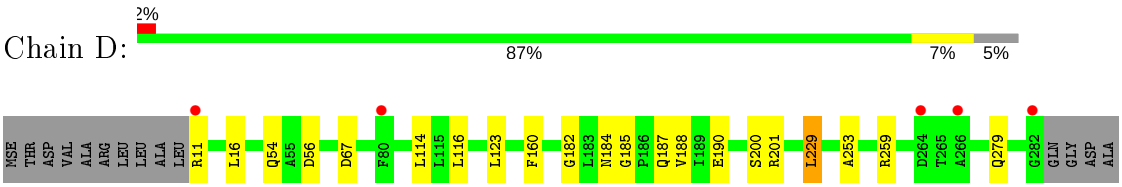
- Molecule 2: MdcD

Chain B:



- Molecule 2: MdcD

Chain D:



THR

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	56.31Å 60.08Å 91.28Å 89.87° 88.77° 68.80°	Depositor
Resolution (Å)	45.94 – 1.86 45.94 – 1.86	Depositor EDS
% Data completeness (in resolution range)	93.8 (45.94-1.86) 93.8 (45.94-1.86)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.09 (at 1.86Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.181 , 0.207 0.187 , 0.209	Depositor DCC
R_{free} test set	2012 reflections (2.28%)	wwPDB-VP
Wilson B-factor (Å^2)	20.3	Xtriage
Anisotropy	0.016	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 52.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.004 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7693	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.51% 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	0.34	0/1938	0.49	0/2627
1	C	0.39	0/1922	0.55	1/2608 (0.0%)
2	B	0.36	0/1676	0.52	0/2266
2	D	0.44	1/2053 (0.0%)	0.57	0/2782
All	All	0.39	1/7589 (0.0%)	0.53	1/10283 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	190	GLU	CD-OE1	-6.20	1.18	1.25

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	209	ARG	NE-CZ-NH1	6.03	123.31	120.30

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	1905	0	1918	13	0
1	C	1887	0	1892	13	0
2	B	1661	0	1674	6	0
2	D	2024	0	2023	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	38	0	0	0	0
3	B	51	0	0	0	0
3	C	51	0	0	0	0
3	D	76	0	0	1	0
All	All	7693	0	7507	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:209:ARG:HH11	1:A:209:ARG:HG2	1.37	0.90
2:B:247:ARG:NH2	2:B:255:ASP:OD2	2.15	0.80
1:C:209:ARG:HG2	1:C:209:ARG:HH11	1.49	0.78
1:A:80:GLU:OE1	1:A:83:ARG:NH1	2.19	0.70
2:D:200:SER:O	2:D:201:ARG:NH1	2.24	0.70
2:D:187:GLN:OE1	3:D:301:HOH:O	2.12	0.67
1:C:209:ARG:CG	1:C:209:ARG:HH11	2.07	0.66
2:D:259:ARG:NH2	2:D:279:GLN:O	2.23	0.64
1:A:209:ARG:NH1	1:A:209:ARG:HG2	2.14	0.58
1:C:15:ALA:HB1	1:C:224:LEU:HD11	1.86	0.58
1:A:176:ILE:HD11	2:B:123:LEU:HD22	1.86	0.56
1:A:209:ARG:CG	1:A:209:ARG:HH11	2.13	0.56
1:A:204:LEU:HD22	2:B:128:LEU:HD21	1.90	0.53
1:C:14:GLN:HG2	1:C:20:LEU:HA	1.92	0.51
1:A:172:ALA:O	1:A:176:ILE:HG23	2.11	0.51
2:B:214:GLN:OE1	2:B:215:ARG:HD3	2.12	0.49
1:C:100:GLN:NE2	2:D:184:ASN:OD1	2.40	0.49
1:A:209:ARG:CG	1:A:209:ARG:NH1	2.72	0.48
2:D:11:ARG:HH11	2:D:11:ARG:HG3	1.78	0.48
2:B:142:ASP:OD1	2:B:145:ARG:NH2	2.43	0.46
2:D:201:ARG:HA	2:D:201:ARG:HD3	1.61	0.46
2:B:160:PHE:HA	2:B:182:GLY:O	2.18	0.44
2:D:16:LEU:HD11	2:D:229:LEU:HD11	2.00	0.44
1:C:44:PHE:CE1	1:C:92:LEU:HD12	2.53	0.44
1:C:209:ARG:NH1	1:C:209:ARG:CG	2.71	0.44
2:D:54:GLN:HG2	2:D:56:ASP:OD1	2.18	0.44
1:C:207:LEU:HD13	1:C:210:THR:HG22	2.00	0.44
2:D:160:PHE:HA	2:D:182:GLY:O	2.18	0.44
1:C:12:TRP:CZ3	1:C:227:VAL:HG21	2.53	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:13:PHE:CE2	1:A:31:ARG:HB3	2.53	0.43
1:C:211:LEU:HD12	1:C:212:PRO:HD2	2.00	0.43
1:A:179:ARG:HG2	1:A:183:GLU:HB2	2.01	0.43
2:D:185:GLY:O	2:D:188:VAL:HG12	2.18	0.43
1:C:69:TRP:CE3	2:D:253:ALA:HA	2.54	0.42
1:A:55:PHE:HA	1:A:56:PRO:HD3	1.92	0.41
1:C:194:MSE:HG2	2:D:123:LEU:O	2.21	0.41
1:A:12:TRP:CZ3	1:A:227:VAL:HG21	2.56	0.41
1:C:28:ALA:O	1:C:31:ARG:NH1	2.42	0.40
1:A:42:VAL:HG12	1:A:90:ALA:HB3	2.03	0.40
2:D:114:LEU:HB3	2:D:116:LEU:HD21	2.03	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	249/284 (88%)	242 (97%)	7 (3%)	0	100	100
1	C	250/284 (88%)	245 (98%)	5 (2%)	0	100	100
2	B	222/287 (77%)	221 (100%)	1 (0%)	0	100	100
2	D	270/287 (94%)	267 (99%)	3 (1%)	0	100	100
All	All	991/1142 (87%)	975 (98%)	16 (2%)	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	181/203 (89%)	178 (98%)	3 (2%)	60	47
1	C	178/203 (88%)	176 (99%)	2 (1%)	73	65
2	B	159/207 (77%)	159 (100%)	0	100	100
2	D	197/207 (95%)	195 (99%)	2 (1%)	76	69
All	All	715/820 (87%)	708 (99%)	7 (1%)	76	69

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	150	TYR
1	A	179	ARG
1	A	183	GLU
1	C	37	LEU
1	C	150	TYR
2	D	67	ASP
2	D	229	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

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	251/284 (88%)	0.05	10 (3%) 38 36	13, 26, 45, 55	0
1	C	250/284 (88%)	-0.08	6 (2%) 59 57	10, 21, 40, 56	0
2	B	227/287 (79%)	0.02	7 (3%) 49 47	12, 20, 39, 53	0
2	D	271/287 (94%)	-0.22	5 (1%) 68 68	10, 18, 37, 48	0
All	All	999/1142 (87%)	-0.06	28 (2%) 53 52	10, 21, 41, 56	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	5	PHE	6.4
2	B	195	ILE	5.2
1	C	20	LEU	4.1
1	A	19	SER	3.7
1	A	25	GLY	3.6
1	A	175	ARG	3.6
2	B	261	GLY	3.4
1	A	24	PRO	3.1
2	D	266	ALA	3.1
2	D	264	ASP	2.9
2	B	190	GLU	2.9
1	A	20	LEU	2.8
1	C	250	GLY	2.8
2	D	11	ARG	2.7
2	B	262	ASP	2.7
1	A	187	LEU	2.7
1	A	21	ALA	2.6
1	C	186	ALA	2.5
1	C	252	ALA	2.5
2	D	282	GLY	2.4
1	A	84	GLU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	37	LEU	2.2
1	A	248	ARG	2.2
2	B	36	ASP	2.1
1	A	188	ALA	2.1
2	D	80	PHE	2.1
2	B	198	TYR	2.0
2	B	191	GLN	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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.