Full wwPDB X-ray Structure Validation Report (i)

Aug 29, 2023 – 12:57 PM EDT

PDB ID	:	3PVJ
Title	:	Crystal structure of the Fe(II)/alpha-ketoglutarate dependent taurine dioxygenase from Pseudomonas putida KT2440
Authors	:	Knauer, S.H.; Spiegelhauer, O.; Schwarzinger, S.; Haenzelmann, P.; Dobbek, H.
Deposited on	:	2010-12-07
Resolution	:	1.85 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

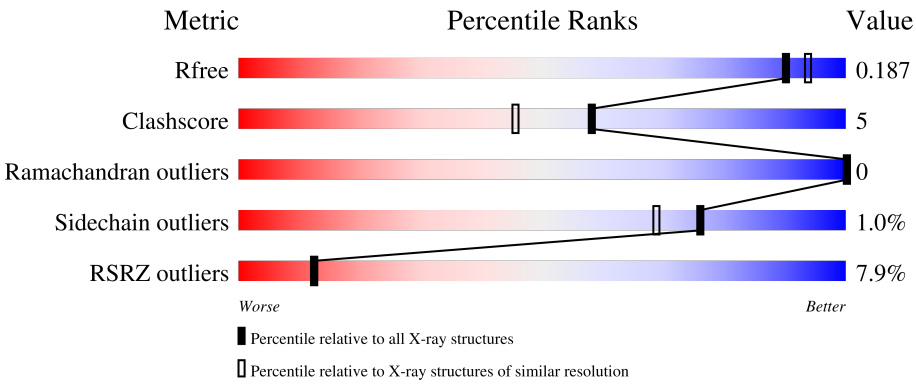
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.85 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 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	277	% 91% 8% •
1	B	277	12% 89% 10%
1	C	277	9% 87% 12% •
1	D	277	10% 89% 10%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 19060 atoms, of which 8831 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 Alpha-ketoglutarate-dependent taurine dioxygenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	275	4467	1425	2225	404	410	3	0	10	0
1	B	276	4446	1421	2207	406	410	2	0	8	0
1	C	276	4432	1417	2202	403	408	2	0	6	0
1	D	276	4424	1411	2197	407	407	2	0	5	0

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
2	A	2	2	2	0	0

- Molecule 3 is water.

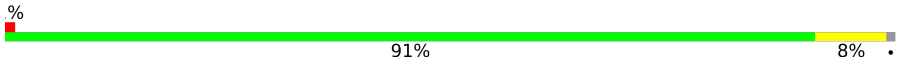
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	388	388	388	0	16
3	B	271	271	271	0	3
3	C	333	333	333	0	11
3	D	297	297	297	0	2

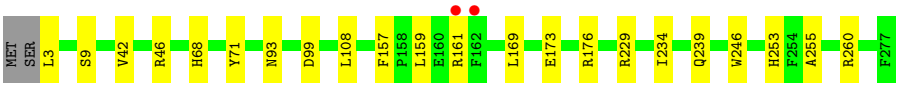
3 Residue-property plots (i)

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: Alpha-ketoglutarate-dependent taurine dioxygenase

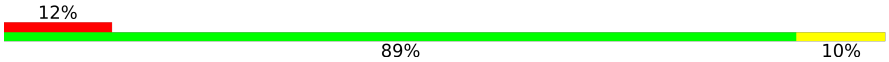
Chain A:

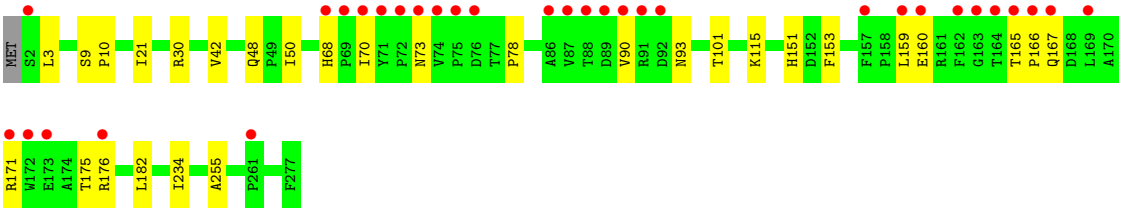




• Molecule 1: Alpha-ketoglutarate-dependent taurine dioxygenase

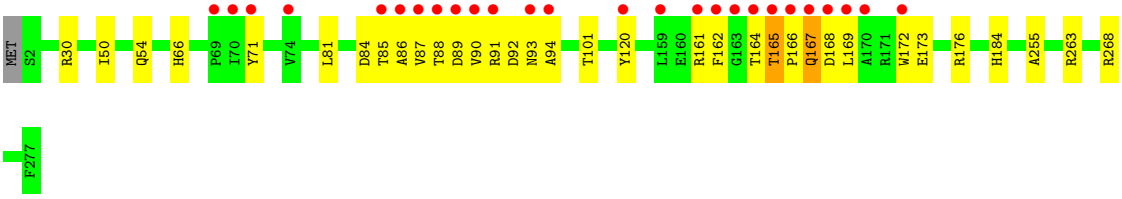
Chain B:





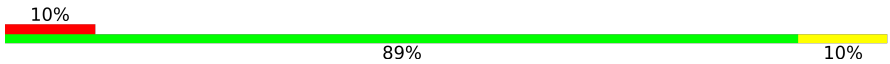
• Molecule 1: Alpha-ketoglutarate-dependent taurine dioxygenase

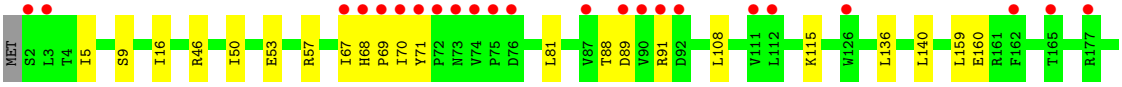
Chain C:

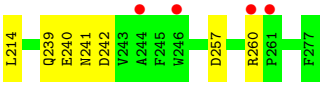


• Molecule 1: Alpha-ketoglutarate-dependent taurine dioxygenase

Chain D:






L214
Q239
E240
N241
D242
V243
A244
F245
W246
D257
R260
P261
F277

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	75.25Å 105.34Å 171.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.57 – 1.85 29.92 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.5 (28.57-1.85) 99.5 (29.92-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	2.82 (at 1.85Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.1_357)	Depositor
R, R_{free}	0.157 , 0.191 0.152 , 0.187	Depositor DCC
R_{free} test set	5806 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	20.7	Xtriage
Anisotropy	0.306	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 55.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	19060	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.94% 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: CL

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	0.46	0/2334	0.58	0/3189
1	B	0.40	0/2324	0.54	0/3176
1	C	0.45	0/2307	0.59	0/3153
1	D	0.40	0/2304	0.53	0/3148
All	All	0.43	0/9269	0.56	0/12666

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	2242	2225	2207	15	0
1	B	2239	2207	2184	25	0
1	C	2230	2202	2188	31	0
1	D	2227	2197	2176	18	0
2	A	2	0	0	0	0
3	A	388	0	0	3	0
3	B	271	0	0	4	0
3	C	333	0	0	5	0
3	D	297	0	0	1	0
All	All	10229	8831	8755	85	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 5.

All (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:169:LEU:O	1:C:173:GLU:HG2	1.93	0.68
1:C:164:THR:HG23	1:C:165:THR:HG22	1.75	0.68
1:B:153[A]:PHE:CE2	1:B:176:ARG:HA	2.29	0.67
1:A:93:ASN:HB3	1:A:255:ALA:HB3	1.78	0.66
1:C:94:ALA:O	3:C:631:HOH:O	2.14	0.65
1:A:159:LEU:HD23	1:B:176:ARG:NH1	2.12	0.64
1:C:93:ASN:HB3	1:C:255:ALA:HB3	1.78	0.64
1:D:69:PRO:HB2	1:D:70:ILE:HD12	1.83	0.60
1:C:71:TYR:HB2	1:C:81:LEU:HD11	1.84	0.59
1:A:42[B]:VAL:HG11	1:A:234:ILE:HD11	1.85	0.58
1:B:73:ASN:OD1	1:B:78:PRO:HA	2.03	0.58
1:B:151:HIS:HE1	3:B:520:HOH:O	1.88	0.57
1:B:90:VAL:O	1:B:90:VAL:HG23	2.03	0.57
1:D:5:ILE:HG12	1:D:16:ILE:CD1	2.35	0.56
1:C:165:THR:OG1	1:C:167:GLN:HG3	2.05	0.56
1:C:164:THR:HG23	1:C:165:THR:N	2.21	0.56
1:C:184:HIS:HD2	3:C:319:HOH:O	1.88	0.55
1:D:69:PRO:HB2	1:D:70:ILE:CD1	2.37	0.54
1:B:21:ILE:HG12	1:B:48:GLN:CG	2.38	0.53
1:A:99:ASP:OD1	1:A:253:HIS:HE1	1.91	0.53
1:A:176:ARG:NH2	1:B:159:LEU:HD22	2.23	0.53
1:D:71:TYR:HB2	1:D:81:LEU:HD11	1.91	0.53
1:B:101:THR:OG1	1:B:151:HIS:HD2	1.92	0.52
1:C:87:VAL:O	1:C:87:VAL:HG12	2.08	0.52
1:C:176:ARG:NH2	1:D:159:LEU:HD21	2.25	0.52
1:D:136:LEU:HD22	1:D:140:LEU:HD23	1.91	0.52
1:B:3:LEU:HD23	1:B:3:LEU:C	2.31	0.50
1:A:42[B]:VAL:HG23	1:A:246:TRP:HB3	1.93	0.50
1:B:153[A]:PHE:HE2	1:B:176:ARG:HA	1.74	0.50
1:C:86:ALA:HA	3:C:626:HOH:O	2.12	0.49
1:C:172:TRP:CH2	1:C:176:ARG:HD3	2.48	0.48
1:C:88:THR:CG2	1:C:89:ASP:N	2.76	0.48
1:D:53:GLU:HG3	1:D:57:ARG:NH1	2.29	0.48
1:D:67:ILE:HD12	1:D:67:ILE:N	2.29	0.48
1:B:165:THR:HB	1:B:166:PRO:HD2	1.96	0.48
1:C:84:ASP:HB3	1:C:87:VAL:H	1.77	0.48
1:D:46:ARG:HD3	1:D:242:ASP:OD1	2.14	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:85:THR:HB	1:C:263:ARG:HG3	1.96	0.47
1:C:92:ASP:HB3	1:C:161:ARG:HH12	1.78	0.47
1:A:46:ARG:NH2	1:A:239[B]:GLN:HG3	2.29	0.47
1:C:85:THR:CB	1:C:263:ARG:HA	2.45	0.47
1:C:162:PHE:CE1	1:C:172:TRP:HA	2.50	0.47
1:C:176:ARG:NH2	1:C:176:ARG:HB3	2.30	0.47
1:A:42[A]:VAL:HG22	1:A:246:TRP:HB2	1.96	0.47
1:C:66:HIS:HD2	3:C:536:HOH:O	1.98	0.47
1:B:160:GLU:HG3	3:B:536:HOH:O	2.15	0.46
1:B:182:LEU:HD12	1:B:182:LEU:C	2.36	0.46
1:B:68:HIS:CE1	1:B:70:ILE:HG22	2.51	0.46
1:B:153[A]:PHE:CZ	1:B:176:ARG:HA	2.50	0.46
3:B:400:HOH:O	1:D:214:LEU:HD11	2.15	0.46
1:C:89:ASP:OD2	1:C:91:ARG:HG3	2.16	0.45
1:D:68:HIS:HA	1:D:69:PRO:HD3	1.81	0.45
1:B:167:GLN:O	1:B:171:ARG:HG3	2.16	0.45
1:C:50[B]:ILE:HG13	1:C:54:GLN:HB2	1.98	0.45
1:D:50:ILE:O	1:D:115:LYS:NZ	2.49	0.45
1:C:90:VAL:O	1:C:90:VAL:HG23	2.17	0.45
1:D:108:LEU:C	1:D:108:LEU:HD23	2.36	0.45
1:A:3:LEU:N	3:A:671:HOH:O	2.50	0.45
1:B:165:THR:HB	1:B:166:PRO:CD	2.47	0.45
1:D:257:ASP:OD2	1:D:260:ARG:NH1	2.50	0.45
1:B:21:ILE:HG12	1:B:48:GLN:HG2	1.98	0.44
1:A:157:PHE:HA	1:A:161:ARG:HD3	2.00	0.44
1:A:229:ARG:NH1	3:A:472:HOH:O	2.49	0.44
1:B:93:ASN:HB3	1:B:255:ALA:HB3	2.00	0.44
1:C:30:ARG:HD3	1:C:30:ARG:C	2.38	0.43
1:A:68:HIS:CE1	1:A:71:TYR:CD2	3.06	0.43
1:C:164:THR:CG2	1:C:165:THR:N	2.82	0.43
1:B:42:VAL:HG11	1:B:234:ILE:HD11	2.01	0.42
1:D:89:ASP:OD2	1:D:91:ARG:HB2	2.18	0.42
1:B:10:PRO:HG3	1:C:120[A]:TYR:CE1	2.54	0.42
1:C:120[A]:TYR:C	1:C:120[A]:TYR:CD1	2.94	0.42
1:C:165:THR:CB	1:C:166:PRO:CD	2.97	0.42
1:C:176:ARG:HB3	1:C:176:ARG:CZ	2.49	0.42
1:B:50:ILE:O	1:B:115:LYS:NZ	2.53	0.41
1:A:260:ARG:HG2	3:A:381:HOH:O	2.20	0.41
1:A:169:LEU:O	1:A:173:GLU:HG3	2.21	0.41
1:C:176:ARG:HD2	3:C:609:HOH:O	2.21	0.41
1:D:46:ARG:NH1	1:D:239[A]:GLN:HG3	2.36	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:240:GLU:O	1:D:241:ASN:HB2	2.21	0.41
1:A:108:LEU:C	1:A:108:LEU:HD23	2.41	0.40
1:B:160:GLU:CG	3:B:536:HOH:O	2.69	0.40
1:D:160:GLU:CG	3:D:352:HOH:O	2.69	0.40
1:C:101:THR:O	1:C:184:HIS:HE1	2.05	0.40
1:B:153[B]:PHE:CZ	1:B:175:THR:HG21	2.57	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	283/277 (102%)	277 (98%)	6 (2%)	0	100	100
1	B	282/277 (102%)	273 (97%)	9 (3%)	0	100	100
1	C	280/277 (101%)	273 (98%)	7 (2%)	0	100	100
1	D	279/277 (101%)	273 (98%)	6 (2%)	0	100	100
All	All	1124/1108 (101%)	1096 (98%)	28 (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	243/235 (103%)	241 (99%)	2 (1%)	81	76
1	B	241/235 (103%)	239 (99%)	2 (1%)	81	76
1	C	239/235 (102%)	235 (98%)	4 (2%)	60	47
1	D	239/235 (102%)	237 (99%)	2 (1%)	81	76
All	All	962/940 (102%)	952 (99%)	10 (1%)	76	69

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	9[A]	SER
1	A	9[B]	SER
1	B	9	SER
1	B	30	ARG
1	C	165	THR
1	C	167	GLN
1	C	168	ASP
1	C	268	ARG
1	D	9	SER
1	D	88	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	253	HIS
1	B	151	HIS
1	B	201	ASN
1	C	184	HIS
1	C	201	ASN
1	C	209	ASN
1	D	201	ASN

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)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

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 (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	275/277 (99%)	-0.33	2 (0%) 87 88	11, 18, 45, 72	0
1	B	276/277 (99%)	0.20	32 (11%) 4 4	13, 26, 77, 132	0
1	C	276/277 (99%)	0.10	26 (9%) 8 8	13, 20, 88, 207	0
1	D	276/277 (99%)	0.17	27 (9%) 7 7	14, 24, 66, 160	0
All	All	1103/1108 (99%)	0.03	87 (7%) 12 12	11, 22, 70, 207	0

All (87) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	70	ILE	9.6
1	C	87	VAL	8.0
1	C	85	THR	7.7
1	D	71	TYR	6.9
1	B	164	THR	6.4
1	C	88	THR	6.0
1	B	71	TYR	5.9
1	D	69	PRO	5.9
1	C	170	ALA	5.8
1	C	86	ALA	5.7
1	B	70	ILE	5.6
1	C	91	ARG	5.5
1	B	162	PHE	5.3
1	B	169	LEU	5.2
1	C	165	THR	5.2
1	D	75	PRO	5.1
1	C	93	ASN	5.1
1	B	165	THR	4.9
1	B	87	VAL	4.8
1	C	166	PRO	4.7
1	C	164	THR	4.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	162	PHE	4.3
1	C	168	ASP	4.2
1	C	169	LEU	4.1
1	B	90	VAL	4.1
1	B	74	VAL	4.0
1	B	163	GLY	3.9
1	C	90	VAL	3.8
1	C	94	ALA	3.7
1	D	67	ILE	3.7
1	D	162	PHE	3.6
1	B	166	PRO	3.6
1	D	76	ASP	3.5
1	C	71	TYR	3.5
1	D	91	ARG	3.5
1	A	161	ARG	3.4
1	A	162	PHE	3.4
1	D	87	VAL	3.4
1	D	2	SER	3.3
1	D	72	PRO	3.3
1	D	246	TRP	3.2
1	B	160	GLU	3.2
1	B	159	LEU	3.2
1	C	172	TRP	3.2
1	B	76	ASP	3.1
1	D	126	TRP	3.1
1	B	89	ASP	3.1
1	D	68	HIS	3.1
1	B	69	PRO	3.0
1	B	72	PRO	3.0
1	D	74	VAL	3.0
1	D	90	VAL	3.0
1	B	75	PRO	2.8
1	B	167	GLN	2.7
1	C	69	PRO	2.7
1	D	261	PRO	2.7
1	C	70	ILE	2.7
1	D	112	LEU	2.7
1	C	161	ARG	2.7
1	B	2	SER	2.6
1	D	73	ASN	2.6
1	C	159	LEU	2.6
1	D	3	LEU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	173	GLU	2.5
1	D	92	ASP	2.4
1	D	165	THR	2.4
1	D	111	VAL	2.4
1	B	68	HIS	2.4
1	B	73	ASN	2.4
1	C	163	GLY	2.4
1	B	88	THR	2.3
1	B	172	TRP	2.3
1	B	261	PRO	2.3
1	D	177	ARG	2.3
1	C	74	VAL	2.3
1	D	89	ASP	2.3
1	B	171	ARG	2.2
1	C	120[A]	TYR	2.2
1	B	157	PHE	2.2
1	B	91	ARG	2.2
1	B	86	ALA	2.1
1	C	167	GLN	2.1
1	B	92	ASP	2.1
1	D	244	ALA	2.1
1	B	176	ARG	2.0
1	D	260	ARG	2.0
1	C	89	ASP	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	CL	A	279	1/1	0.99	0.09	24,24,24,24	0
2	CL	A	278	1/1	1.00	0.10	21,21,21,21	0

6.5 Other polymers (i)

There are no such residues in this entry.