wwPDB X-ray Structure Validation Summary Report (i)

Feb 10, 2024 – 12:00 PM EST

PDB ID	:	2IFC
Title	:	The Structure of the Binary Complex of Oxalateacetate with Citrate Synthase from the Thermophilic Archaeon Thermolasma acidophilum
Authors	:	Lehmann, C.
Deposited on	:	2006-09-20
Resolution	:	1.70 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 (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 (i)) were used in the production of this report:

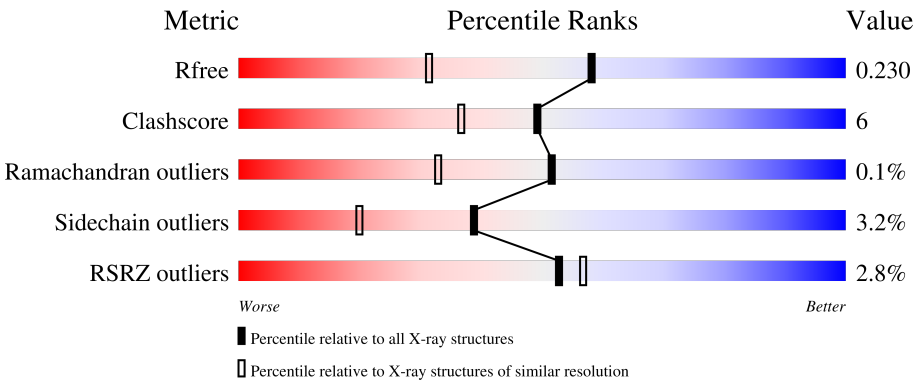
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	385	.% 87% 11% •
1	B	385	3% 87% 9% ••
1	C	385	.% 88% 10% ••
1	D	385	6% 86% 10% ••

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 14321 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.

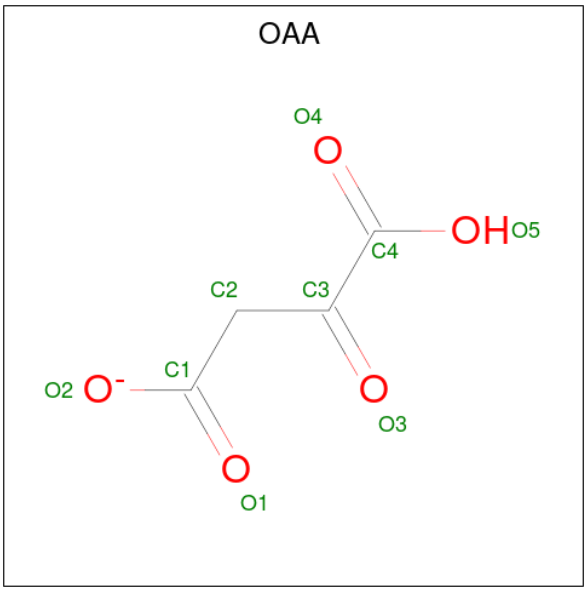
- Molecule 1 is a protein called Citrate Synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	380	3019	1932	511	563	13	0	3	0
1	B	380	3012	1926	510	564	12	0	2	0
1	C	382	3057	1950	522	573	12	0	6	0
1	D	376	2969	1901	501	555	12	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P21553
B	0	MET	-	initiating methionine	UNP P21553
C	0	MET	-	initiating methionine	UNP P21553
D	0	MET	-	initiating methionine	UNP P21553

- Molecule 2 is OXALOACETATE ION (three-letter code: OAA) (formula: $C_4H_3O_5$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	9	4	5	0	0
2	B	1	9	4	5	0	0
2	C	1	9	4	5	0	0
2	D	1	9	4	5	0	0

- Molecule 3 is water.

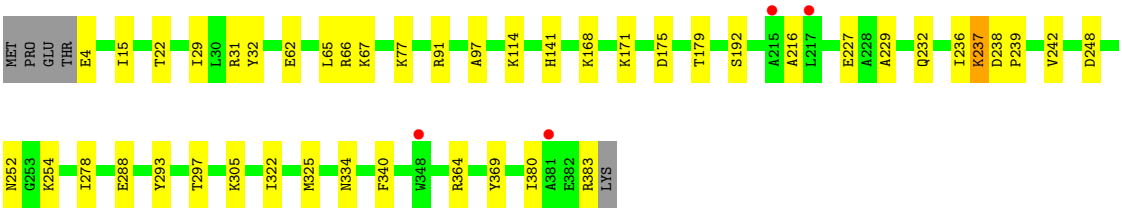
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	670	670	670	0	0
3	B	472	472	472	0	0
3	C	628	628	628	0	0
3	D	458	458	458	0	0

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.

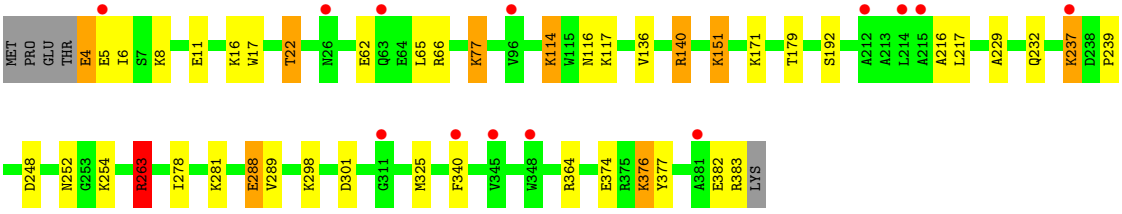
- Molecule 1: Citrate Synthase

Chain A:



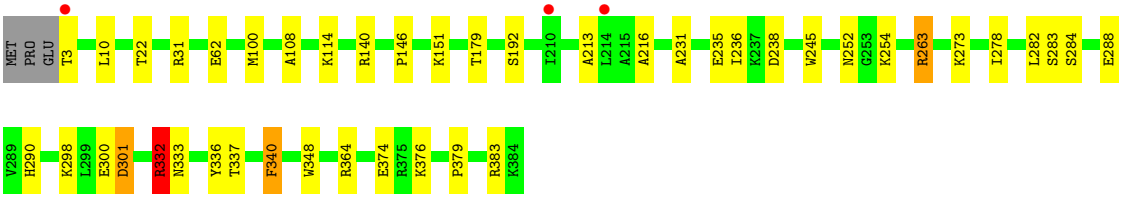
- Molecule 1: Citrate Synthase

Chain B:



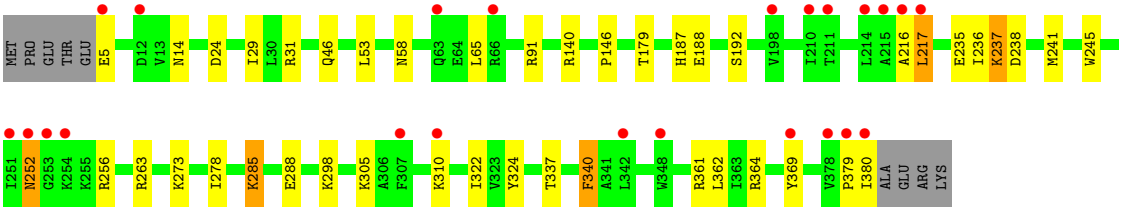
- Molecule 1: Citrate Synthase

Chain C:



- Molecule 1: Citrate Synthase

Chain D:


MET PRO GLU THR GLU E5 D12 V13 N14 D24 I29 L30 R31 Q46 L53 N58 Q63 E64 L65 R66 R91 R140 P146 T179 H187 E188 S192 V198 I210 T211 L214 A215 A216 L217 E235 I236 K237 D238 M241 W245
I251 N252 G253 K254 K255 R256 R263 K273 I278 K285 E288 K298 K305 A306 F307 K310 I322 V323 Y324 T337 F340 A341 L342 W348 R361 L362 I363 R364 Y369 V378 P379 I380 ALA GLU ARG LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	78.53Å 97.45Å 106.94Å 90.00° 93.10° 90.00°	Depositor
Resolution (Å)	41.38 – 1.70 41.38 – 1.70	Depositor EDS
% Data completeness (in resolution range)	96.8 (41.38-1.70) 96.8 (41.38-1.70)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.13 (at 1.70Å)	Xtriage
Refinement program	REFMAC, CNS	Depositor
R, R_{free}	0.181 , 0.234 0.177 , 0.230	Depositor DCC
R_{free} test set	8560 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	23.4	Xtriage
Anisotropy	0.109	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 55.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	14321	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.89% 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: OAA

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	1/3083 (0.0%)	0.90	3/4165 (0.1%)
1	B	0.76	0/3076	0.76	6/4156 (0.1%)
1	C	1.09	6/3133 (0.2%)	0.94	6/4232 (0.1%)
1	D	0.85	1/3033 (0.0%)	0.75	1/4100 (0.0%)
All	All	0.95	8/12325 (0.1%)	0.84	16/16653 (0.1%)

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	C	0	2
1	D	0	1
All	All	0	3

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	332[A]	ARG	C-N	-7.04	1.17	1.34
1	C	332[B]	ARG	C-N	-7.04	1.17	1.34
1	C	301[A]	ASP	C-N	-6.89	1.18	1.34
1	C	301[B]	ASP	C-N	-6.89	1.18	1.34
1	C	336	TYR	CD2-CE2	5.64	1.47	1.39

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	383	ARG	NE-CZ-NH1	-10.49	115.06	120.30
1	C	332[A]	ARG	NE-CZ-NH1	9.23	124.91	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	332[B]	ARG	NE-CZ-NH1	9.23	124.91	120.30
1	A	383	ARG	NE-CZ-NH2	8.12	124.36	120.30
1	B	263	ARG	NE-CZ-NH2	-7.87	116.37	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	332[A]	ARG	Sidechain
1	C	332[B]	ARG	Sidechain
1	D	379	PRO	Peptide

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	3019	0	3026	37	0
1	B	3012	0	3013	41	0
1	C	3057	0	3047	41	1
1	D	2969	0	2971	37	0
2	A	9	0	2	0	0
2	B	9	0	2	0	0
2	C	9	0	2	0	0
2	D	9	0	2	0	0
3	A	670	0	0	23	0
3	B	472	0	0	13	1
3	C	628	0	0	13	0
3	D	458	0	0	13	0
All	All	14321	0	12065	142	1

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 6.

The worst 5 of 142 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:4:GLU:HG3	1:B:5:GLU:N	1.76	0.99
1:A:237[B]:LYS:HA	1:A:237[B]:LYS:HE3	1.45	0.94
1:A:288:GLU:HG3	3:A:822:HOH:O	1.69	0.91
1:B:252:ASN:OD1	3:B:855:HOH:O	1.90	0.89
1:B:151:LYS:HG2	3:B:821:HOH:O	1.74	0.87

All (1) 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)
1:C:332[B]:ARG:NH2	3:B:706:HOH:O[2_553]	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	381/385 (99%)	368 (97%)	13 (3%)	0	100	100
1	B	380/385 (99%)	372 (98%)	8 (2%)	0	100	100
1	C	386/385 (100%)	376 (97%)	10 (3%)	0	100	100
1	D	375/385 (97%)	364 (97%)	9 (2%)	2 (0%)	29	13
All	All	1522/1540 (99%)	1480 (97%)	40 (3%)	2 (0%)	51	33

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	252	ASN
1	D	188	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	313/315 (99%)	305 (97%)	8 (3%)	46	28
1	B	312/315 (99%)	295 (95%)	17 (5%)	22	7
1	C	318/315 (101%)	311 (98%)	7 (2%)	52	34
1	D	308/315 (98%)	297 (96%)	11 (4%)	35	16
All	All	1251/1260 (99%)	1208 (97%)	43 (3%)	39	18

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	282	LEU
1	D	65	LEU
1	C	284	SER
1	D	5	GLU
1	D	217	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	58	ASN
1	D	247	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)

4 ligands are 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	OAA	B	385	-	8,8,8	2.50	4 (50%)	9,10,10	1.65	2 (22%)
2	OAA	C	385	-	8,8,8	2.51	3 (37%)	9,10,10	1.89	2 (22%)
2	OAA	D	385	-	8,8,8	2.74	4 (50%)	9,10,10	2.48	3 (33%)
2	OAA	A	385	-	8,8,8	3.00	3 (37%)	9,10,10	3.94	4 (44%)

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	OAA	B	385	-	-	0/8/8/8	-
2	OAA	C	385	-	-	0/8/8/8	-
2	OAA	D	385	-	-	0/8/8/8	-
2	OAA	A	385	-	-	0/8/8/8	-

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	385	OAA	O3-C3	6.76	1.36	1.23
2	D	385	OAA	O3-C3	4.97	1.33	1.23
2	C	385	OAA	O3-C3	4.96	1.33	1.23
2	B	385	OAA	O3-C3	4.53	1.32	1.23
2	D	385	OAA	O4-C4	3.91	1.33	1.22

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	385	OAA	O4-C4-C3	-7.59	111.58	121.72
2	A	385	OAA	O3-C3-C2	7.45	131.36	120.58
2	D	385	OAA	O3-C3-C2	5.92	129.15	120.58
2	C	385	OAA	O3-C3-C2	4.63	127.28	120.58
2	A	385	OAA	C2-C3-C4	-3.64	111.71	117.85

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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	C	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	301:ASP	C	302:PHE	N	1.18
1	C	332:ARG	C	333:ASN	N	1.17

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	380/385 (98%)	-0.05	4 (1%)	80	83	11, 18, 31, 44	0
1	B	380/385 (98%)	0.18	13 (3%)	45	50	18, 28, 39, 47	0
1	C	382/385 (99%)	-0.10	3 (0%)	86	88	11, 18, 30, 45	0
1	D	376/385 (97%)	0.29	23 (6%)	21	23	17, 25, 40, 53	0
All	All	1518/1540 (98%)	0.08	43 (2%)	53	57	11, 23, 37, 53	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	253	GLY	5.3
1	D	378	VAL	3.9
1	C	3	THR	3.5
1	D	252	ASN	3.4
1	D	214	LEU	3.1

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	OAA	B	385	9/9	0.95	0.07	25,27,31,32	0
2	OAA	D	385	9/9	0.95	0.14	20,25,28,30	0
2	OAA	C	385	9/9	0.96	0.09	18,19,21,23	0
2	OAA	A	385	9/9	0.96	0.10	14,16,20,21	0

6.5 Other polymers (i)

There are no such residues in this entry.