Full wwPDB X-ray Structure Validation Report (i)

Aug 20, 2020 – 07:32 PM BST

PDB ID	:	4CON
Title	:	Crystal structure of the anaerobic ribonucleotide reductase from Thermotoga maritima with citrate in the active site
Authors	:	Aurelius, O.; Johansson, R.; Bagenholm, V.; Beck, T.; Balhuizen, A.; Lundin, D.; Sjoberg, B.M.; Mulliez, E.; Logan, D.T.
Deposited on	:	2014-01-29
Resolution	:	2.12 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 following versions of software and data (see references (1)) were used in the production of this report:

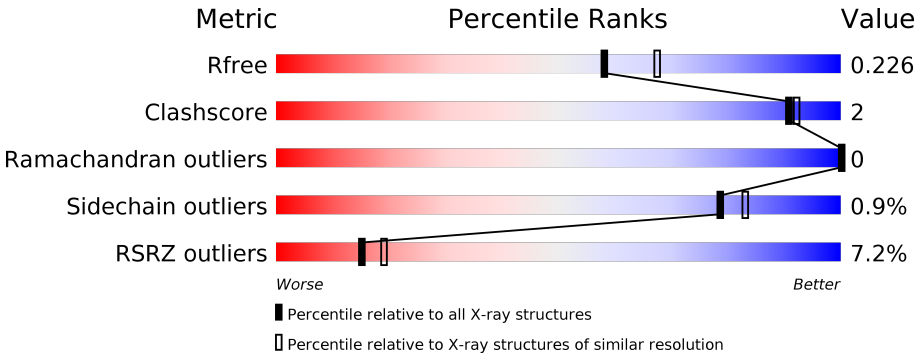
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.12 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 <=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	651	7% 79% . 17%
1	B	651	5% 81% . 15%

2 Entry composition

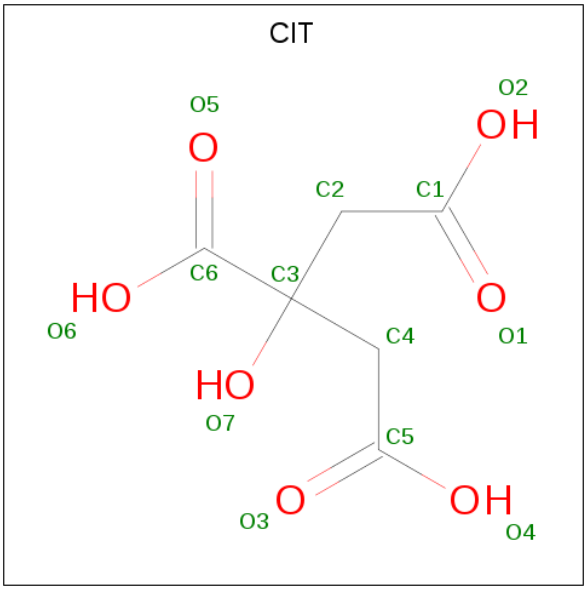
There are 3 unique types of molecules in this entry. The entry contains 9230 atoms, of which 10 are hydrogens and 0 are deuteriums.

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 ANAEROBIC RIBONUCLEOSIDE-TRIPHOSPHATE REDUCTASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	543	Total	C	N	O	S	0	1	0
			4467	2894	730	822	21			
1	B	553	Total	C	N	O	S	0	1	0
			4539	2935	742	840	22			

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			18	6	5	7		
2	B	1	Total	C	H	O	0	0
			18	6	5	7		

- Molecule 3 is water.

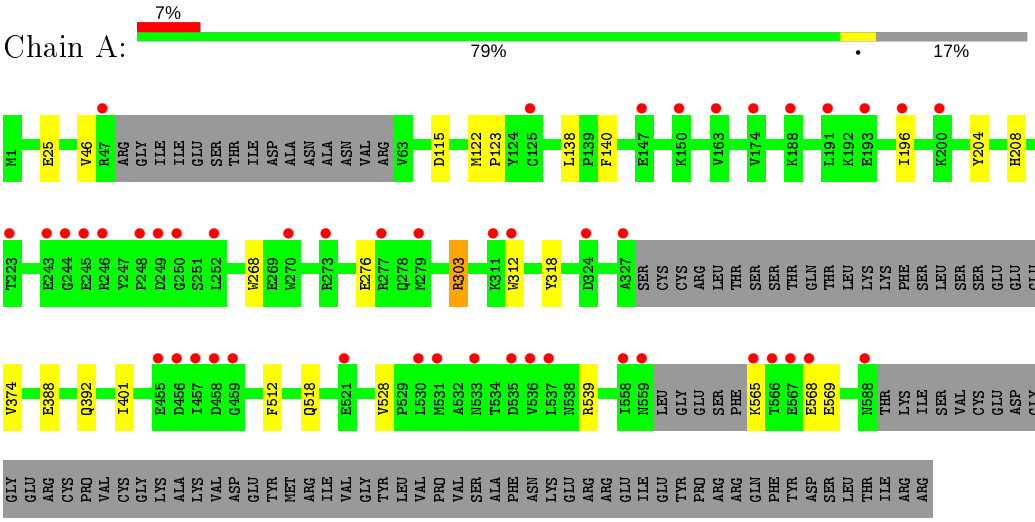
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	88	Total O 88 88	0	0
3	B	100	Total O 100 100	0	0

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.

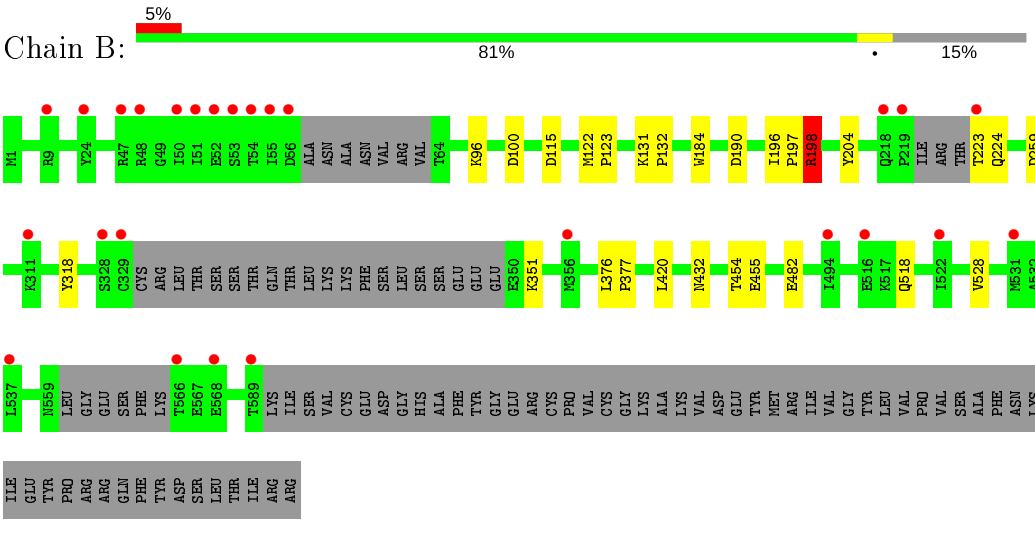
• Molecule 1: ANAEROBIC RIBONUCLEOSIDE-TRIPHOSPHATE REDUCTASE

Chain A:



• Molecule 1: ANAEROBIC RIBONUCLEOSIDE-TRIPHOSPHATE REDUCTASE

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	78.55Å 97.74Å 86.58Å 90.00° 112.13° 90.00°	Depositor
Resolution (Å)	41.73 – 2.12 42.54 – 2.12	Depositor EDS
% Data completeness (in resolution range)	99.5 (41.73-2.12) 99.5 (42.54-2.12)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.80 (at 2.12Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.182 , 0.218 0.196 , 0.226	Depositor DCC
R_{free} test set	1713 reflections (2.51%)	wwPDB-VP
Wilson B-factor (Å^2)	36.5	Xtriage
Anisotropy	0.836	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 50.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	9230	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.16% 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

Bond lengths and bond angles in the following residue types are not validated in this section: CIT

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	0.44	0/4577	0.54	0/6175
1	B	0.43	0/4649	0.55	1/6273 (0.0%)
All	All	0.44	0/9226	0.55	1/12448 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	198	ARG	NE-CZ-NH1	5.34	122.97	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	4467	0	4402	15	0
1	B	4539	0	4468	16	0
2	A	13	5	5	0	0
2	B	13	5	5	0	0
3	A	88	0	0	2	0
3	B	100	0	0	0	0
All	All	9220	10	8880	30	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.

All (30) 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:388:GLU:OE1	3:A:2062:HOH:O	2.12	0.66
1:A:196:ILE:HG21	1:A:204:TYR:CD2	2.42	0.54
1:B:190:ASP:HB3	1:B:196:ILE:CD1	2.38	0.54
1:A:303:ARG:NH2	1:A:568:GLU:OE1	2.41	0.53
1:B:96:LYS:NZ	1:B:100:ASP:OD2	2.42	0.52
1:A:374:VAL:HG22	1:A:401:ILE:CD1	2.41	0.51
1:B:196:ILE:HG12	1:B:204:TYR:CD2	2.46	0.51
1:A:565:LYS:N	1:A:569:GLU:OE1	2.46	0.48
1:A:122:MET:HB2	1:A:123:PRO:HD2	1.94	0.48
1:A:528:VAL:CG1	1:A:539:ARG:HG3	2.44	0.48
1:B:198:ARG:CG	1:B:198:ARG:HH11	2.28	0.46
1:B:454:THR:OG1	1:B:455:GLU:N	2.50	0.44
1:A:392:GLN:HG3	3:A:2063:HOH:O	2.17	0.44
1:B:131:LYS:HB3	1:B:132:PRO:HD3	2.00	0.44
1:B:190:ASP:HB3	1:B:196:ILE:HD12	1.99	0.44
1:A:46:VAL:HG23	1:B:420:LEU:HD11	2.00	0.43
1:B:351:LYS:NZ	1:B:482:GLU:OE1	2.37	0.43
1:A:276:GLU:HG3	1:A:312:TRP:HZ2	1.83	0.43
1:B:122:MET:HB2	1:B:123:PRO:HD2	1.99	0.43
1:B:196:ILE:HG22	1:B:197:PRO:O	2.19	0.43
1:A:25:GLU:HG3	1:A:512:PHE:CZ	2.54	0.42
1:B:528:VAL:HG11	1:B:534:THR:HG21	2.00	0.42
1:A:374:VAL:HG22	1:A:401:ILE:HD12	2.02	0.41
1:A:138:LEU:HD22	1:A:140:PHE:CZ	2.55	0.41
1:B:184:TRP:CE2	1:B:259:ASP:HB3	2.56	0.41
1:B:432:ASN:OD1	1:B:432:ASN:N	2.54	0.41
1:A:204:TYR:CZ	1:A:208:HIS:HE1	2.39	0.41
1:B:376:LEU:HB2	1:B:377:PRO:HD3	2.02	0.40
1:B:223:THR:HG23	1:B:224:GLN:N	2.36	0.40
1:A:25:GLU:HG3	1:A:512:PHE:CE1	2.56	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	534/651 (82%)	522 (98%)	12 (2%)	0	100	100
1	B	544/651 (84%)	533 (98%)	11 (2%)	0	100	100
All	All	1078/1302 (83%)	1055 (98%)	23 (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	481/577 (83%)	476 (99%)	5 (1%)	76	81
1	B	490/577 (85%)	486 (99%)	4 (1%)	81	86
All	All	971/1154 (84%)	962 (99%)	9 (1%)	78	83

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	115	ASP
1	A	268	TRP
1	A	303	ARG
1	A	318	TYR
1	A	518	GLN
1	B	115	ASP
1	B	198	ARG
1	B	318	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	518	GLN

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

Mol	Chain	Res	Type
1	A	158	HIS
1	B	158	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

2 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	CIT	B	1590	-	3,12,12	0.38	0	3,17,17	2.51	1 (33%)
2	CIT	A	1589	-	3,12,12	1.24	0	3,17,17	0.79	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
2	CIT	B	1590	-	-	3/6/16/16	-
2	CIT	A	1589	-	-	0/6/16/16	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1590	CIT	C4-C3-C2	3.89	119.72	109.33

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1590	CIT	C1-C2-C3-O7
2	B	1590	CIT	C1-C2-C3-C6
2	B	1590	CIT	O7-C3-C4-C5

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

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	543/651 (83%)	0.47	48 (8%)	10	12	34, 61, 107, 129	0
1	B	553/651 (84%)	0.33	31 (5%)	24	29	34, 55, 92, 127	0
All	All	1096/1302 (84%)	0.40	79 (7%)	15	19	34, 57, 102, 129	0

All (79) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	457	ILE	6.7
1	B	279	MET	5.8
1	B	50	ILE	5.3
1	B	55	ILE	5.2
1	A	248	PRO	4.8
1	B	328	SER	4.5
1	B	24	TYR	4.5
1	A	559	ASN	4.4
1	B	47	ARG	4.4
1	B	537	LEU	4.2
1	A	196	ILE	4.2
1	A	533	ASN	4.2
1	A	246	ARG	4.0
1	A	565	LYS	3.9
1	A	566	THR	3.9
1	A	327	ALA	3.8
1	A	312	TRP	3.7
1	B	53	SER	3.6
1	A	458	ASP	3.5
1	B	48	ARG	3.4
1	A	252	LEU	3.3
1	A	455	GLU	3.3
1	B	52	GLU	3.3
1	A	223	THR	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	531	MET	3.2
1	A	459	GLY	3.2
1	A	243	GLU	3.1
1	A	531	MET	3.1
1	A	530	LEU	3.1
1	B	9	ARG	3.0
1	A	536	VAL	3.0
1	B	536	VAL	2.9
1	A	324	ASP	2.9
1	B	219	PRO	2.8
1	B	533	ASN	2.8
1	A	47	ARG	2.8
1	A	147	GLU	2.7
1	B	56	ASP	2.7
1	A	558	ILE	2.7
1	B	51	ILE	2.7
1	A	456	ASP	2.7
1	A	188	LYS	2.6
1	A	588	ASN	2.6
1	A	245	GLU	2.6
1	A	537	LEU	2.6
1	A	277	ARG	2.6
1	A	244	GLY	2.5
1	B	54	THR	2.5
1	B	223	THR	2.5
1	B	356[A]	MET	2.5
1	A	125	CYS	2.5
1	A	150	LYS	2.4
1	A	174	VAL	2.4
1	B	277	ARG	2.4
1	A	270	TRP	2.4
1	A	273	ARG	2.4
1	A	163	VAL	2.3
1	B	589	THR	2.3
1	A	200	LYS	2.3
1	A	567	GLU	2.3
1	A	193	GLU	2.3
1	A	521	GLU	2.3
1	A	311	LYS	2.2
1	A	356[A]	MET	2.2
1	B	218	GLN	2.1
1	B	568	GLU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	249	ASP	2.1
1	B	329	CYS	2.1
1	A	250	GLY	2.1
1	B	522	ILE	2.1
1	B	516	GLU	2.1
1	B	311	LYS	2.1
1	A	279	MET	2.0
1	A	535	ASP	2.0
1	A	191	LEU	2.0
1	B	566	THR	2.0
1	A	568	GLU	2.0
1	B	494	ILE	2.0
1	B	535	ASP	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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, 95^{th} 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	CIT	B	1590	13/13	0.88	0.15	62,70,90,90	0
2	CIT	A	1589	13/13	0.94	0.14	53,66,85,85	0

6.5 Other polymers

There are no such residues in this entry.