wwPDB X-ray Structure Validation Summary Report (i)

Aug 20, 2020 – 12:43 PM BST

PDB ID	:	4D4O
Title	:	Crystal Structure of the Kti11 Kti13 heterodimer Spacegroup P64
Authors	:	Glatt, S.; Mueller, C.W.
Deposited on	:	2014-10-30
Resolution	:	2.90 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 following versions of software and data (see references (1)) were used in the production of this report:

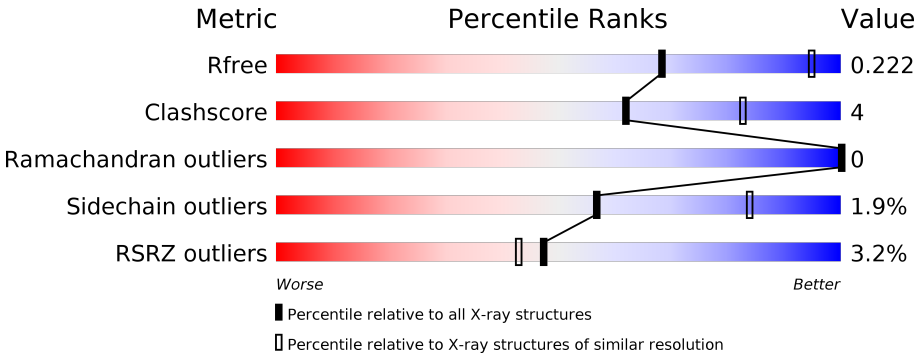
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	427	4% 83% 14%
1	B	427	2% 70% 7% 22%
1	C	427	16% 82%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6346 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 PROTEIN ATS1, DIPHTHAMIDE BIOSYNTHESIS PROTEIN 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	414	Total	C	N	O	S	0	0	0
			3185	1987	555	616	27			
1	B	331	Total	C	N	O	S	0	0	0
			2538	1582	464	474	18			
1	C	75	Total	C	N	O	S	0	0	0
			606	384	85	128	9			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP P31386
A	0	ALA	-	expression tag	UNP P31386
A	334	GLY	-	linker	UNP P31386
A	335	SER	-	linker	UNP P31386
A	336	GLY	-	linker	UNP P31386
A	337	SER	-	linker	UNP P31386
A	338	GLY	-	linker	UNP P31386
A	339	SER	-	linker	UNP P31386
A	340	GLY	-	linker	UNP P31386
A	341	SER	-	linker	UNP P31386
A	342	GLY	-	linker	UNP P31386
A	343	SER	-	linker	UNP P31386
B	-1	GLY	-	expression tag	UNP P31386
B	0	ALA	-	expression tag	UNP P31386
B	334	GLY	-	linker	UNP P31386
B	335	SER	-	linker	UNP P31386
B	336	GLY	-	linker	UNP P31386
B	337	SER	-	linker	UNP P31386
B	338	GLY	-	linker	UNP P31386
B	339	SER	-	linker	UNP P31386
B	340	GLY	-	linker	UNP P31386
B	341	SER	-	linker	UNP P31386

Continued on next page...

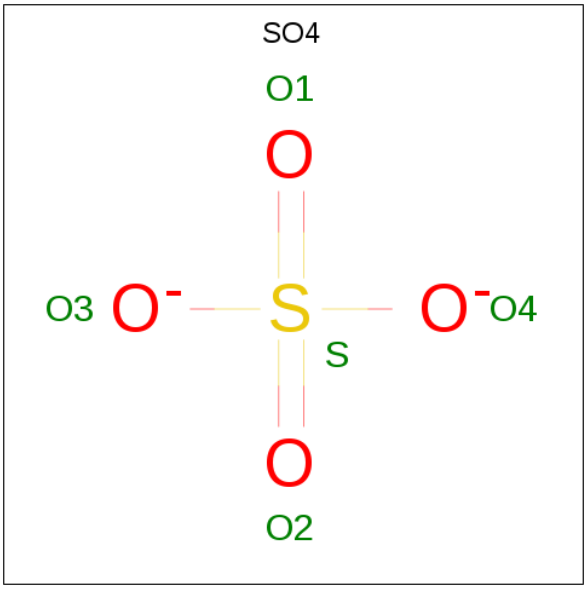
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	342	GLY	-	linker	UNP P31386
B	343	SER	-	linker	UNP P31386
C	-1	GLY	-	expression tag	UNP P31386
C	0	ALA	-	expression tag	UNP P31386
C	334	GLY	-	linker	UNP P31386
C	335	SER	-	linker	UNP P31386
C	336	GLY	-	linker	UNP P31386
C	337	SER	-	linker	UNP P31386
C	338	GLY	-	linker	UNP P31386
C	339	SER	-	linker	UNP P31386
C	340	GLY	-	linker	UNP P31386
C	341	SER	-	linker	UNP P31386
C	342	GLY	-	linker	UNP P31386
C	343	SER	-	linker	UNP P31386

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

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

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



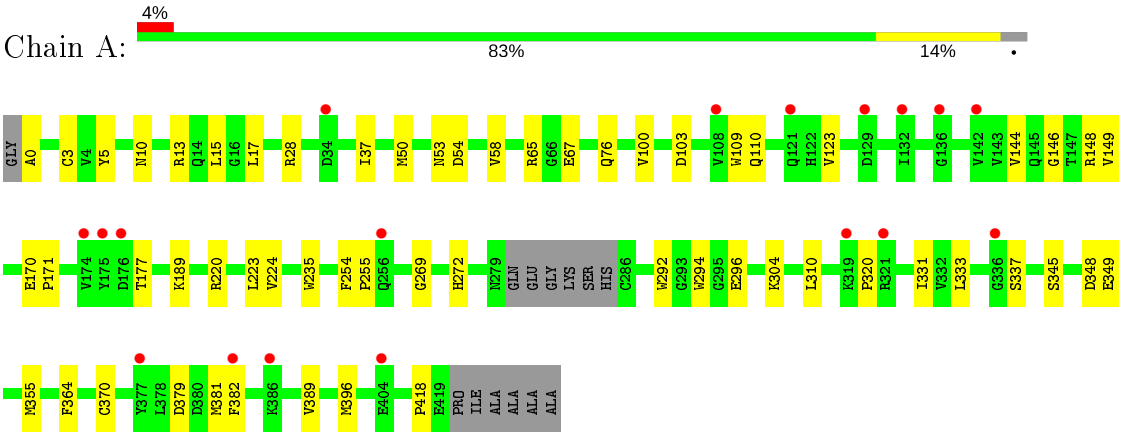
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0

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.

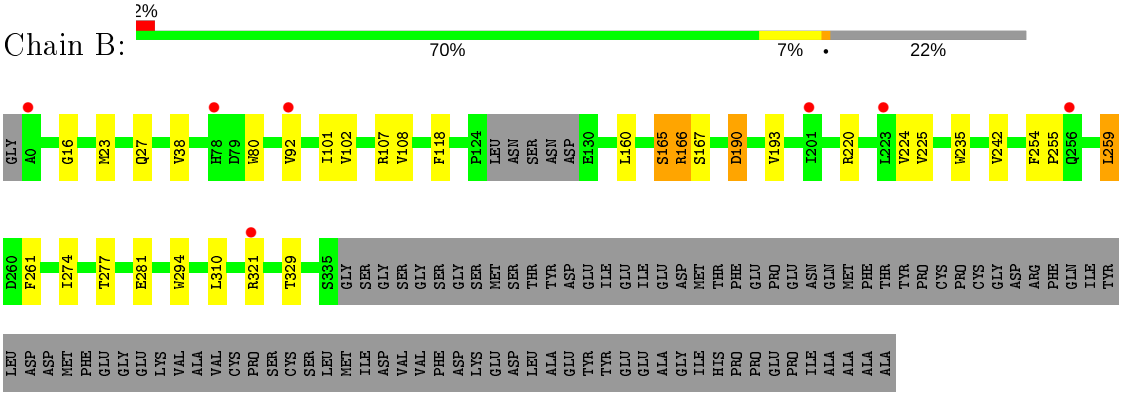
• Molecule 1: PROTEIN ATS1, DIPHTHAMIDE BIOSYNTHESIS PROTEIN 3

Chain A:



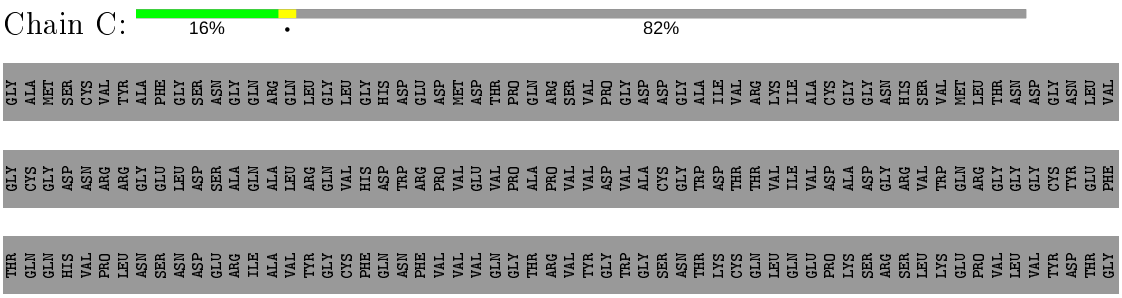
• Molecule 1: PROTEIN ATS1, DIPHTHAMIDE BIOSYNTHESIS PROTEIN 3

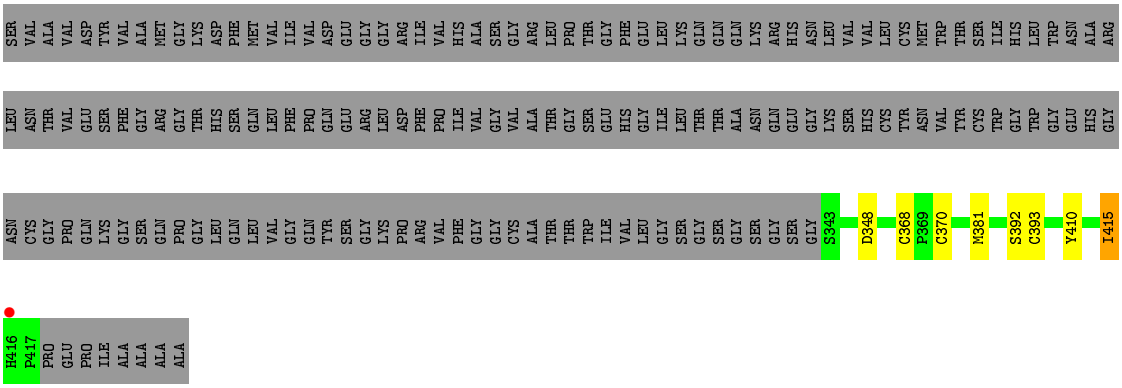
Chain B:



• Molecule 1: PROTEIN ATS1, DIPHTHAMIDE BIOSYNTHESIS PROTEIN 3

Chain C:




SER VAL ALA VAL ASP TYR VAL ALA MET GLY LYS ASP PHE MET VAL ILE VAL ASP GLU GLY GLY ARG ILE VAL HIS ALA SER GLY ARG LEU PRO THR GLY PHE GLU LEU LYS GLN GLN GLN LYS ARG HIS ASN LEU VAL VAL LEU CYS MET TRP THR SER ILE HIS LEU TRP ASN ALA ARG
LEU ASN THR VAL GLU SER PHE GLY ARG GLY THR HIS SER GLN LEU PHE PRO GLN GLU ARG LEU ASP PHE PRO ILE VAL GLY VAL ALA THR GLY SER GLU HIS GLY ILE LEU LEU THR THR ALA ASN GLN GLU GLY LYS SER HIS CYS TYR ASN VAL TYR CYS TRP GLY TRP GLY GLU HIS GLY
ASN CYS GLY PRO GLN LYS GLY SER GLN PRO GLY LEU GLN LEU VAL GLY GLN TYR SER GLY LYS PRO ARG VAL PHE GLY GLY CYS ALA THR THR TRP ILE VAL LEU GLY SER GLY SER GLY SER GLY SER GLY S343 D348 C368 P369 C370 M381 S392 C393 Y410 I415
H416 P417 PRO GLU PRO ILE ALA ALA ALA ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, α, β, γ	151.27Å 151.27Å 107.14Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	75.64 – 2.90 82.94 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.4 (75.64-2.90) 99.4 (82.94-2.90)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.53 (at 2.91Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.198 , 0.219 0.199 , 0.222	Depositor DCC
R_{free} test set	1546 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	83.8	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 16.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.034 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6346	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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: FE, SO4

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	0.22	0/3261	0.39	0/4426
1	B	0.22	0/2598	0.39	0/3525
1	C	0.42	0/621	0.46	0/841
All	All	0.25	0/6480	0.40	0/8792

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

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	3185	0	3009	32	0
1	B	2538	0	2438	19	0
1	C	606	0	541	6	0
2	A	1	0	0	0	0
2	C	1	0	0	1	0
3	A	5	0	0	0	0
3	B	10	0	0	2	0
All	All	6346	0	5988	53	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 4.

The worst 5 of 53 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:107:ARG:NH1	3:B:1336:SO4:O3	2.12	0.82
1:B:38:VAL:O	1:B:321:ARG:NH2	2.14	0.78
1:A:224:VAL:HB	1:A:235:TRP:HB3	1.67	0.75
1:B:224:VAL:HB	1:B:235:TRP:HB3	1.73	0.71
1:A:294:TRP:HH2	1:C:370:CYS:HB3	1.60	0.66

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	410/427 (96%)	383 (93%)	27 (7%)	0	100	100
1	B	327/427 (77%)	312 (95%)	15 (5%)	0	100	100
1	C	73/427 (17%)	70 (96%)	3 (4%)	0	100	100
All	All	810/1281 (63%)	765 (94%)	45 (6%)	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	344/351 (98%)	340 (99%)	4 (1%)	71	91

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	270/351 (77%)	263 (97%)	7 (3%)	46 77
1	C	69/351 (20%)	67 (97%)	2 (3%)	42 76
All	All	683/1053 (65%)	670 (98%)	13 (2%)	57 84

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	160	LEU
1	B	165	SER
1	B	281	GLU
1	B	118	PHE
1	B	259	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 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	$\#\|Z\| > 2$
3	SO4	B	1337	-	4,4,4	0.27	0	6,6,6	0.17	0
3	SO4	B	1336	-	4,4,4	0.34	0	6,6,6	0.28	0
3	SO4	A	1420	-	4,4,4	0.52	0	6,6,6	0.46	0

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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1337	SO4	1	0
3	B	1336	SO4	1	0

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	414/427 (96%)	0.52	18 (4%) 35 31	20, 48, 90, 174	0
1	B	331/427 (77%)	0.51	7 (2%) 63 61	20, 43, 86, 102	0
1	C	75/427 (17%)	0.41	1 (1%) 77 77	24, 46, 81, 89	0
All	All	820/1281 (64%)	0.51	26 (3%) 47 43	20, 46, 87, 174	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	176	ASP	4.0
1	A	382	PHE	3.8
1	A	34	ASP	3.7
1	B	201	ILE	3.2
1	A	108	VAL	3.1

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
3	SO4	B	1337	5/5	0.89	0.17	82,82,82,82	0
3	SO4	B	1336	5/5	0.89	0.15	114,114,114,114	0
3	SO4	A	1420	5/5	0.93	0.10	107,107,107,107	0
2	FE	C	501	1/1	0.98	0.17	47,47,47,47	0
2	FE	A	501	1/1	0.99	0.18	45,45,45,45	0

6.5 Other polymers (i)

There are no such residues in this entry.