



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 9, 2023 – 04:34 PM EDT

PDB ID : 4IVQ
Title : Crystal structure of thymidine kinase from Herpes simplex virus type 1 in complex with IN43/5
Authors : Pernot, L.; Novakovic, I.; Westermaier, Y.; Perozzo, R.; Raic-Malic, S.; Scapozza, L.
Deposited on : 2013-01-23
Resolution : 1.90 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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)
Xtrriage (Phenix) : 1.13
EDS : 2.35.1
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) : 2.35.1

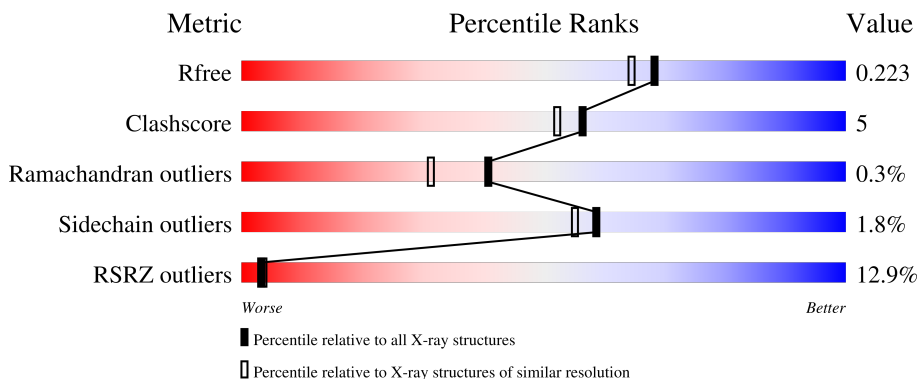
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 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 $\leq 5\%$. 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	331	 14% 80% 14% 5%
1	B	331	 11% 85% 8% 7%

2 Entry composition [i](#)

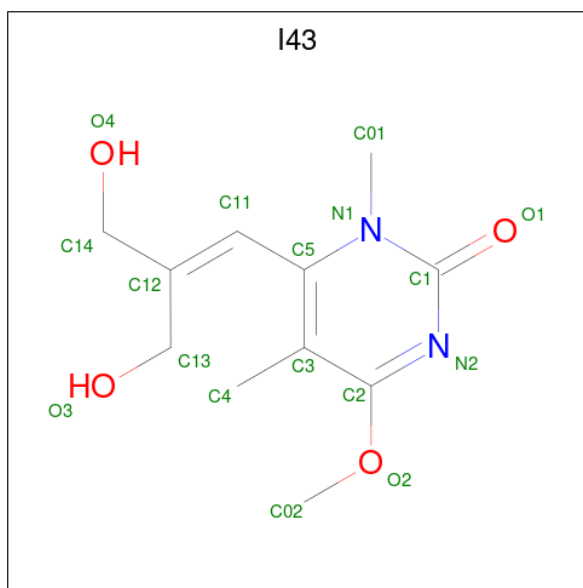
There are 4 unique types of molecules in this entry. The entry contains 4998 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Thymidine kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	313	Total 2376	C 1519	N 403	O 434	S 20	0	5	0
1	B	309	Total 2370	C 1519	N 409	O 423	S 19	0	8	0

- Molecule 2 is 6-[3-hydroxy-2-(hydroxymethyl)prop-1-en-1-yl]-4-methoxy-1,5-dimethylpyrimidin-2(1H)-one (three-letter code: I43) (formula: C₁₁H₁₆N₂O₄).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total 17	C 11	N 2	O 4	0	0
2	B	1	Total 17	C 11	N 2	O 4	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

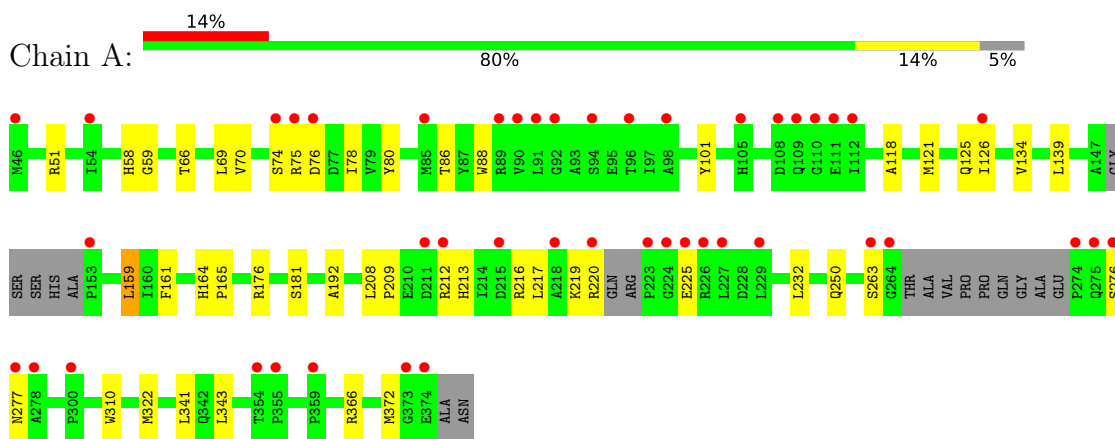
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	96	Total O 96 96	0	0
4	B	107	Total O 107 107	0	0

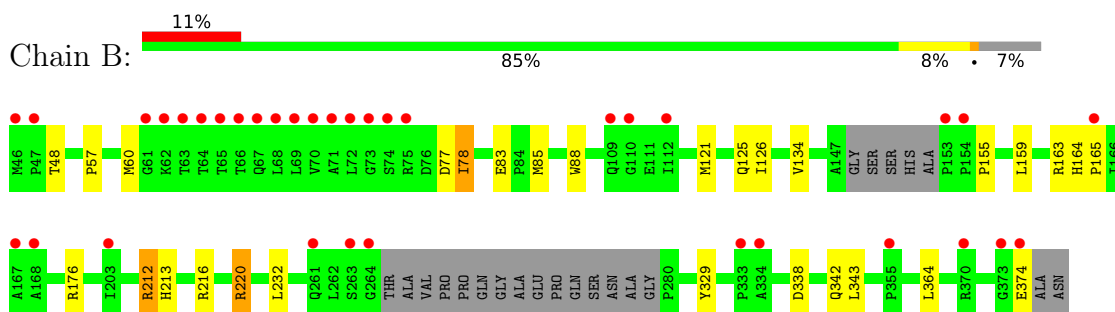
3 Residue-property plots [i](#)

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: Thymidine kinase



- Molecule 1: Thymidine kinase



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	113.44Å 117.07Å 107.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.16 – 1.90 27.16 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.5 (27.16-1.90) 98.5 (27.16-1.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.99 (at 1.91Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.192 , 0.222 0.192 , 0.223	Depositor DCC
R_{free} test set	2000 reflections (3.58%)	wwPDB-VP
Wilson B-factor (Å ²)	28.8	Xtrriage
Anisotropy	0.578	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 55.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.022 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	4998	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.09% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: I43, 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	0.35	0/2450	0.54	0/3345
1	B	0.36	0/2452	0.55	0/3347
All	All	0.36	0/4902	0.54	0/6692

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	2376	0	2370	31	0
1	B	2370	0	2385	20	0
2	A	17	0	16	5	0
2	B	17	0	16	5	0
3	A	10	0	0	0	0
3	B	5	0	0	1	0
4	A	96	0	0	2	0
4	B	107	0	0	1	0
All	All	4998	0	4787	48	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 (48) 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:176:ARG:HH21	2:B:401:I43:H4	1.39	0.88
1:A:176:ARG:HH21	2:A:501:I43:H4	1.40	0.86
1:A:76:ASP:HB3	1:A:366:ARG:HH22	1.62	0.65
1:A:58:HIS:HB3	2:A:501:I43:H13	1.79	0.64
1:B:163:ARG:HE	2:B:401:I43:H14	1.66	0.60
1:B:121:MET:O	1:B:125:GLN:HG2	2.04	0.58
1:B:212:ARG:NH2	4:B:584:HOH:O	2.35	0.58
1:A:220:ARG:NH2	4:A:622:HOH:O	2.32	0.56
1:A:51:ARG:HD2	1:A:139[A]:LEU:HD21	1.86	0.55
1:A:126:ILE:HG13	1:B:126[A]:ILE:HD11	1.90	0.52
1:A:176:ARG:NH2	2:A:501:I43:H4	2.16	0.52
1:B:88:TRP:CZ3	2:B:401:I43:H16	2.45	0.52
1:A:88:TRP:CZ3	2:A:501:I43:H15	2.45	0.52
1:A:74:SER:OG	1:A:75:ARG:N	2.43	0.51
1:A:250:GLN:NE2	1:A:322:MET:O	2.45	0.50
1:A:101:TYR:CZ	1:A:225:GLU:HG3	2.47	0.49
1:B:83:GLU:OE2	2:B:401:I43:H13	2.12	0.49
1:A:121:MET:HG3	1:A:181:SER:HB2	1.95	0.48
2:B:401:I43:H5	2:B:401:I43:H8	1.96	0.48
1:A:59:GLY:O	1:A:216:ARG:HD3	2.14	0.47
1:B:216:ARG:O	1:B:220:ARG:HG2	2.14	0.47
1:B:77:ASP:OD1	1:B:78:ILE:HG22	2.15	0.46
1:A:86:THR:HB	1:A:372:MET:HA	1.97	0.46
1:B:85:MET:HB3	1:B:374:GLU:HG3	1.99	0.45
1:A:66:THR:HG21	1:A:80:TYR:CE1	2.52	0.44
1:A:213:HIS:CE1	1:A:232:LEU:HD11	2.52	0.44
1:B:213:HIS:CE1	1:B:232:LEU:HD11	2.52	0.44
1:A:164:HIS:CG	1:A:165:PRO:HD2	2.53	0.44
1:B:220:ARG:HD3	3:B:402:SO4:O1	2.17	0.44
1:A:88:TRP:CH2	2:A:501:I43:H15	2.53	0.44
1:A:125:GLN:NE2	4:A:673:HOH:O	2.51	0.43
1:A:159:LEU:HD13	1:A:161:PHE:CZ	2.54	0.43
1:A:118:ALA:HA	1:A:181:SER:O	2.19	0.43
1:B:338:ASP:O	1:B:342:GLN:HG2	2.18	0.43
1:A:217:LEU:HD22	1:A:232:LEU:HD13	2.00	0.42
1:A:276:SER:OG	1:A:277:ASN:N	2.52	0.42
1:A:276:SER:O	1:A:277:ASN:HB2	2.18	0.42
1:A:310:TRP:CZ2	1:B:364:LEU:HD22	2.54	0.42
1:A:70:VAL:HA	1:A:78:ILE:HD13	2.00	0.42
1:B:57:PRO:O	1:B:60:MET:HE2	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:164:HIS:CG	1:B:165:PRO:HD2	2.55	0.42
1:B:48:THR:OG1	1:B:155:PRO:HA	2.20	0.41
1:B:216:ARG:NH1	1:B:329:TYR:O	2.51	0.41
1:A:176:ARG:HA	1:A:176:ARG:HD3	1.90	0.41
1:A:192:ALA:HB2	1:B:134[A]:VAL:HG12	2.03	0.41
1:A:208:LEU:HD12	1:A:209:PRO:HD2	2.03	0.40
1:A:126:ILE:CG1	1:B:126[A]:ILE:HD11	2.52	0.40
1:A:69:LEU:HD11	1:A:341:LEU:HD13	2.04	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	310/331 (94%)	302 (97%)	6 (2%)	2 (1%)	25	15
1	B	311/331 (94%)	304 (98%)	7 (2%)	0	100	100
All	All	621/662 (94%)	606 (98%)	13 (2%)	2 (0%)	41	31

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	263	SER
1	A	219	LYS

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	251/264 (95%)	247 (98%)	4 (2%)	62	60
1	B	249/264 (94%)	244 (98%)	5 (2%)	55	51
All	All	500/528 (95%)	491 (98%)	9 (2%)	59	55

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

Mol	Chain	Res	Type
1	A	134	VAL
1	A	159	LEU
1	A	212	ARG
1	A	343	LEU
1	B	78	ILE
1	B	159	LEU
1	B	212	ARG
1	B	220	ARG
1	B	343	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	125	GLN

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](#)

5 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
3	SO4	A	503	-	4,4,4	0.13	0	6,6,6	0.16	0
2	I43	B	401	-	17,17,17	1.44	2 (11%)	15,23,23	1.74	3 (20%)
3	SO4	A	502	-	4,4,4	0.12	0	6,6,6	0.31	0
2	I43	A	501	-	17,17,17	1.58	3 (17%)	15,23,23	1.85	3 (20%)
3	SO4	B	402	-	4,4,4	0.13	0	6,6,6	0.16	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	I43	A	501	-	-	8/8/10/10	0/1/1/1
2	I43	B	401	-	-	5/8/10/10	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	I43	C1-N1	4.85	1.46	1.41
2	B	401	I43	C1-N1	4.30	1.45	1.41
2	B	401	I43	C1-N2	2.45	1.41	1.36
2	A	501	I43	C1-N2	2.20	1.40	1.36
2	A	501	I43	C11-C5	2.15	1.50	1.41

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	I43	C01-N1-C1	-5.17	113.61	118.05
2	B	401	I43	C01-N1-C1	-4.51	114.17	118.05
2	A	501	I43	C01-N1-C5	3.97	125.08	120.25
2	B	401	I43	C01-N1-C5	3.47	124.47	120.25
2	A	501	I43	O3-C13-C12	-2.14	106.60	112.35
2	B	401	I43	O3-C13-C12	-2.13	106.62	112.35

There are no chirality outliers.

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	I43	N2-C2-O2-C02
2	A	501	I43	C3-C2-O2-C02
2	A	501	I43	C5-C11-C12-C13
2	A	501	I43	C11-C12-C13-O3
2	A	501	I43	C14-C12-C13-O3
2	A	501	I43	C11-C12-C14-O4
2	A	501	I43	C13-C12-C14-O4
2	B	401	I43	N2-C2-O2-C02
2	B	401	I43	C3-C2-O2-C02
2	B	401	I43	C5-C11-C12-C13
2	B	401	I43	C14-C12-C13-O3
2	B	401	I43	C13-C12-C14-O4
2	A	501	I43	C5-C11-C12-C14

There are no ring outliers.

3 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	401	I43	5	0
2	A	501	I43	5	0
3	B	402	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 [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	313/331 (94%)	0.74	45 (14%) 2 2	16, 30, 63, 78	0
1	B	309/331 (93%)	0.64	35 (11%) 5 5	17, 28, 56, 81	0
All	All	622/662 (93%)	0.69	80 (12%) 3 3	16, 29, 60, 81	0

All (80) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	223	PRO	9.9
1	B	64	THR	8.0
1	A	277	ASN	7.0
1	A	46	MET	6.7
1	B	71	ALA	6.6
1	A	227	LEU	6.3
1	A	110	GLY	6.0
1	B	63	THR	5.8
1	B	264	GLY	5.6
1	A	224	GLY	5.5
1	B	153	PRO	5.5
1	B	74	SER	5.4
1	A	263	SER	5.1
1	A	226	ARG	5.0
1	A	264	GLY	4.9
1	B	65	THR	4.8
1	A	74	SER	4.8
1	B	70	VAL	4.8
1	B	374	GLU	4.8
1	A	109	GLN	4.7
1	A	112	ILE	4.7
1	A	355	PRO	4.4
1	A	300	PRO	4.1
1	B	69	LEU	4.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	111	GLU	4.0
1	B	75	ARG	3.9
1	B	67	GLN	3.7
1	A	211	ASP	3.7
1	A	96	THR	3.6
1	B	154	PRO	3.6
1	A	105	HIS	3.5
1	A	274	PRO	3.5
1	A	220	ARG	3.4
1	A	75	ARG	3.4
1	A	374	GLU	3.3
1	B	334	ALA	3.3
1	A	212	ARG	3.3
1	B	73	GLY	3.2
1	A	225	GLU	3.1
1	B	112	ILE	3.0
1	B	68	LEU	3.0
1	A	275	GLN	3.0
1	A	85	MET	2.9
1	A	276	SER	2.8
1	A	91	LEU	2.8
1	B	72	LEU	2.8
1	A	354	THR	2.7
1	B	333	PRO	2.7
1	A	373	GLY	2.7
1	B	355	PRO	2.7
1	A	98	ALA	2.6
1	B	62	LYS	2.6
1	A	54	ILE	2.6
1	B	370	ARG	2.5
1	B	109	GLN	2.5
1	A	92	GLY	2.5
1	B	373	GLY	2.5
1	A	76	ASP	2.5
1	A	215	ASP	2.5
1	B	203[A]	ILE	2.4
1	A	126	ILE	2.4
1	A	359	PRO	2.4
1	B	261	GLN	2.4
1	B	46	MET	2.3
1	A	278	ALA	2.3
1	B	66	THR	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	167	ALA	2.2
1	A	108	ASP	2.2
1	A	229	LEU	2.1
1	B	110	GLY	2.1
1	B	263	SER	2.1
1	A	89	ARG	2.1
1	B	165	PRO	2.1
1	B	61	GLY	2.1
1	A	153	PRO	2.1
1	B	47	PRO	2.1
1	A	218	ALA	2.0
1	A	90	VAL	2.0
1	B	168	ALA	2.0
1	A	94	SER	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	I43	A	501	17/17	0.79	0.23	39,44,55,57	0
2	I43	B	401	17/17	0.90	0.17	30,34,51,51	0
3	SO4	A	503	5/5	0.97	0.26	44,48,55,55	0
3	SO4	B	402	5/5	0.97	0.11	34,43,49,55	0
3	SO4	A	502	5/5	0.98	0.14	33,34,40,42	0

6.5 Other polymers [i](#)

There are no such residues in this entry.