Full wwPDB X-ray Structure Validation Report (i)

May 29, 2020 – 04:22 am BST

PDB ID	:	4A57
Title	:	CRYSTAL STRUCTURE OF TOXOPLASMA GONDII NUCLEOSIDE TRIPHOSPHATE DIPHOSPHOHYDROLASE 3 (NTPDASE3)
Authors	:	Krug, U.; Zebisch, M.; Straeter, N.
Deposited on	:	2011-10-24
Resolution	:	2.00 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 (i)) were used in the production of this report:

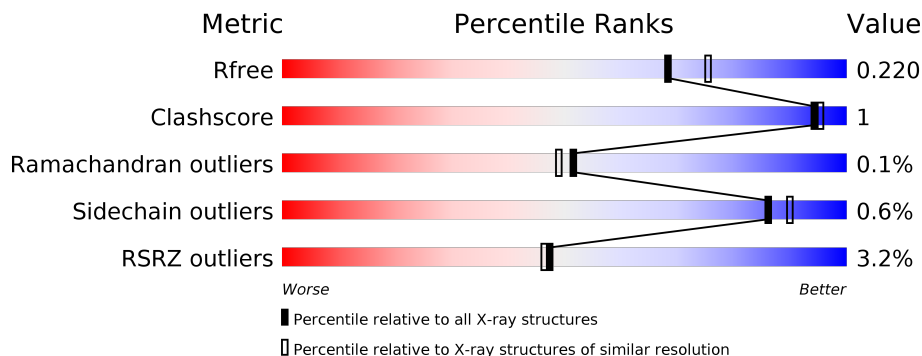|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.00 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	Percentile Ranks	Value
Rfree		0.220
Clashscore		1
Ramachandran outliers		0.1%
Sidechain outliers		0.6%
RSRZ outliers		3.2%

Worse — *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	611	3% / 92% / 5% ·
1	B	611	5% / 93% · ·
1	C	611	% / 94% · ·
1	D	611	3% / 94% · ·

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 19870 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 NUCLEOSIDE-TRIPHOSPHATASE 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	594	Total 4640	C 2916	N 823	O 874	S 27	0	6	0
1	B	594	Total 4617	C 2905	N 813	O 872	S 27	0	3	0
1	C	591	Total 4611	C 2898	N 817	O 869	S 27	0	4	0
1	D	595	Total 4624	C 2905	N 817	O 875	S 27	0	2	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	25	MET	-	expression tag	UNP Q27893
A	629	GLU	-	expression tag	UNP Q27893
A	630	HIS	-	expression tag	UNP Q27893
A	631	HIS	-	expression tag	UNP Q27893
A	632	HIS	-	expression tag	UNP Q27893
A	633	HIS	-	expression tag	UNP Q27893
A	634	HIS	-	expression tag	UNP Q27893
A	635	HIS	-	expression tag	UNP Q27893
B	25	MET	-	expression tag	UNP Q27893
B	629	GLU	-	expression tag	UNP Q27893
B	630	HIS	-	expression tag	UNP Q27893
B	631	HIS	-	expression tag	UNP Q27893
B	632	HIS	-	expression tag	UNP Q27893
B	633	HIS	-	expression tag	UNP Q27893
B	634	HIS	-	expression tag	UNP Q27893
B	635	HIS	-	expression tag	UNP Q27893
C	25	MET	-	expression tag	UNP Q27893
C	629	GLU	-	expression tag	UNP Q27893
C	630	HIS	-	expression tag	UNP Q27893
C	631	HIS	-	expression tag	UNP Q27893
C	632	HIS	-	expression tag	UNP Q27893

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	633	HIS	-	expression tag	UNP Q27893
C	634	HIS	-	expression tag	UNP Q27893
C	635	HIS	-	expression tag	UNP Q27893
D	25	MET	-	expression tag	UNP Q27893
D	629	GLU	-	expression tag	UNP Q27893
D	630	HIS	-	expression tag	UNP Q27893
D	631	HIS	-	expression tag	UNP Q27893
D	632	HIS	-	expression tag	UNP Q27893
D	633	HIS	-	expression tag	UNP Q27893
D	634	HIS	-	expression tag	UNP Q27893
D	635	HIS	-	expression tag	UNP Q27893

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	3	Total Cl 3 3	0	0
2	C	1	Total Cl 1 1	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	451	Total O 452 452	0	1
3	B	306	Total O 307 307	0	1
3	C	366	Total O 366 366	0	0
3	D	248	Total O 249 249	0	1

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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.
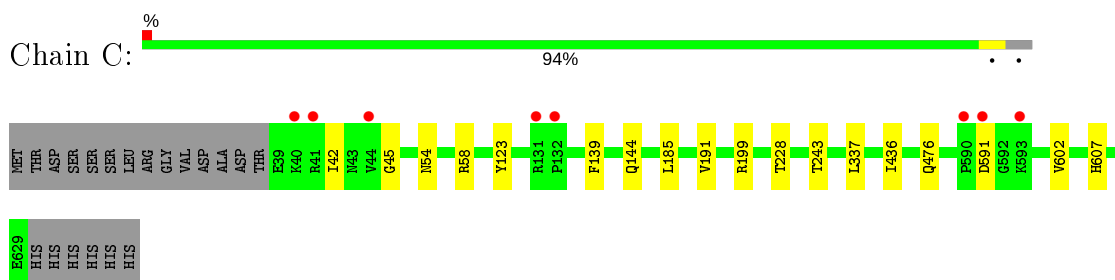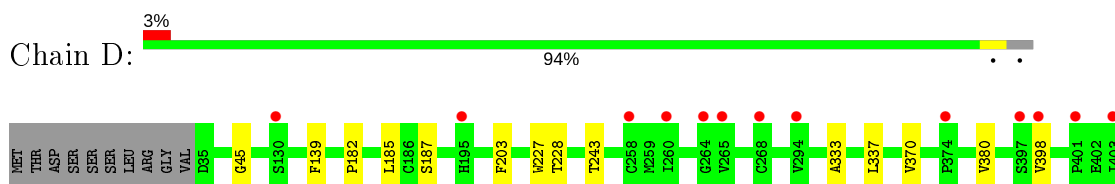
- Molecule 1: NUCLEOSIDE-TRIPHOSPHATASE 1
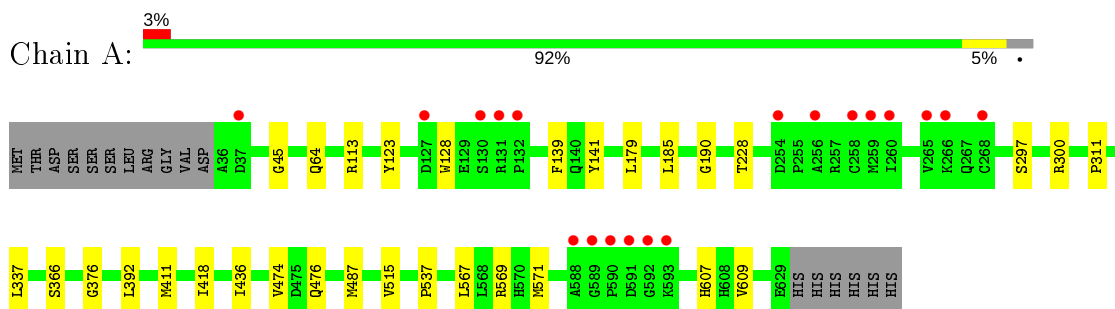
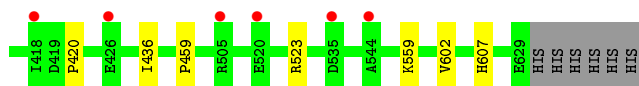- Molecule 1: NUCLEOSIDE-TRIPHOSPHATASE 1
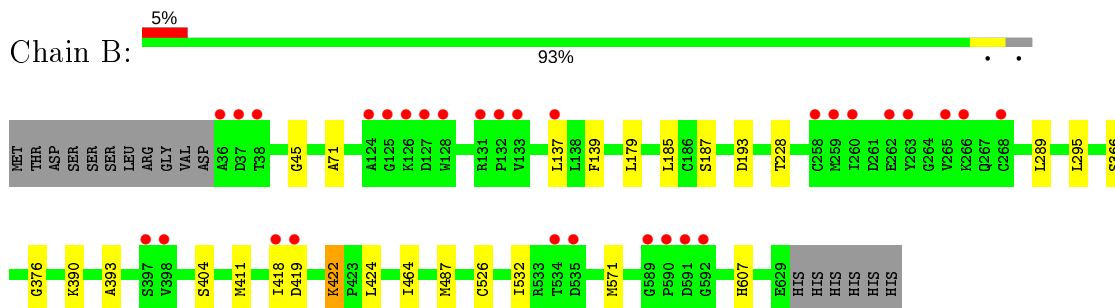
- Molecule 1: NUCLEOSIDE-TRIPHOSPHATASE 1

- Molecule 1: NUCLEOSIDE-TRIPHOSPHATASE 1

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	88.96Å 165.94Å 97.51Å 90.00° 97.03° 90.00°	Depositor
Resolution (Å)	39.43 – 2.00 39.43 – 2.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (39.43-2.00) 97.2 (39.43-2.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.57 (at 2.00Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, R_{free}	0.174 , 0.210 0.179 , 0.220	Depositor DCC
R_{free} test set	1844 reflections (1.01%)	wwPDB-VP
Wilson B-factor (Å^2)	22.8	Xtriage
Anisotropy	0.170	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 51.2	EDS
L-test for twinning[2]	$<|L|> = 0.49$, $<L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	19870	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.71% 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	0.51	0/4748	0.63	0/6424
1	B	0.49	0/4716	0.63	0/6384
1	C	0.50	0/4711	0.63	0/6374
1	D	0.49	0/4719	0.61	0/6386
All	All	0.50	0/18894	0.63	0/25568

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	4640	0	4632	17	0
1	B	4617	0	4600	15	0
1	C	4611	0	4590	10	0
1	D	4624	0	4599	11	0
2	A	3	0	0	2	0
2	C	1	0	0	1	0
3	A	452	0	0	0	0
3	B	307	0	0	1	0
3	C	366	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	249	0	0	1	0
All	All	19870	0	18421	47	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 1.

All (47) 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:419:ASP:HB3	1:B:422:LYS:HD3	1.55	0.89
1:A:300[A]:ARG:HG3	1:B:464:ILE:HD12	1.68	0.76
1:A:487:MET:HG3	1:A:571:MET:HE1	1.72	0.70
1:A:476:GLN:HG3	2:A:1630:CL:CL	2.30	0.69
1:B:289:LEU:HD11	1:B:295[A]:LEU:HD21	1.75	0.67
1:A:113:ARG:HB2	1:A:190:GLY:O	2.00	0.62
1:B:185:LEU:HB3	1:B:228:THR:HG23	1.88	0.55
1:A:45:GLY:HA3	1:D:139:PHE:O	2.08	0.54
1:D:459:PRO:HD2	3:D:2182:HOH:O	2.09	0.53
1:B:45:GLY:HA3	1:C:139:PHE:O	2.11	0.50
1:D:243:THR:HG23	1:D:602:VAL:HB	1.95	0.48
1:A:185:LEU:HB3	1:A:228:THR:HG23	1.96	0.48
1:A:411:MET:HB3	1:A:411:MET:HE2	1.73	0.48
1:C:337:LEU:HD22	1:C:436:ILE:HD13	1.96	0.48
1:B:487:MET:HG3	1:B:571:MET:HE1	1.96	0.47
1:D:185:LEU:HB3	1:D:228:THR:HG23	1.95	0.47
1:C:123:TYR:OH	1:C:144:GLN:HG2	2.15	0.46
1:A:64:GLN:HB2	1:A:179:LEU:HD12	1.97	0.46
1:B:376:GLY:HA3	1:B:532:ILE:HD12	1.96	0.46
1:C:54:ASN:HB3	1:C:58:ARG:HE	1.81	0.46
1:C:191:VAL:O	1:C:199:ARG:HD3	2.16	0.46
1:C:476:GLN:HG3	2:C:1630:CL:CL	2.53	0.45
1:A:515:VAL:HG11	1:A:569:ARG:HG3	1.99	0.45
1:D:398:VAL:HG13	1:D:420:PRO:HB2	1.99	0.45
1:B:139:PHE:O	1:C:45:GLY:HA3	2.16	0.45
1:A:311:PRO:HB2	2:A:1631:CL:CL	2.54	0.45
1:A:392:LEU:HB3	1:A:418:ILE:HD13	1.99	0.45
1:B:390:LYS:HE2	1:B:424:LEU:HD21	1.99	0.44
1:D:187:SER:HG	1:D:203:PHE:HZ	1.64	0.44
1:D:370:VAL:HG22	1:D:380:VAL:HG22	2.00	0.43
1:A:123:TYR:HB3	1:A:141:TYR:CD2	2.53	0.43
1:A:474:VAL:HG11	1:A:567:LEU:CD2	2.49	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:411:MET:HB3	1:B:411:MET:HE2	1.78	0.43
1:B:393:ALA:HA	1:B:418:ILE:HG21	2.01	0.43
1:B:487:MET:HG3	1:B:571:MET:CE	2.49	0.43
1:D:337:LEU:HD22	1:D:436:ILE:HD13	2.00	0.43
1:A:139:PHE:O	1:D:45:GLY:HA3	2.19	0.43
1:D:333:ALA:HA	1:D:559:LYS:HE2	2.01	0.42
1:B:137:LEU:HD23	1:C:42:ILE:HD11	2.02	0.42
1:C:185:LEU:HB3	1:C:228:THR:HG23	2.02	0.42
1:C:243:THR:HG23	1:C:602:VAL:HB	2.01	0.41
1:D:182:PRO:HB3	1:D:227:TRP:CZ3	2.55	0.41
1:A:474:VAL:HG11	1:A:567:LEU:HD23	2.03	0.41
1:A:376:GLY:HA2	1:A:537:PRO:HG3	2.02	0.41
1:A:337:LEU:HD22	1:A:436:ILE:HD13	2.03	0.41
1:B:71:ALA:O	1:B:187:SER:HA	2.20	0.41
1:B:526:CYS:HB3	3:B:2175:HOH:O	2.21	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	598/611 (98%)	588 (98%)	8 (1%)	2 (0%)	41	37
1	B	595/611 (97%)	587 (99%)	8 (1%)	0	100	100
1	C	593/611 (97%)	584 (98%)	9 (2%)	0	100	100
1	D	595/611 (97%)	588 (99%)	7 (1%)	0	100	100
All	All	2381/2444 (97%)	2347 (99%)	32 (1%)	2 (0%)	51	49

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	128	TRP
1	A	297	SER

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	511/521 (98%)	508 (99%)	3 (1%)	86	90
1	B	508/521 (98%)	501 (99%)	7 (1%)	67	72
1	C	507/521 (97%)	505 (100%)	2 (0%)	91	93
1	D	508/521 (98%)	506 (100%)	2 (0%)	91	93
All	All	2034/2084 (98%)	2020 (99%)	14 (1%)	86	88

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	366	SER
1	A	607	HIS
1	A	609	VAL
1	B	179[A]	LEU
1	B	179[B]	LEU
1	B	193	ASP
1	B	366	SER
1	B	404	SER
1	B	422	LYS
1	B	607	HIS
1	C	591	ASP
1	C	607	HIS
1	D	523	ARG
1	D	607	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	348	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	284	GLN
1	C	548	GLN
1	D	348	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 4 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	594/611 (97%)	-0.23	19 (3%)	47	46	13, 24, 55, 92	0
1	B	594/611 (97%)	-0.04	30 (5%)	28	27	16, 30, 64, 126	0
1	C	591/611 (96%)	-0.28	8 (1%)	75	74	14, 25, 50, 88	0
1	D	595/611 (97%)	-0.17	19 (3%)	47	46	13, 31, 60, 81	0
All	All	2374/2444 (97%)	-0.18	76 (3%)	47	46	13, 28, 59, 126	0

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	132	PRO	5.6
1	C	590	PRO	5.2
1	A	592	GLY	4.9
1	B	137	LEU	4.8
1	A	590	PRO	4.4
1	A	258	CYS	4.2
1	B	126	LYS	4.2
1	A	591	ASP	4.2
1	B	591	ASP	4.1
1	B	590	PRO	4.0
1	A	260	ILE	3.8
1	B	125	GLY	3.8
1	B	262	GLU	3.7
1	B	37	ASP	3.6
1	B	259	MET	3.6
1	B	265	VAL	3.6
1	B	258	CYS	3.5
1	B	592	GLY	3.5
1	A	127	ASP	3.4
1	B	418	ILE	3.4
1	B	127	ASP	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	260	ILE	3.4
1	D	398	VAL	3.3
1	B	398	VAL	3.3
1	A	130	SER	3.3
1	A	131	ARG	3.2
1	B	133	VAL	3.2
1	B	131	ARG	3.1
1	B	419	ASP	3.1
1	B	535	ASP	3.1
1	D	374	PRO	2.9
1	C	591	ASP	2.9
1	A	266	LYS	2.9
1	D	265	VAL	2.8
1	B	36	ALA	2.8
1	C	593	LYS	2.7
1	A	259	MET	2.7
1	A	256	ALA	2.7
1	D	418	ILE	2.7
1	D	195	HIS	2.7
1	C	40	LYS	2.7
1	C	132	PRO	2.6
1	B	266	LYS	2.6
1	D	130	SER	2.6
1	B	589	GLY	2.6
1	D	520	GLU	2.6
1	B	397	SER	2.5
1	D	505	ARG	2.5
1	B	124	ALA	2.5
1	D	403	ILE	2.4
1	A	589	GLY	2.4
1	D	258	CYS	2.4
1	D	260	ILE	2.3
1	D	264	GLY	2.3
1	C	41	ARG	2.3
1	D	401	PRO	2.3
1	B	263	TYR	2.3
1	D	426	GLU	2.3
1	A	254	ASP	2.2
1	D	535	ASP	2.2
1	A	268	CYS	2.2
1	D	268	CYS	2.2
1	A	37	ASP	2.2

Continued on next page...

Full wwPDB X-ray Structure Validation Report

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	132	PRO	2.2
1	B	38	THR	2.1
1	B	128	TRP	2.1
1	A	593	LYS	2.1
1	A	588	ALA	2.1
1	C	131	ARG	2.1
1	D	544	ALA	2.1
1	B	268	CYS	2.1
1	A	265	VAL	2.1
1	D	397	SER	2.0
1	B	534	THR	2.0
1	C	44	VAL	2.0
1	D	294	VAL	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 carbohydrates 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	1632	1/1	0.84	0.16	36,36,36,36	1
2	CL	A	1631	1/1	0.98	0.07	29,29,29,29	1
2	CL	A	1630	1/1	0.99	0.05	32,32,32,32	0
2	CL	C	1630	1/1	0.99	0.05	34,34,34,34	0

6.5 Other polymers (i)

There are no such residues in this entry.