



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2017 – 07:06 AM EDT

PDB ID : 2R6F
Title : Crystal Structure of Bacillus stearothermophilus UvrA
Authors : Inuzuka, Y.; Pakotiprapha, D.; Bowman, B.R.; Jeruzalmi, D.; Verdine, G.L.
Deposited on : unknown
Resolution : 3.20 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030345

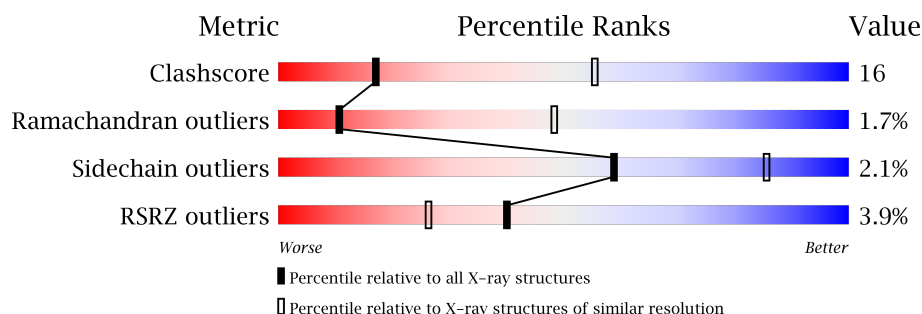
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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))
Clashscore	112137	1009 (3.20-3.20)
Ramachandran outliers	110173	1118 (3.22-3.18)
Sidechain outliers	110143	1117 (3.22-3.18)
RSRZ outliers	101464	1020 (3.22-3.18)

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. 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	972	 3% 66% 25% • 8%
1	B	972	 4% 67% 22% • 10%

2 Entry composition [i](#)

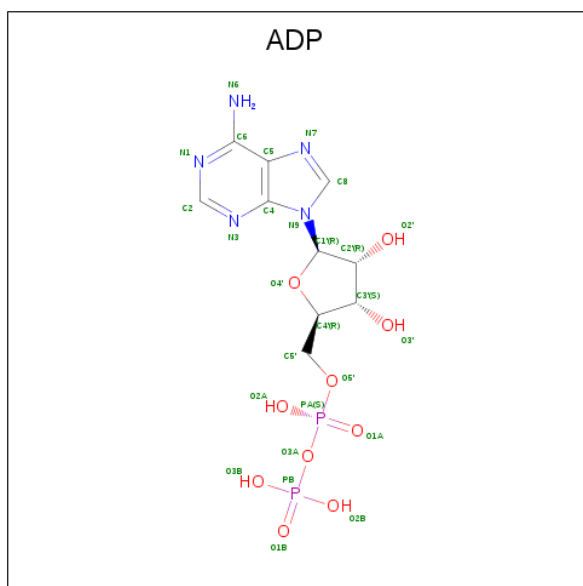
There are 4 unique types of molecules in this entry. The entry contains 13986 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Excinuclease ABC subunit A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	899	Total	C	N	O	S	Se	0	0	0
			7006	4416	1231	1332	12	15			
1	B	879	Total	C	N	O	S	Se	0	0	0
			6862	4318	1212	1305	12	15			

- Molecule 2 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			27	10	5	10	2		
2	A	1	Total	C	N	O	P	0	0
			27	10	5	10	2		
2	B	1	Total	C	N	O	P	0	0
			27	10	5	10	2		
2	B	1	Total	C	N	O	P	0	0
			27	10	5	10	2		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	3	Total 3	Zn 3	0	0
3	A	3	Total 3	Zn 3	0	0

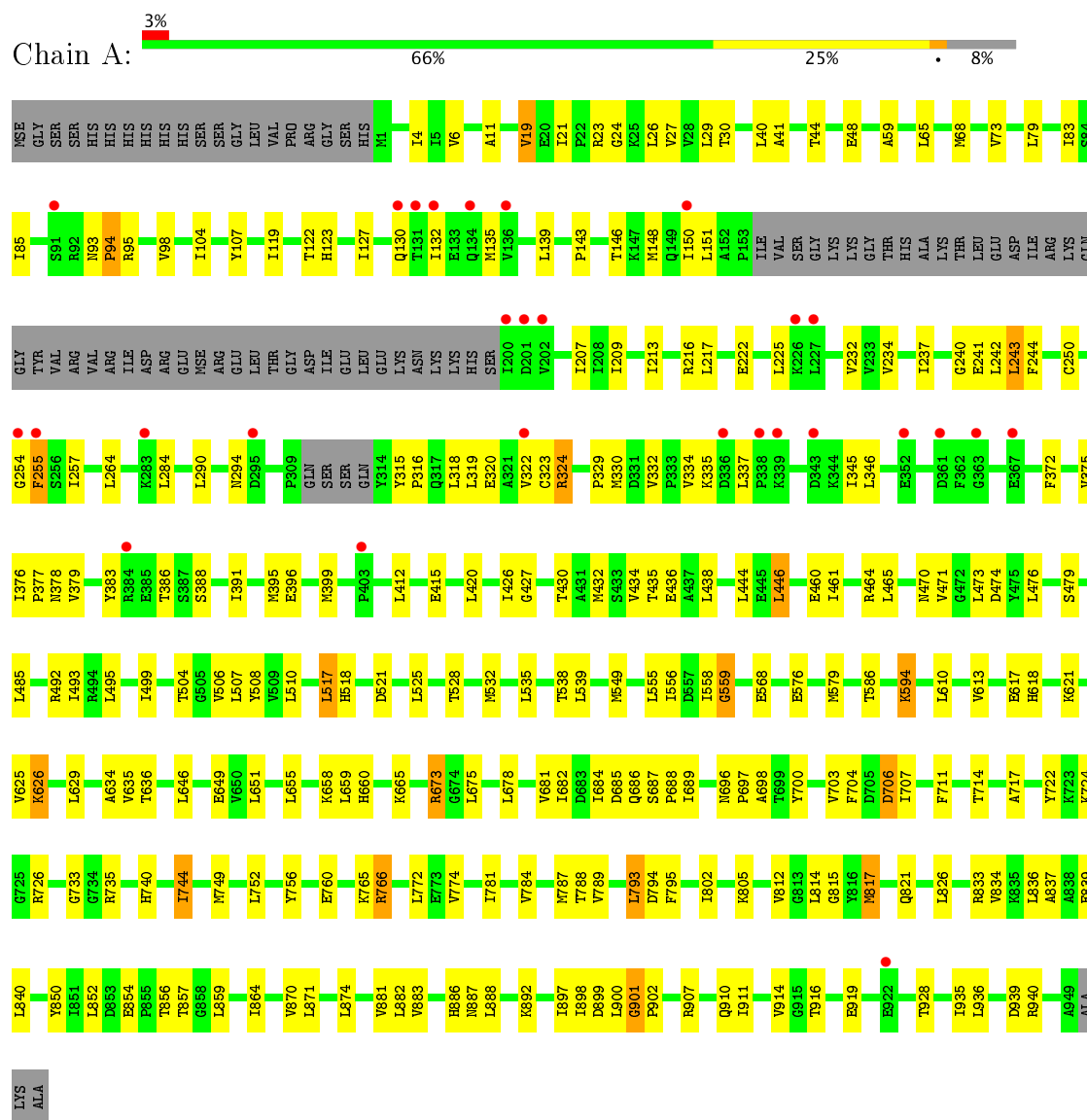
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total 2	O 2	0	0
4	B	2	Total 2	O 2	0	0

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 ($\text{RSRZ} > 2$). 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: Excinuclease ABC subunit A



- Molecule 1: Excinuclease ABC subunit A



L859	I747	L651	D521	E396	K397	LEU	VAL	TYR	MSE
D862	H748	V652	I526	H397	PHE	SER	ARG	GLY	GLY
D863	M749	V652	I526	V398	SER	SER	ARG	SER	SER
L864	H750	L655	M532	M399	GLU	GLY	ARG	HIS	HIS
L867	F751	L655	M532	A400	K247	GLY	ILE	ILE	HIS
L871	D754	L659	T538	L412	C250	GLY	ASP	ASP	HIS
H872	V755	H661	L539	S416	GLU	GLY	ARG	GLY	HIS
V881	E760	A662	T548	V421	I287	MSE	GLU	GLY	HIS
N887	K765	XG65	Y554	H425	G258	ARG	GLU	GLY	SER
L888	R766	L673	L555	I426	L264	LEU	LEU	LEU	GLY
D889	T771	G674	G559	E428	C277	THR	THR	THR	VAL
T893	V774	G674	P560	V429	L284	GLY	ASP	GLY	PRO
I897	I781	L678	E568	T430	L290	GLY	ILE	GLY	ARG
L898	V794	V681	V569	A431	P293	LYS	GLY	LEU	SER
D899	M787	L682	V570	S433	P309	ASN	LYS	GLY	HIS
L900	M787	L684	P575	V434	Q310	LYS	LYS	T131	I4
G901	L793	S687	V578	E436	S311	HIS	HIS	I132	V6
I911	D794	P883	M579	A437	Y315	SER	SER	I127	K7
V920	F795	L689	L585	I457	P316	ILE	ILE	Q128	G8
I935	F796	H694	T586	E460	ASP	ASP	ASP	D137	A11
A949	K805	S695	Y589	I461	P329	VAL	VAL	R138	
ALA	V812	H696	L590	R464	V332	VAL	VAL	L139	
LYS	G813	P897	I596	L465	V332	D205	D205	L140	V19
ALA	L814	A698	K694	L465	K335	R206	R206	S141	E20
	M817	V700	F595	V471	D336	I207	I207	Y142	I21
	Q821	F704	P597	G472	L342	I213	I213	P443	P22
	P822	I707	I598	D474	I345	R216	R216	Q149	R23
	A823	V707	R602	Y475	L346	L217	L217	I150	G24
	L826	V710	V612	L476	R357	I221	I221	L151	R25
	R833	F711	V613	S479	R357	ALA	ALA	VAL	L26
	V834	H714	E617	L485	V365	VAL	VAL	PRO	V27
	R335	M715	H618	R492	F372	A224	A224	ILE	L28
	L836	E716	M619	I493	E373	L225	L225	VAL	L29
	A837	A717	L620	R494	G374	L227	L227	LYS	L40
	L840	K718	V623	I499	V375	GLY	GLY	LYS	T44
	R847	V719	S624	V379	V379	LYS	LYS	THR	E48
	T848	Y722	V625	T504	V379	VAL	VAL	ALA	L57
	L852	R726	A634	G505	R382	VAL	VAL	LYS	M68
	T856	F727	V635	L507	Y383	ASP	ASP	THR	V73
	T857	S728	T636	Y508	T386	VAL	VAL	GLY	I76
		F729		L517	I391	ILE	ILE	ARG	L79
		G733	I646	H518	M395	GLY	GLY	LYS	L79
			V647			GLY	GLY	GLY	I83

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	102.71Å 94.72Å 130.48Å 90.00° 108.80° 90.00°	Depositor
Resolution (Å)	50.00 – 3.20 33.03 – 3.20	Depositor EDS
% Data completeness (in resolution range)	98.5 (50.00-3.20) 98.6 (33.03-3.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.57 (at 3.18Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.253 , 0.292 0.255 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	82.0	Xtriage
Anisotropy	0.477	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 31.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	13986	wwPDB-VP
Average B, all atoms (Å ²)	100.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.26% 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: ZN, ADP

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	0.38	0/7114	0.63	0/9597
1	B	0.36	0/6968	0.60	0/9396
All	All	0.37	0/14082	0.61	0/18993

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	7006	0	7058	241	0
1	B	6862	0	6903	200	0
2	A	54	0	24	0	0
2	B	54	0	24	0	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
All	All	13986	0	14009	436	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 16.

The worst 5 of 436 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:549:MSE:HE1	1:A:586:THR:HG21	1.30	1.07
1:A:379:VAL:HG13	1:A:395:MSE:HE3	1.34	1.06
1:B:812:VAL:HG21	1:B:837:ALA:HB2	1.36	1.06
1:A:284:LEU:HD22	1:A:399:MSE:HE3	1.39	1.05
1:A:812:VAL:HG22	1:A:836:LEU:HD12	1.44	0.98

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	893/972 (92%)	823 (92%)	56 (6%)	14 (2%)	11	50
1	B	873/972 (90%)	787 (90%)	70 (8%)	16 (2%)	10	47
All	All	1766/1944 (91%)	1610 (91%)	126 (7%)	30 (2%)	11	48

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	559	GLY
1	A	617	GLU
1	A	901	GLY
1	B	11	ALA
1	B	94	PRO

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	755/801 (94%)	736 (98%)	19 (2%)	53	83
1	B	739/801 (92%)	727 (98%)	12 (2%)	68	89
All	All	1494/1602 (93%)	1463 (98%)	31 (2%)	59	85

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	744	ILE
1	A	793	LEU
1	B	673	ARG
1	A	760	GLU
1	A	817	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	657	GLN
1	B	669	HIS
1	A	763	HIS
1	A	498	GLN
1	A	660	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 10 ligands modelled in this entry, 6 are monoatomic - leaving 4 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
2	ADP	A	1000	-	25,29,29	1.06	2 (8%)	24,45,45	1.83	2 (8%)
2	ADP	A	1001	-	25,29,29	0.98	1 (4%)	24,45,45	1.84	4 (16%)
2	ADP	B	1002	-	25,29,29	0.95	1 (4%)	24,45,45	1.90	3 (12%)
2	ADP	B	1003	-	25,29,29	0.98	1 (4%)	24,45,45	1.87	3 (12%)

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	ADP	A	1000	-	-	0/12/32/32	0/3/3/3
2	ADP	A	1001	-	-	0/12/32/32	0/3/3/3
2	ADP	B	1002	-	-	0/12/32/32	0/3/3/3
2	ADP	B	1003	-	-	0/12/32/32	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1000	ADP	O4'-C1'	2.15	1.44	1.41
2	B	1002	ADP	C5-C4	2.77	1.46	1.40
2	A	1000	ADP	C5-C4	2.80	1.46	1.40
2	B	1003	ADP	C5-C4	2.86	1.46	1.40
2	A	1001	ADP	C5-C4	2.97	1.47	1.40

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1000	ADP	N3-C2-N1	-7.30	122.50	128.86
2	B	1003	ADP	N3-C2-N1	-7.26	122.53	128.86

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	ADP	N3-C2-N1	-7.11	122.67	128.86
2	B	1002	ADP	N3-C2-N1	-7.08	122.69	128.86
2	B	1002	ADP	C4-C5-N7	-3.26	106.26	109.41

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	884/972 (90%)	0.03	28 (3%)	48	32	57, 86, 148, 198	0
1	B	864/972 (88%)	0.17	40 (4%)	33	20	66, 103, 156, 185	0
All	All	1748/1944 (89%)	0.10	68 (3%)	40	26	57, 94, 155, 198	0

The worst 5 of 68 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	130	GLN	6.7
1	B	221	LEU	5.1
1	B	149	GLN	4.9
1	B	132	ILE	4.9
1	B	206	ARG	4.9

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
2	ADP	A	1000	27/27	0.92	0.21	-0.04	59,65,66,67	0
3	ZN	B	1009	1/1	0.98	0.17	-0.14	99,99,99,99	0
2	ADP	B	1003	27/27	0.93	0.17	-0.67	68,71,72,72	0
2	ADP	A	1001	27/27	0.95	0.17	-0.85	57,68,70,70	0
2	ADP	B	1002	27/27	0.93	0.18	-0.93	75,81,82,82	0
3	ZN	A	1004	1/1	0.88	0.09	-0.96	96,96,96,96	0
3	ZN	A	1006	1/1	0.96	0.11	-1.07	89,89,89,89	0
3	ZN	B	1007	1/1	0.97	0.10	-1.25	94,94,94,94	0
3	ZN	B	1008	1/1	0.92	0.05	-1.75	103,103,103,103	0
3	ZN	A	1005	1/1	0.99	0.05	-2.05	92,92,92,92	0

6.5 Other polymers [i](#)

There are no such residues in this entry.