wwPDB X-ray Structure Validation Summary Report (i)

Sep 24, 2023 – 10:15 AM EDT

PDB ID	:	5U1Z
Title	:	X-ray structure of the WlarG aminotransferase, apo form, from Campylobacter jejune
Authors	:	Holden, H.M.; Thoden, J.B.; Dow, G.T.; Gilbert, M.
Deposited on	:	2016-11-29
Resolution	:	1.60 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

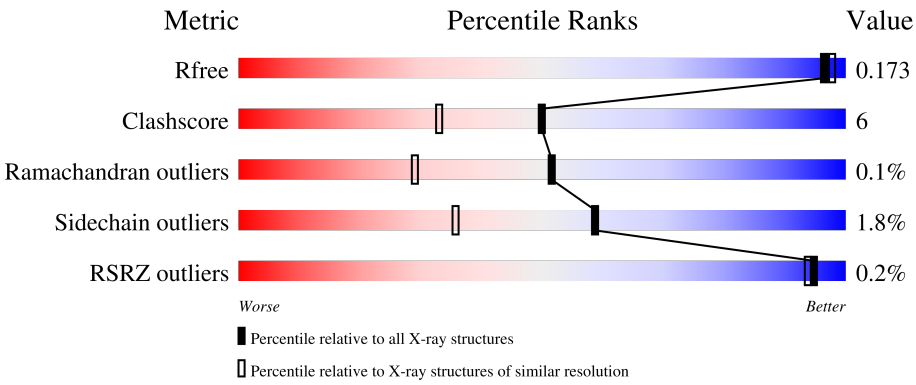
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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 <=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	382	81% 13% • 5%
1	B	382	84% 10% • 5%
1	C	382	84% 11% 5%
1	D	382	80% 13% 6%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12467 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 Putative aminotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	361	2961	1910	501	535	15	0	6	0
1	B	361	2938	1891	498	535	14	0	1	0
1	C	361	2926	1883	495	535	13	0	0	0
1	D	359	2925	1884	495	534	12	0	2	0

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MET	-	initiating methionine	UNP Q9ALS9
A	-20	GLY	-	expression tag	UNP Q9ALS9
A	-19	SER	-	expression tag	UNP Q9ALS9
A	-18	SER	-	expression tag	UNP Q9ALS9
A	-17	HIS	-	expression tag	UNP Q9ALS9
A	-16	HIS	-	expression tag	UNP Q9ALS9
A	-15	HIS	-	expression tag	UNP Q9ALS9
A	-14	HIS	-	expression tag	UNP Q9ALS9
A	-13	HIS	-	expression tag	UNP Q9ALS9
A	-12	HIS	-	expression tag	UNP Q9ALS9
A	-11	SER	-	expression tag	UNP Q9ALS9
A	-10	SER	-	expression tag	UNP Q9ALS9
A	-9	ARG	-	expression tag	UNP Q9ALS9
A	-8	ASN	-	expression tag	UNP Q9ALS9
A	-7	LEU	-	expression tag	UNP Q9ALS9
A	-6	TYR	-	expression tag	UNP Q9ALS9
A	-5	PHE	-	expression tag	UNP Q9ALS9
A	-4	GLN	-	expression tag	UNP Q9ALS9
A	-3	GLY	-	expression tag	UNP Q9ALS9
A	-2	GLY	-	expression tag	UNP Q9ALS9
A	-1	GLY	-	expression tag	UNP Q9ALS9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	HIS	-	expression tag	UNP Q9ALS9
B	-21	MET	-	initiating methionine	UNP Q9ALS9
B	-20	GLY	-	expression tag	UNP Q9ALS9
B	-19	SER	-	expression tag	UNP Q9ALS9
B	-18	SER	-	expression tag	UNP Q9ALS9
B	-17	HIS	-	expression tag	UNP Q9ALS9
B	-16	HIS	-	expression tag	UNP Q9ALS9
B	-15	HIS	-	expression tag	UNP Q9ALS9
B	-14	HIS	-	expression tag	UNP Q9ALS9
B	-13	HIS	-	expression tag	UNP Q9ALS9
B	-12	HIS	-	expression tag	UNP Q9ALS9
B	-11	SER	-	expression tag	UNP Q9ALS9
B	-10	SER	-	expression tag	UNP Q9ALS9
B	-9	ARG	-	expression tag	UNP Q9ALS9
B	-8	ASN	-	expression tag	UNP Q9ALS9
B	-7	LEU	-	expression tag	UNP Q9ALS9
B	-6	TYR	-	expression tag	UNP Q9ALS9
B	-5	PHE	-	expression tag	UNP Q9ALS9
B	-4	GLN	-	expression tag	UNP Q9ALS9
B	-3	GLY	-	expression tag	UNP Q9ALS9
B	-2	GLY	-	expression tag	UNP Q9ALS9
B	-1	GLY	-	expression tag	UNP Q9ALS9
B	0	HIS	-	expression tag	UNP Q9ALS9
C	-21	MET	-	initiating methionine	UNP Q9ALS9
C	-20	GLY	-	expression tag	UNP Q9ALS9
C	-19	SER	-	expression tag	UNP Q9ALS9
C	-18	SER	-	expression tag	UNP Q9ALS9
C	-17	HIS	-	expression tag	UNP Q9ALS9
C	-16	HIS	-	expression tag	UNP Q9ALS9
C	-15	HIS	-	expression tag	UNP Q9ALS9
C	-14	HIS	-	expression tag	UNP Q9ALS9
C	-13	HIS	-	expression tag	UNP Q9ALS9
C	-12	HIS	-	expression tag	UNP Q9ALS9
C	-11	SER	-	expression tag	UNP Q9ALS9
C	-10	SER	-	expression tag	UNP Q9ALS9
C	-9	ARG	-	expression tag	UNP Q9ALS9
C	-8	ASN	-	expression tag	UNP Q9ALS9
C	-7	LEU	-	expression tag	UNP Q9ALS9
C	-6	TYR	-	expression tag	UNP Q9ALS9
C	-5	PHE	-	expression tag	UNP Q9ALS9
C	-4	GLN	-	expression tag	UNP Q9ALS9
C	-3	GLY	-	expression tag	UNP Q9ALS9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	GLY	-	expression tag	UNP Q9ALS9
C	-1	GLY	-	expression tag	UNP Q9ALS9
C	0	HIS	-	expression tag	UNP Q9ALS9
D	-21	MET	-	initiating methionine	UNP Q9ALS9
D	-20	GLY	-	expression tag	UNP Q9ALS9
D	-19	SER	-	expression tag	UNP Q9ALS9
D	-18	SER	-	expression tag	UNP Q9ALS9
D	-17	HIS	-	expression tag	UNP Q9ALS9
D	-16	HIS	-	expression tag	UNP Q9ALS9
D	-15	HIS	-	expression tag	UNP Q9ALS9
D	-14	HIS	-	expression tag	UNP Q9ALS9
D	-13	HIS	-	expression tag	UNP Q9ALS9
D	-12	HIS	-	expression tag	UNP Q9ALS9
D	-11	SER	-	expression tag	UNP Q9ALS9
D	-10	SER	-	expression tag	UNP Q9ALS9
D	-9	ARG	-	expression tag	UNP Q9ALS9
D	-8	ASN	-	expression tag	UNP Q9ALS9
D	-7	LEU	-	expression tag	UNP Q9ALS9
D	-6	TYR	-	expression tag	UNP Q9ALS9
D	-5	PHE	-	expression tag	UNP Q9ALS9
D	-4	GLN	-	expression tag	UNP Q9ALS9
D	-3	GLY	-	expression tag	UNP Q9ALS9
D	-2	GLY	-	expression tag	UNP Q9ALS9
D	-1	GLY	-	expression tag	UNP Q9ALS9
D	0	HIS	-	expression tag	UNP Q9ALS9

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

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

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
3	B	1	1	1	0	0
			Total	Na		
3	C	1	1	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	240	240	240	0	0
			Total	O		
4	B	184	184	184	0	0
			Total	O		
4	C	137	137	137	0	0
			Total	O		
4	D	151	151	151	0	0

3 Residue-property plots (i)

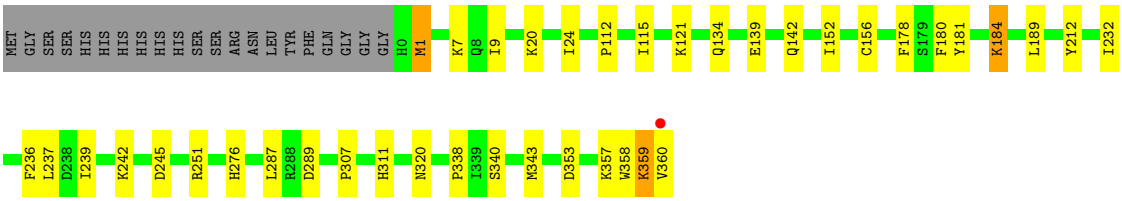
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Putative aminotransferase

Chain A: 81% 13% • 5%

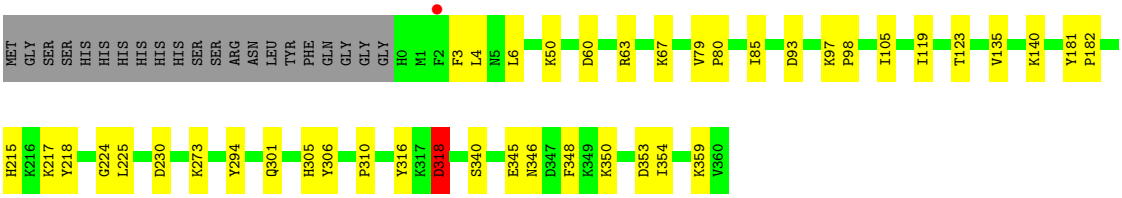
• Molecule 1: Putative aminotransferase

Chain B: 84% 10% • 5%



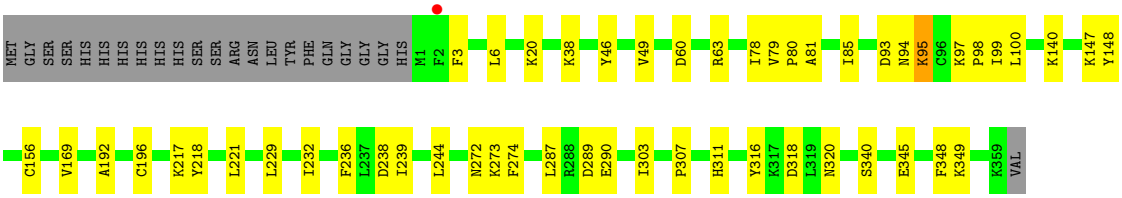
• Molecule 1: Putative aminotransferase

Chain C: 84% 11% 5%



• Molecule 1: Putative aminotransferase

Chain D: 80% 13% 6%


MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER ARG ASN LEU TYR PHE GLN GLY GLY GLY HIS M1 F2 F3 L6 K20 K38 Y46 V49 D60 R63 I78 V79 P80 A81 I85 D93 N94 K95 C96 K97 P98 I99 L100 K140 K147 Y148
C156 V169 A192 C196 K217 Y218 L221 L229 I232 F236 L237 D238 I239 L244 N272 K273 F274 L287 R288 D289 E290 I303 P307 H311 Y316 K317 D318 L319 N320 S340 E345 F348 K349 K359 VAL

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	107.73Å 56.84Å 124.80Å 90.00° 90.04° 90.00°	Depositor
Resolution (Å)	29.97 – 1.60 29.97 – 1.60	Depositor EDS
% Data completeness (in resolution range)	98.6 (29.97-1.60) 99.6 (29.97-1.60)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.13 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.8.0124	Depositor
R, R_{free}	0.148 , 0.188 0.141 , 0.173	Depositor DCC
R_{free} test set	9971 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	11.1	Xtriage
Anisotropy	0.047	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 42.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.023 for h,-k,-l	Xtriage
Reported twinning fraction	0.513 for H, K, L 0.487 for h,-k,-l	Depositor
Outliers	3 of 199163 reflections (0.002%)	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	12467	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.80% 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: NA, 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	0.53	0/3041	0.89	3/4102 (0.1%)
1	B	0.52	0/3004	0.89	4/4055 (0.1%)
1	C	0.49	0/2988	0.85	2/4035 (0.0%)
1	D	0.49	0/2993	0.86	1/4041 (0.0%)
All	All	0.51	0/12026	0.87	10/16233 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	353	ASP	CB-CG-OD1	7.93	125.44	118.30
1	A	168	ARG	NE-CZ-NH2	-6.48	117.06	120.30
1	B	245	ASP	CB-CG-OD1	6.08	123.77	118.30
1	D	289	ASP	CB-CG-OD1	5.98	123.68	118.30
1	B	289	ASP	CB-CG-OD1	5.87	123.58	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	359	LYS	Peptide

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	2961	0	3019	36	0
1	B	2938	0	2969	32	0
1	C	2926	0	2948	34	0
1	D	2925	0	2958	38	0
2	A	2	0	0	0	0
2	D	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	240	0	0	3	0
4	B	184	0	0	0	0
4	C	137	0	0	7	0
4	D	151	0	0	1	0
All	All	12467	0	11894	136	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 6.

The worst 5 of 136 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:24:ILE:HD13	1:B:232:ILE:HD13	1.41	0.99
1:A:288[A]:ARG:NH1	1:A:289:ASP:OD1	1.98	0.97
1:C:63:ARG:CD	1:C:93:ASP:OD2	2.19	0.89
1:C:60:ASP:OD1	1:C:63:ARG:NH2	2.09	0.84
1:C:63:ARG:HD2	1:C:93:ASP:OD2	1.78	0.81

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	365/382 (96%)	349 (96%)	15 (4%)	1 (0%)	41	21
1	B	360/382 (94%)	344 (96%)	16 (4%)	0	100	100
1	C	359/382 (94%)	347 (97%)	11 (3%)	1 (0%)	41	21
1	D	359/382 (94%)	346 (96%)	13 (4%)	0	100	100
All	All	1443/1528 (94%)	1386 (96%)	55 (4%)	2 (0%)	51	29

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	137	ASP
1	C	318	ASP

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	327/339 (96%)	323 (99%)	4 (1%)	71	54
1	B	323/339 (95%)	315 (98%)	8 (2%)	47	22
1	C	320/339 (94%)	313 (98%)	7 (2%)	52	27
1	D	321/339 (95%)	317 (99%)	4 (1%)	71	54
All	All	1291/1356 (95%)	1268 (98%)	23 (2%)	59	36

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	215	HIS
1	C	346	ASN
1	C	340	SER
1	C	359	LYS
1	B	121	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	261	ASN
1	C	279	HIS
1	D	279	HIS
1	D	74	ASN
1	D	272	ASN

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, 5 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	361/382 (94%)	-0.61	0	100	100	5, 10, 20, 43	0
1	B	361/382 (94%)	-0.60	1 (0%)	94	93	5, 11, 21, 52	0
1	C	361/382 (94%)	-0.52	1 (0%)	94	93	6, 13, 25, 38	0
1	D	359/382 (93%)	-0.53	1 (0%)	94	93	6, 12, 26, 42	0
All	All	1442/1528 (94%)	-0.57	3 (0%)	95	94	5, 11, 23, 52	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	360	VAL	5.1
1	C	2	PHE	2.8
1	D	2	PHE	2.5

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	CL	A	401	1/1	0.99	0.08	15,15,15,15	0
2	CL	D	401	1/1	0.99	0.08	16,16,16,16	0
3	NA	B	401	1/1	0.99	0.09	8,8,8,8	0
2	CL	A	402	1/1	1.00	0.03	10,10,10,10	0
3	NA	C	401	1/1	1.00	0.07	9,9,9,9	0

6.5 Other polymers (i)

There are no such residues in this entry.