wwPDB X-ray Structure Validation Summary Report (i)

Sep 26, 2023 – 03:17 AM EDT

PDB ID	:	6BAT
Title	:	Crystal Structure of Wild-Type GltPh in complex with L-aspartate
Authors	:	Font, J.; Scopelliti, A.J.; Vandenberg, R.J.; Boudker, O.; Ryan, R.M.
Deposited on	:	2017-10-15
Resolution	:	3.40 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 (i)) were used in the production of this report:

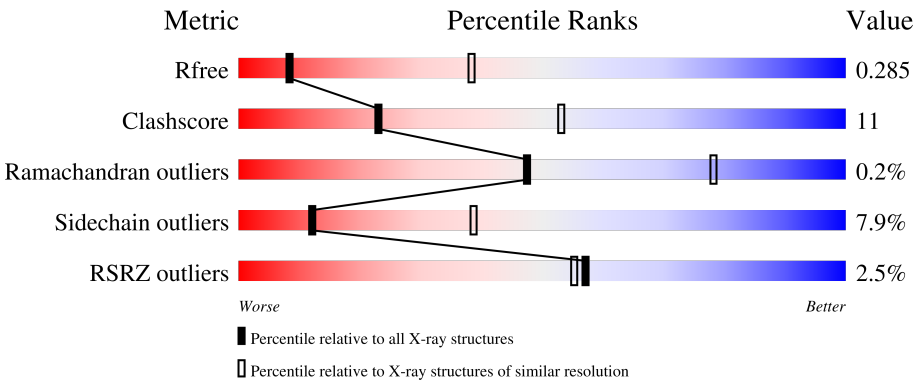
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 3.40 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	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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	420	% 72% 22% • •
1	B	420	5% 68% 28% • •
1	C	420	2% 69% 26% • •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ASP	A	503	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9051 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 Glutamate transporter homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	408	3006	1979	477	533	17	0	0	0
1	B	408	3006	1979	477	533	17	0	0	0
1	C	408	3006	1979	477	533	17	0	0	0

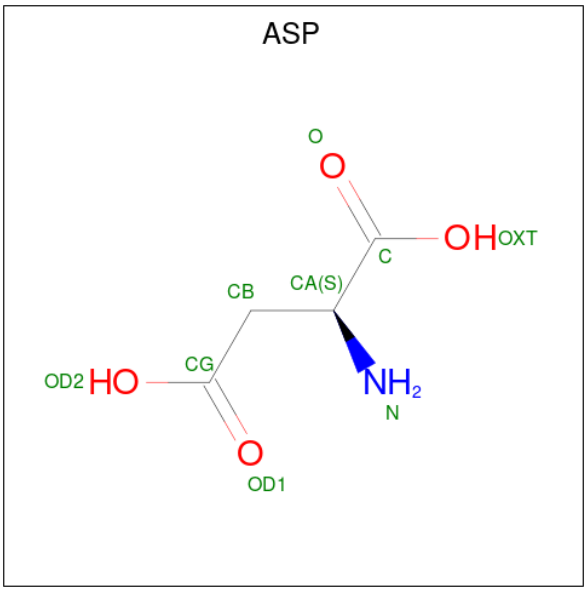
There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	418	THR	-	expression tag	UNP O59010
A	419	LEU	-	expression tag	UNP O59010
A	420	VAL	-	expression tag	UNP O59010
B	418	THR	-	expression tag	UNP O59010
B	419	LEU	-	expression tag	UNP O59010
B	420	VAL	-	expression tag	UNP O59010
C	418	THR	-	expression tag	UNP O59010
C	419	LEU	-	expression tag	UNP O59010
C	420	VAL	-	expression tag	UNP O59010

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
2	A	2	2	2	0	0
2	B	2	2	2	0	0
2	C	2	2	2	0	0

- Molecule 3 is ASPARTIC ACID (three-letter code: ASP) (formula: $C_4H_7NO_4$).



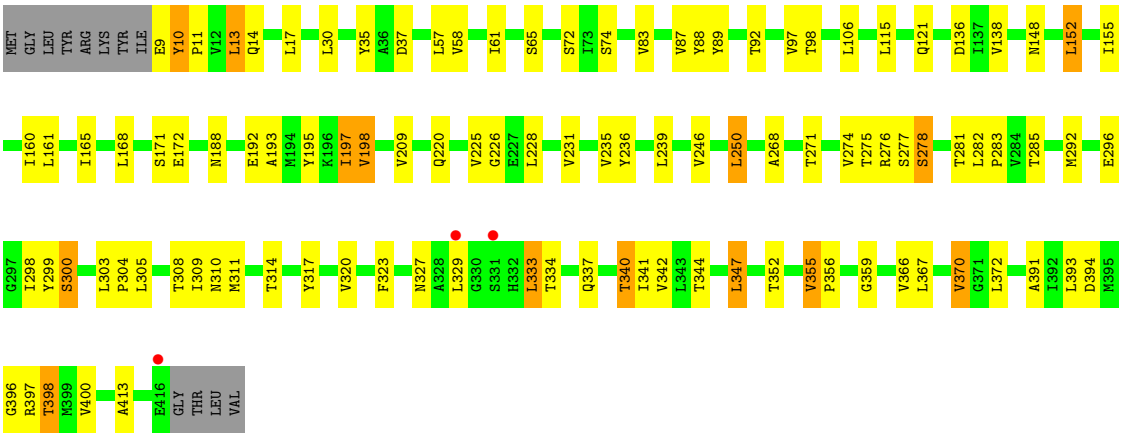
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	9	4	1	4	0	0
			Total	C	N	O		
3	B	1	9	4	1	4	0	0
			Total	C	N	O		
3	C	1	9	4	1	4	0	0

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.

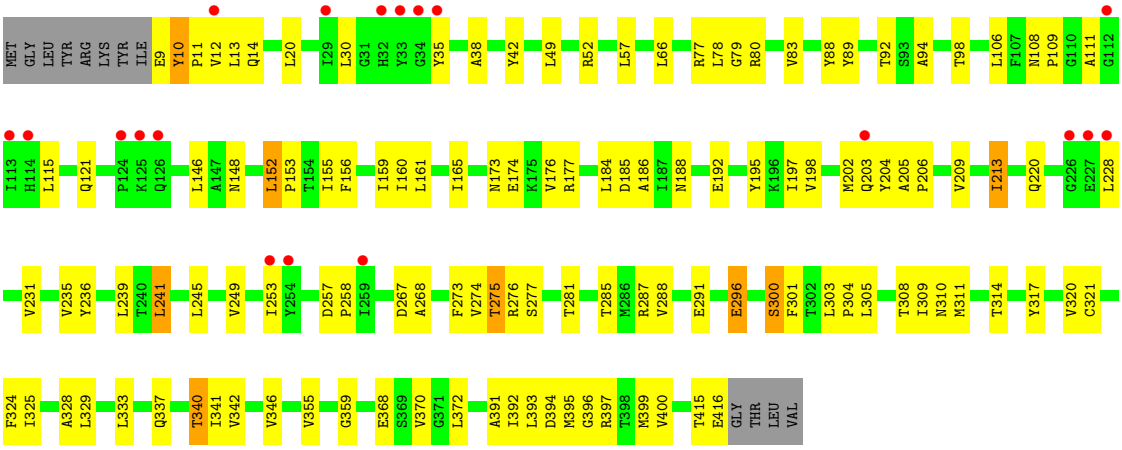
• Molecule 1: Glutamate transporter homolog

Chain A:



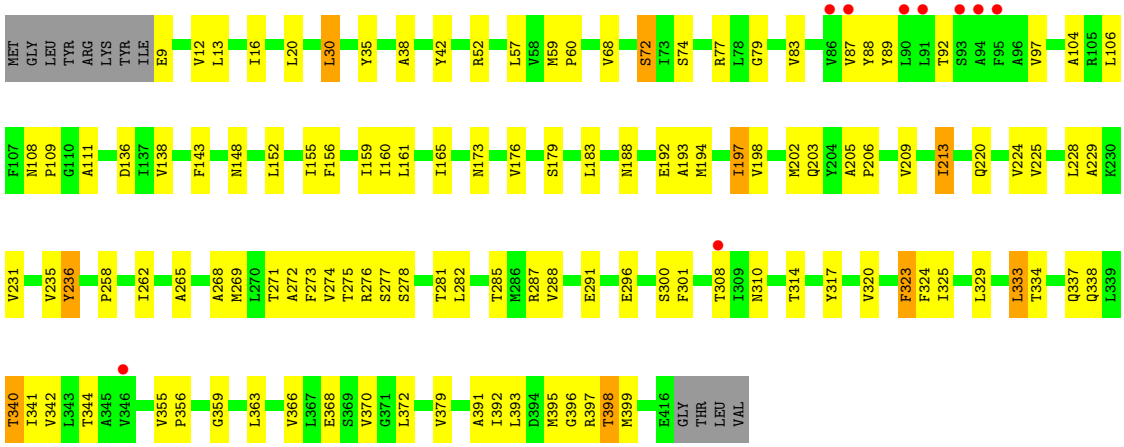
• Molecule 1: Glutamate transporter homolog

Chain B:



• Molecule 1: Glutamate transporter homolog

Chain C:


MET GLY LEU TYR ARG LYS TYR ILE E9 V12 L13 I16 L20 L30 Y35 A38 Y42 R52 L57 V58 M59 P60 V68 S72 I73 S74 R77 L78 G79 V83 V86 V87 Y88 Y89 L90 L91 T92 S93 A94 F95 A96 V97 A104 R105 L106
F107 N108 P109 G110 A111 D136 I137 V138 F143 N148 L152 I155 F156 I159 I160 L161 I165 N173 V176 S179 L183 N188 E192 A193 M194 I197 V198 M202 Q203 Y204 A205 P206 V209 I213 Q220 V224 V225 L228 A229 K230
V231 V235 Y236 P258 I262 A265 A268 M269 L270 T271 A272 F273 V274 T275 R276 S277 S278 T281 L282 T285 M286 R287 V288 E291 E296 S300 F301 T308 I309 N310 T314 Y317 V320 F323 F324 I325 L329 L333 T334 Q337 Q338 L339
T340 I341 V342 L343 T344 A345 V346 V355 P356 G359 L363 V366 L367 E368 S369 V370 G371 L372 V379 A391 I392 L393 D394 M395 G396 R397 T398 M399 E416 GLY THR LEU VAL

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	114.73Å 114.73Å 322.51Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.00 – 3.40 39.36 – 3.40	Depositor EDS
% Data completeness (in resolution range)	97.1 (40.00-3.40) 97.1 (39.36-3.40)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.40 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.247 , 0.286 0.249 , 0.285	Depositor DCC
R_{free} test set	1566 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	120.4	Xtriage
Anisotropy	0.008	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.23 , 47.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.077 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	9051	wwPDB-VP
Average B, all atoms (Å^2)	140.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.12% 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

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	0.39	0/3058	0.55	0/4168
1	B	0.40	0/3058	0.52	0/4168
1	C	0.42	0/3058	0.53	0/4168
All	All	0.40	0/9174	0.54	0/12504

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	3006	0	3198	66	0
1	B	3006	0	3198	75	0
1	C	3006	0	3198	85	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
3	A	9	0	3	4	0
3	B	9	0	3	2	0
3	C	9	0	3	2	0
All	All	9051	0	9603	213	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 11.

The worst 5 of 213 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:268:ALA:HB1	1:A:285:THR:HG22	1.43	1.00
1:B:296:GLU:O	1:B:300:SER:HB2	1.77	0.85
1:C:275:THR:HG23	1:C:277:SER:H	1.43	0.84
1:A:296:GLU:O	1:A:300:SER:HB2	1.79	0.83
1:A:275:THR:HG23	1:A:277:SER:H	1.44	0.82

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	406/420 (97%)	385 (95%)	20 (5%)	1 (0%)	47	78
1	B	406/420 (97%)	386 (95%)	19 (5%)	1 (0%)	47	78
1	C	406/420 (97%)	385 (95%)	20 (5%)	1 (0%)	47	78
All	All	1218/1260 (97%)	1156 (95%)	59 (5%)	3 (0%)	47	78

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	10	TYR
1	C	224	VAL
1	A	10	TYR

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	316/328 (96%)	285 (90%)	31 (10%)	8 28
1	B	316/328 (96%)	293 (93%)	23 (7%)	14 43
1	C	316/328 (96%)	295 (93%)	21 (7%)	16 46
All	All	948/984 (96%)	873 (92%)	75 (8%)	12 39

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	72	SER
1	C	370	VAL
1	C	148	ASN
1	C	323	PHE
1	A	334	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	220	GLN
1	C	327	ASN
1	B	203	GLN
1	B	220	GLN
1	B	327	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 9 ligands modelled in this entry, 6 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	ASP	A	503	-	6,8,8	1.23	1 (16%)	8,10,10	1.43	2 (25%)
3	ASP	B	503	-	6,8,8	1.24	1 (16%)	8,10,10	1.26	2 (25%)
3	ASP	C	503	-	6,8,8	1.18	1 (16%)	8,10,10	1.34	2 (25%)

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
3	ASP	A	503	-	-	2/8/8/8	-
3	ASP	B	503	-	-	4/8/8/8	-
3	ASP	C	503	-	-	4/8/8/8	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	503	ASP	OXT-C	-2.21	1.23	1.30
3	B	503	ASP	OXT-C	-2.18	1.23	1.30
3	C	503	ASP	OXT-C	-2.06	1.23	1.30

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	503	ASP	OXT-C-O	-2.86	117.60	124.09

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	503	ASP	OXT-C-O	-2.58	118.23	124.09
3	B	503	ASP	OXT-C-O	-2.35	118.75	124.09
3	C	503	ASP	OXT-C-CA	2.32	121.30	113.38
3	A	503	ASP	OXT-C-CA	2.28	121.16	113.38

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	503	ASP	C-CA-CB-CG
3	B	503	ASP	N-CA-CB-CG
3	B	503	ASP	C-CA-CB-CG
3	C	503	ASP	C-CA-CB-CG
3	C	503	ASP	OXT-C-CA-N

There are no ring outliers.

3 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	503	ASP	4	0
3	B	503	ASP	2	0
3	C	503	ASP	2	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	408/420 (97%)	-0.31	3 (0%)	87	87	86, 125, 174, 190	0
1	B	408/420 (97%)	-0.23	19 (4%)	31	31	92, 142, 206, 228	0
1	C	408/420 (97%)	-0.42	9 (2%)	62	60	94, 142, 199, 223	0
All	All	1224/1260 (97%)	-0.32	31 (2%)	57	55	86, 137, 196, 228	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	33	TYR	9.4
1	B	125	LYS	5.2
1	B	114	HIS	5.1
1	B	12	VAL	4.6
1	B	32	HIS	4.5

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	NA	B	502	1/1	0.94	0.11	104,104,104,104	0
2	NA	A	502	1/1	0.95	0.09	87,87,87,87	0
2	NA	C	501	1/1	0.97	0.23	116,116,116,116	0
3	ASP	A	503	9/9	0.97	0.11	94,94,95,96	0
2	NA	C	502	1/1	0.98	0.25	105,105,105,105	0
2	NA	A	501	1/1	0.98	0.12	104,104,104,104	0
3	ASP	B	503	9/9	0.98	0.11	111,112,113,114	0
3	ASP	C	503	9/9	0.98	0.12	107,108,110,110	0
2	NA	B	501	1/1	1.00	0.13	119,119,119,119	0

6.5 Other polymers (i)

There are no such residues in this entry.