wwPDB X-ray Structure Validation Summary Report (i)

Feb 19, 2024 – 04:29 AM EST

PDB ID	:	4IZM
Title	:	Crystal structure of GltPh L66C-S300C mutant crosslinked with divalent mercury
Authors	:	Reyes, N.; Boudker, O.
Deposited on	:	2013-01-30
Resolution	:	4.50 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:

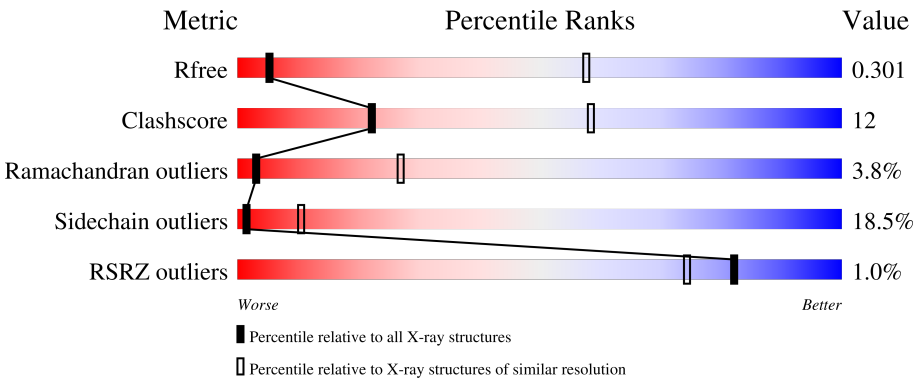
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 4.50 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	1055 (5.20-3.80)
Clashscore	141614	1123 (5.20-3.80)
Ramachandran outliers	138981	1069 (5.20-3.80)
Sidechain outliers	138945	1050 (5.20-3.80)
RSRZ outliers	127900	1101 (5.30-3.70)

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	421	58% 29% 8% •
1	B	421	2% 65% 29% • •
1	C	421	% 66% 26% • •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8727 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 425aa long hypothetical proton glutamate symport protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	404	2897	1903	465	511	18	0	0	0
1	B	404	2897	1903	465	511	18	0	0	0
1	C	404	2897	1903	465	511	18	0	0	0

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	HIS	ASP	engineered mutation	UNP O59010
A	40	HIS	LYS	engineered mutation	UNP O59010
A	66	CYS	LEU	engineered mutation	UNP O59010
A	125	HIS	LYS	engineered mutation	UNP O59010
A	132	HIS	LYS	engineered mutation	UNP O59010
A	223	HIS	LYS	engineered mutation	UNP O59010
A	264	HIS	LYS	engineered mutation	UNP O59010
A	300	CYS	SER	engineered mutation	UNP O59010
A	321	ALA	CYS	engineered mutation	UNP O59010
A	368	HIS	GLU	engineered mutation	UNP O59010
A	418	THR	-	expression tag	UNP O59010
A	419	LEU	-	expression tag	UNP O59010
A	420	VAL	-	expression tag	UNP O59010
A	421	PRO	-	expression tag	UNP O59010
B	37	HIS	ASP	engineered mutation	UNP O59010
B	40	HIS	LYS	engineered mutation	UNP O59010
B	66	CYS	LEU	engineered mutation	UNP O59010
B	125	HIS	LYS	engineered mutation	UNP O59010
B	132	HIS	LYS	engineered mutation	UNP O59010
B	223	HIS	LYS	engineered mutation	UNP O59010
B	264	HIS	LYS	engineered mutation	UNP O59010
B	300	CYS	SER	engineered mutation	UNP O59010
B	321	ALA	CYS	engineered mutation	UNP O59010

Continued on next page...

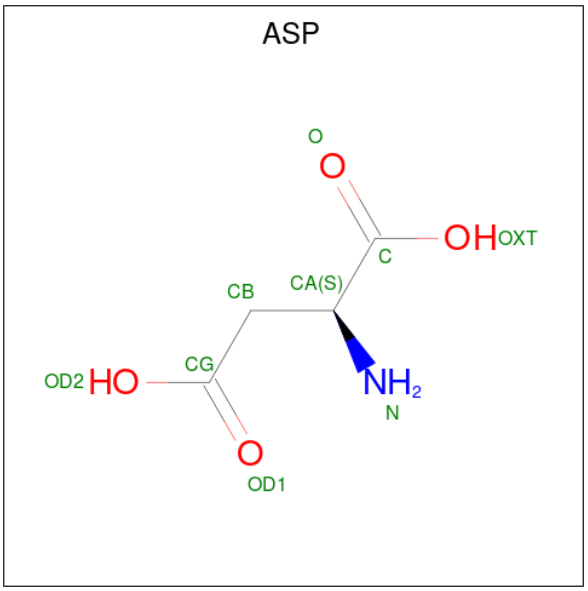
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	368	HIS	GLU	engineered mutation	UNP O59010
B	418	THR	-	expression tag	UNP O59010
B	419	LEU	-	expression tag	UNP O59010
B	420	VAL	-	expression tag	UNP O59010
B	421	PRO	-	expression tag	UNP O59010
C	37	HIS	ASP	engineered mutation	UNP O59010
C	40	HIS	LYS	engineered mutation	UNP O59010
C	66	CYS	LEU	engineered mutation	UNP O59010
C	125	HIS	LYS	engineered mutation	UNP O59010
C	132	HIS	LYS	engineered mutation	UNP O59010
C	223	HIS	LYS	engineered mutation	UNP O59010
C	264	HIS	LYS	engineered mutation	UNP O59010
C	300	CYS	SER	engineered mutation	UNP O59010
C	321	ALA	CYS	engineered mutation	UNP O59010
C	368	HIS	GLU	engineered mutation	UNP O59010
C	418	THR	-	expression tag	UNP O59010
C	419	LEU	-	expression tag	UNP O59010
C	420	VAL	-	expression tag	UNP O59010
C	421	PRO	-	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
			Total	Na		
2	B	2	2	2	0	0
			Total	Na		
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

- Molecule 4 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

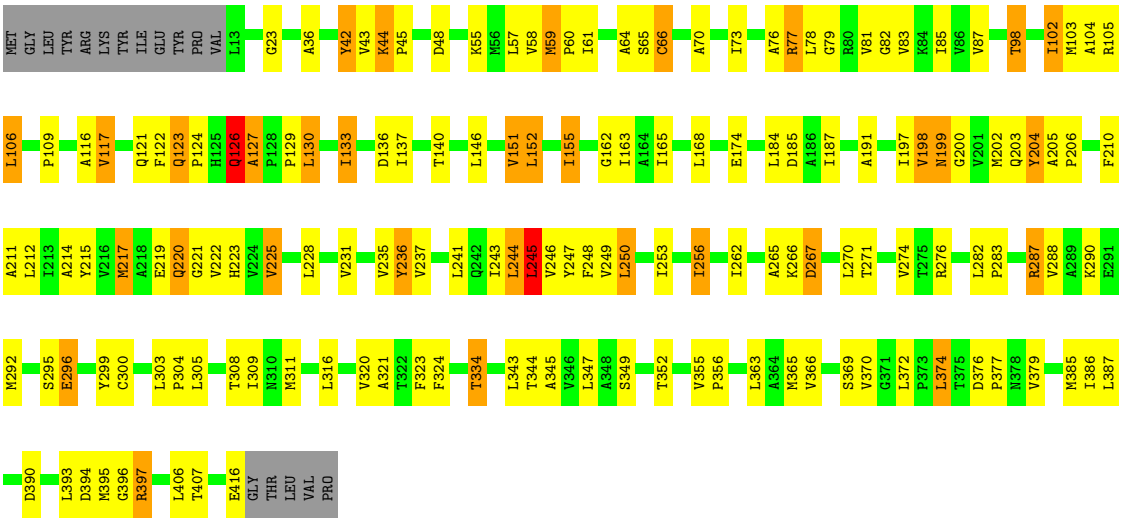
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Hg		
4	A	1	1	1	0	0
			Total	Hg		
4	B	1	1	1	0	0
			Total	Hg		
4	C	1	1	1	0	0

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.

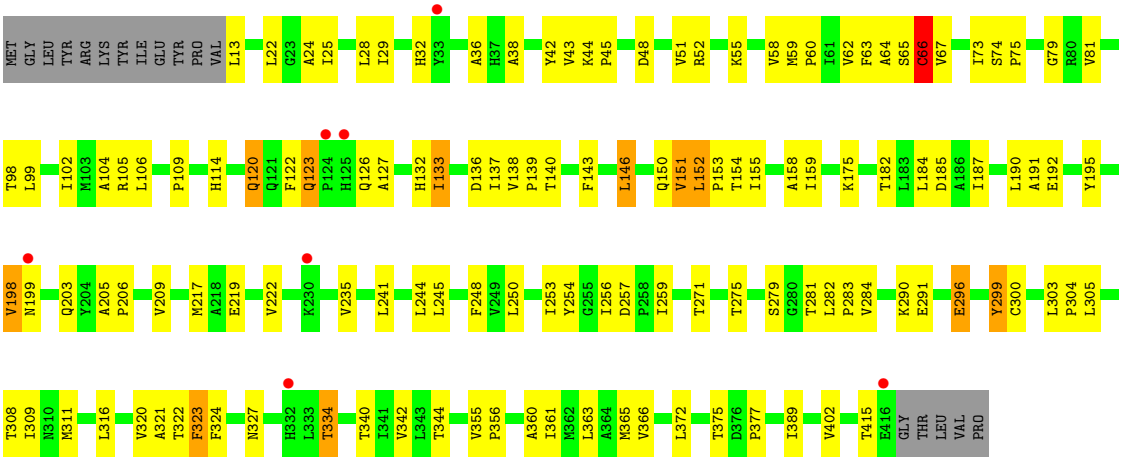
• Molecule 1: 425aa long hypothetical proton glutamate symport protein

Chain A:



• Molecule 1: 425aa long hypothetical proton glutamate symport protein

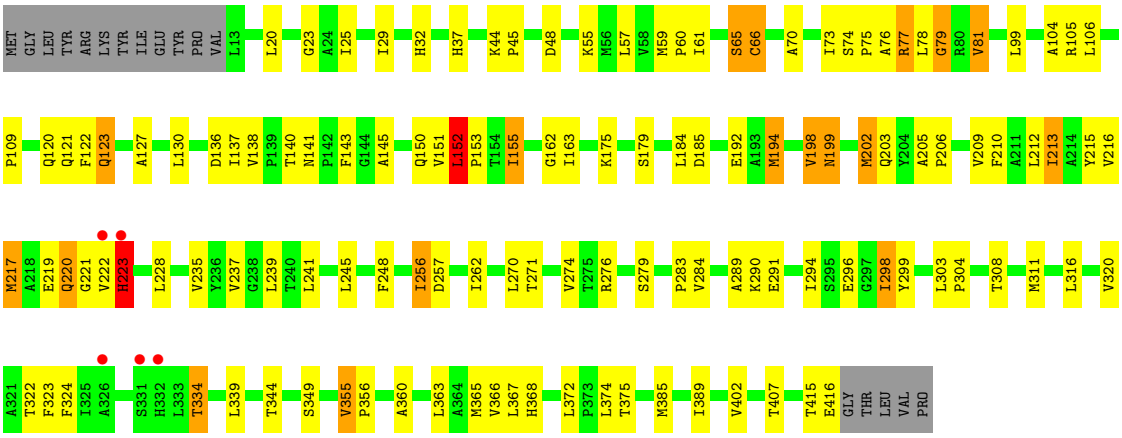
Chain B:



• Molecule 1: 425aa long hypothetical proton glutamate symport protein

Chain C:

%
66%
26%


MET GLY LEU TYR ARG LYS TYR ILE GLU TYR PRO VAL L13 L20 G23 A24 I25 I29 H32 H37 K44 P45 D48 K55 M56 L57 V58 M59 P60 I61 S65 C66 A70 I73 S74 P75 A76 R77 L78 G79 R80 V81 L99 A104 R105 L106
P109 Q120 Q121 F122 Q123 A127 L130 D136 I137 V138 P139 T140 N141 P142 F143 G144 A145 Q150 V151 L152 P153 T154 I155 G162 I163 K175 S179 L184 D185 E192 A193 M194 V198 N199 M202 Q203 Y204 A205 P206 V209 F210 A211 L212 I213 A214 Y215 V216
M217 A218 E219 Q220 G221 V222 H223 L228 V235 Y236 V237 G238 L239 T240 L241 L245 F248 I256 D257 I262 L270 T271 V274 T275 R276 S279 P283 V284 A289 K290 E291 I294 S295 E296 G297 I298 Y299 L303 P304 T308 M311 L316 V320
A321 T322 F323 F324 I325 A326 S331 H332 L333 T334 L339 T344 S349 V355 P356 A360 L363 A364 M365 V366 L367 H368 L372 P373 L374 T375 M385 I389 V402 T407 T415 E416 GLY THR LEU VAL PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	119.71Å 119.71Å 333.72Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 4.50 29.93 – 4.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.00-4.50) 99.9 (29.93-4.50)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.95 (at 4.42Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.250 , 0.299 0.250 , 0.301	Depositor DCC
R_{free} test set	794 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	150.8	Xtriage
Anisotropy	0.485	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.17 , 106.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.098 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	8727	wwPDB-VP
Average B, all atoms (Å^2)	195.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.13% 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, HG

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	0/2952	0.57	1/4040 (0.0%)
1	B	0.36	0/2952	0.52	1/4040 (0.0%)
1	C	0.37	0/2952	0.54	0/4040
All	All	0.37	0/8856	0.54	2/12120 (0.0%)

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	A	0	1
1	B	0	1
1	C	0	1
All	All	0	3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	146	LEU	CA-CB-CG	5.86	128.78	115.30
1	A	146	LEU	CA-CB-CG	5.74	128.50	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	126	GLN	Peptide
1	B	114	HIS	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	C	202	MET	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	2897	0	2960	102	0
1	B	2897	0	2961	56	0
1	C	2897	0	2960	53	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	1	0
3	B	9	0	3	0	0
3	C	9	0	3	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
All	All	8727	0	8890	206	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 12.

The worst 5 of 206 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:296:GLU:HA	1:A:299:TYR:HE1	1.20	1.07
1:A:123:GLN:HB2	1:A:124:PRO:HD3	1.42	1.01
1:A:355:VAL:CG1	1:A:356:PRO:HD2	1.91	1.01
1:A:355:VAL:HG12	1:A:356:PRO:HD2	1.48	0.95
1:A:296:GLU:HA	1:A:299:TYR:CE1	2.01	0.94

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	402/421 (96%)	325 (81%)	60 (15%)	17 (4%)	3	25
1	B	402/421 (96%)	344 (86%)	48 (12%)	10 (2%)	5	35
1	C	402/421 (96%)	342 (85%)	41 (10%)	19 (5%)	2	24
All	All	1206/1263 (96%)	1011 (84%)	149 (12%)	46 (4%)	3	27

5 of 46 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	66	CYS
1	A	122	PHE
1	A	123	GLN
1	A	256	ILE
1	B	152	LEU

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	287/328 (88%)	234 (82%)	53 (18%)	1	10
1	B	287/328 (88%)	233 (81%)	54 (19%)	1	10
1	C	287/328 (88%)	234 (82%)	53 (18%)	1	10
All	All	861/984 (88%)	701 (81%)	160 (19%)	1	10

5 of 160 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	106	LEU
1	C	291	GLU
1	C	136	ASP
1	C	198	VAL
1	C	355	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	148	ASN
1	B	173	ASN
1	C	199	ASN
1	C	148	ASN
1	C	150	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)

Of 12 ligands modelled in this entry, 9 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	C	503	-	6,8,8	1.22	1 (16%)	8,10,10	1.46	2 (25%)
3	ASP	A	503	-	6,8,8	1.18	1 (16%)	8,10,10	1.47	2 (25%)
3	ASP	B	503	-	6,8,8	1.19	1 (16%)	8,10,10	1.45	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	C	503	-	-	3/8/8/8	-
3	ASP	A	503	-	-	2/8/8/8	-
3	ASP	B	503	-	-	4/8/8/8	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	503	ASP	OXT-C	-2.22	1.23	1.30
3	B	503	ASP	OXT-C	-2.21	1.23	1.30
3	A	503	ASP	OXT-C	-2.12	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	C	503	ASP	OXT-C-O	-2.95	117.39	124.09
3	A	503	ASP	OXT-C-O	-2.81	117.70	124.09
3	B	503	ASP	OXT-C-O	-2.74	117.87	124.09
3	B	503	ASP	OXT-C-CA	2.26	121.09	113.38
3	A	503	ASP	OXT-C-CA	2.26	121.07	113.38

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	503	ASP	CA-CB-CG-OD1
3	A	503	ASP	CA-CB-CG-OD2
3	B	503	ASP	CA-CB-CG-OD1
3	B	503	ASP	CA-CB-CG-OD2
3	C	503	ASP	CA-CB-CG-OD1

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	503	ASP	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	404/421 (95%)	-0.55	0 100 100	128, 153, 214, 285	0
1	B	404/421 (95%)	-0.34	7 (1%) 70 61	143, 212, 283, 308	0
1	C	404/421 (95%)	-0.40	5 (1%) 79 70	142, 206, 293, 328	0
All	All	1212/1263 (95%)	-0.43	12 (0%) 82 74	128, 191, 280, 328	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	331	SER	3.5
1	B	33	TYR	3.1
1	B	416	GLU	3.0
1	C	222	VAL	2.9
1	B	199	ASN	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
3	ASP	C	503	9/9	0.71	0.31	177,182,185,188	0
3	ASP	B	503	9/9	0.81	0.44	187,189,190,192	0
3	ASP	A	503	9/9	0.94	0.25	130,131,131,131	0
2	NA	C	502	1/1	0.97	0.27	184,184,184,184	0
2	NA	A	501	1/1	0.97	0.10	127,127,127,127	0
2	NA	A	502	1/1	0.97	0.28	124,124,124,124	0
2	NA	C	501	1/1	0.97	0.11	182,182,182,182	0
4	HG	B	504	1/1	0.97	0.04	208,208,208,208	0
2	NA	B	501	1/1	0.99	0.12	178,178,178,178	0
2	NA	B	502	1/1	0.99	0.26	187,187,187,187	0
4	HG	C	504	1/1	0.99	0.03	215,215,215,215	0
4	HG	A	504	1/1	1.00	0.06	172,172,172,172	0

6.5 Other polymers (i)

There are no such residues in this entry.