



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 02:09 am GMT

PDB ID : 4FPA
Title : Crystal Structure of recombinant human Hexokinase type I mutant D413N
Glucose 6-Phosphate
Authors : Shen, L.; Honzatko, R.B.
Deposited on : 2012-06-21
Resolution : 2.48 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 : trunk28620
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) : recalc28949

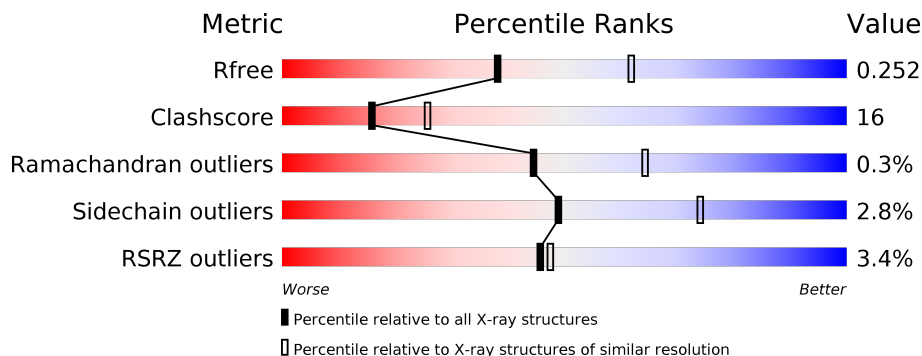
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.48 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}	100719	4719 (2.50-2.46)
Clashscore	112137	5483 (2.50-2.46)
Ramachandran outliers	110173	5388 (2.50-2.46)
Sidechain outliers	110143	5390 (2.50-2.46)
RSRZ outliers	101464	4754 (2.50-2.46)

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	917	<div> <div>3%</div> <div> <div></div> <div>69%</div> <div>27%</div> <div>..</div> </div> </div>
1	B	917	<div> <div>3%</div> <div> <div></div> <div>71%</div> <div>26%</div> <div>..</div> </div> </div>

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
2	BGC	B	1001	-	-	-	X
5	CIT	B	1007	-	-	-	X

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 14391 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.

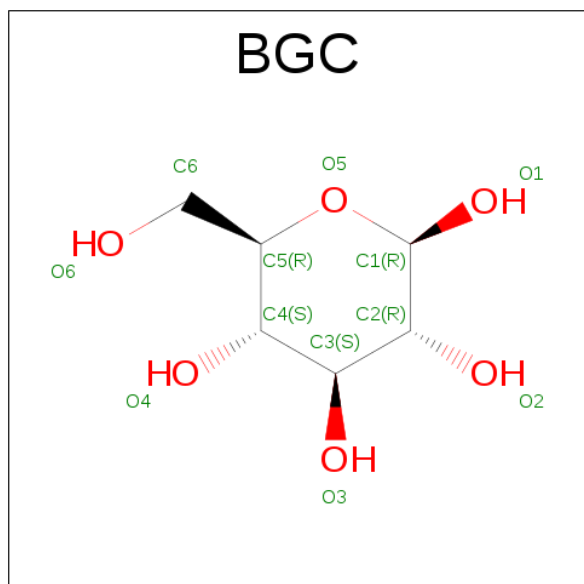
- Molecule 1 is a protein called Hexokinase-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	899	Total	C	N	O	S	0	0	0
			7032	4407	1241	1331	53			
1	B	899	Total	C	N	O	S	0	0	0
			7032	4407	1241	1331	53			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	413	ASN	ASP	ENGINEERED MUTATION	UNP P19367
B	413	ASN	ASP	ENGINEERED MUTATION	UNP P19367

- Molecule 2 is SUGAR (BETA-D-GLUCOSE) (three-letter code: BGC) (formula: C₆H₁₂O₆).



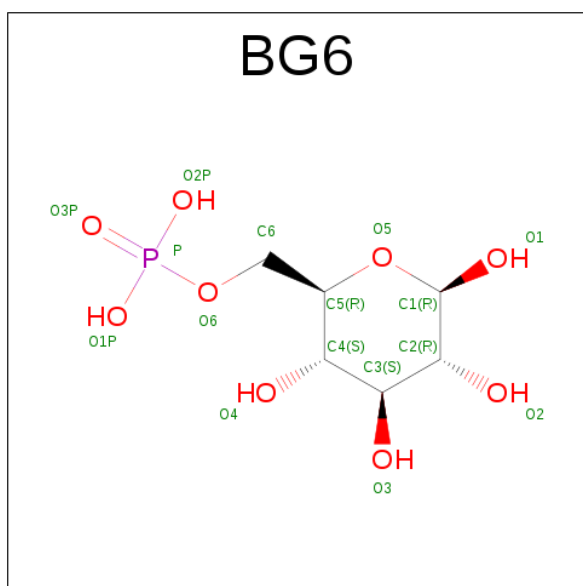
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		
2	B	1	Total	C	O	0	0
			12	6	6		
2	B	1	Total	C	O	0	0
			12	6	6		

- Molecule 3 is SUGAR (BETA-D-GLUCOSE-6-PHOSPHATE) (three-letter code: BG6) (formula: $C_6H_{13}O_9P$).

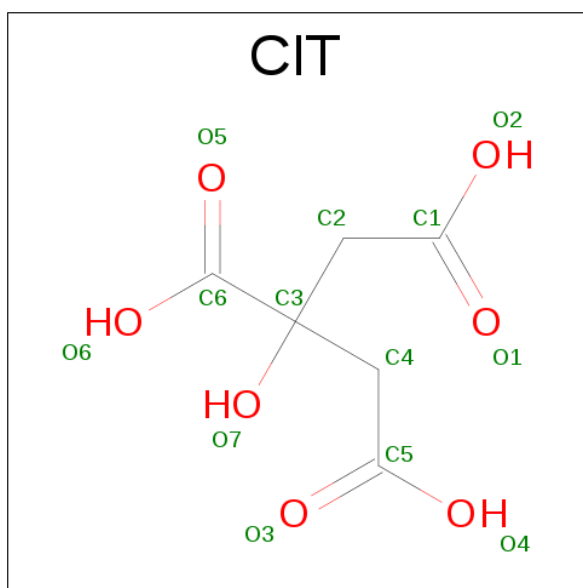


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	O	P	0	0
			16	6	9	1		
3	A	1	Total	C	O	P	0	0
			16	6	9	1		
3	B	1	Total	C	O	P	0	0
			16	6	9	1		
3	B	1	Total	C	O	P	0	0
			16	6	9	1		

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

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

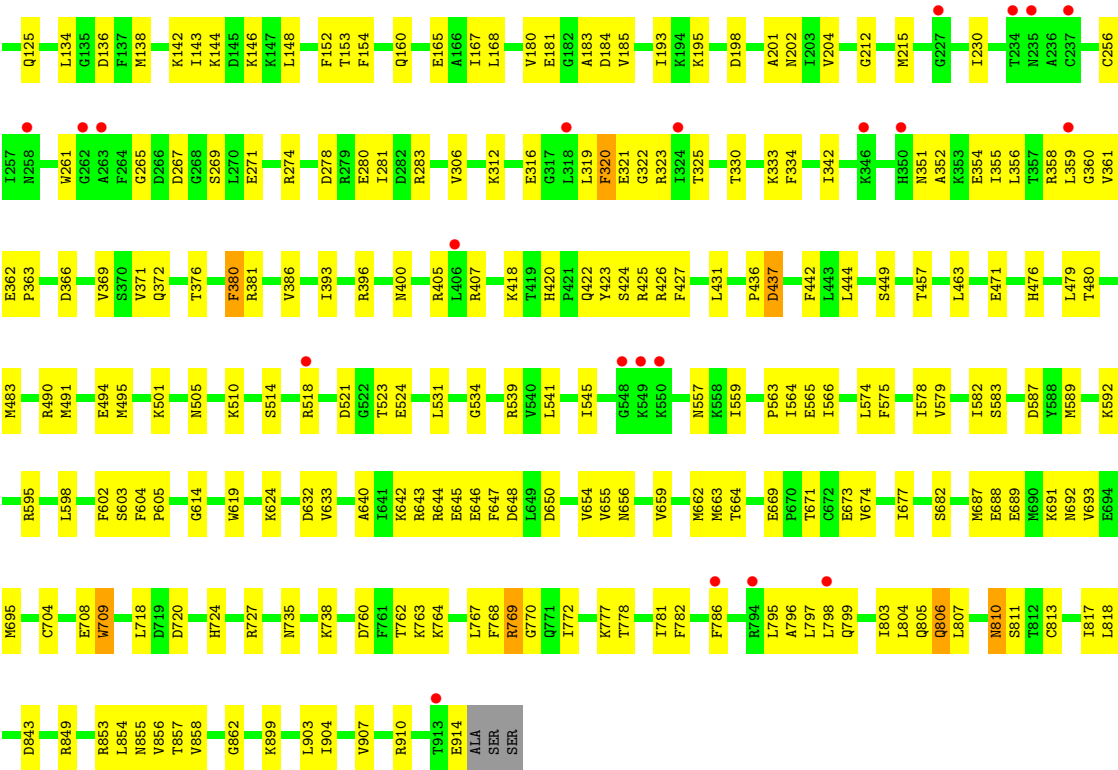
- Molecule 5 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			13	6	7		
5	B	1	Total	C	O	0	0
			13	6	7		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	77	Total	O	0	0
			77	77		
6	B	108	Total	O	0	0
			108	108		



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	82.61Å 121.14Å 119.89Å 90.00° 92.69° 90.00°	Depositor
Resolution (Å)	44.15 – 2.48 44.15 – 2.48	Depositor EDS
% Data completeness (in resolution range)	98.5 (44.15-2.48) 98.5 (44.15-2.48)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.51 (at 2.48Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, R_{free}	0.238 , 0.256 0.235 , 0.252	Depositor DCC
R_{free} test set	4126 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	51.7	Xtriage
Anisotropy	0.121	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 32.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.007 for -h,-l,-k 0.000 for -h,l,k 0.019 for h,-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	14391	wwPDB-VP
Average B, all atoms (Å ²)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.30% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, BGC, BG6, CIT

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	0.26	0/7138	0.45	0/9606
1	B	0.25	0/7138	0.44	0/9606
All	All	0.26	0/14276	0.44	0/19212

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	7032	0	7092	224	0
1	B	7032	0	7092	246	0
2	A	24	0	24	2	0
2	B	24	0	24	0	0
3	A	32	0	22	2	0
3	B	32	0	22	1	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
5	A	13	0	5	1	0
5	B	13	0	5	5	0
6	A	77	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	108	0	0	6	0
All	All	14391	0	14286	465	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 465 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:132:GLU:HG3	1:A:196:ARG:NH2	1.62	1.15
1:B:797:LEU:HD11	1:B:817:ILE:HD11	1.26	1.13
1:B:518:ARG:HH12	1:B:521:ASP:HB3	1.14	1.09
1:B:518:ARG:NH1	1:B:521:ASP:HB3	1.67	1.06
1:B:480:THR:N	1:B:483:MET:HE3	1.70	1.06

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	897/917 (98%)	840 (94%)	55 (6%)	2 (0%)	51 71
1	B	897/917 (98%)	845 (94%)	49 (6%)	3 (0%)	44 64
All	All	1794/1834 (98%)	1685 (94%)	104 (6%)	5 (0%)	44 64

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	104	ASN
1	B	102	ASN
1	A	243	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	450	GLY
1	B	436	PRO

5.3.2 Protein sidechains ⓘ

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	774/788 (98%)	749 (97%)	25 (3%)	44	69
1	B	774/788 (98%)	756 (98%)	18 (2%)	56	79
All	All	1548/1576 (98%)	1505 (97%)	43 (3%)	49	74

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	565	GLU
1	A	913	THR
1	B	805	GLN
1	A	709	TRP
1	A	810	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	806	GLN
1	B	104	ASN
1	B	810	ASN
1	A	810	ASN
1	A	887	ASN

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 14 ligands modelled in this entry, 4 are monoatomic - leaving 10 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	BGC	A	1001	-	12,12,12	0.51	0	17,17,17	1.50	2 (11%)
3	BG6	A	1002	-	16,16,16	1.10	1 (6%)	24,24,24	0.82	1 (4%)
2	BGC	A	1003	-	12,12,12	0.37	0	17,17,17	1.26	2 (11%)
3	BG6	A	1004	-	16,16,16	1.01	1 (6%)	24,24,24	0.97	1 (4%)
5	CIT	A	1007	-	3,12,12	1.22	0	3,17,17	1.60	1 (33%)
2	BGC	B	1001	-	12,12,12	0.37	0	17,17,17	1.42	2 (11%)
3	BG6	B	1002	-	16,16,16	1.02	1 (6%)	24,24,24	0.73	1 (4%)
2	BGC	B	1003	-	12,12,12	0.48	0	17,17,17	1.35	3 (17%)
3	BG6	B	1004	-	16,16,16	1.06	1 (6%)	24,24,24	0.83	1 (4%)
5	CIT	B	1007	-	3,12,12	1.28	0	3,17,17	1.74	1 (33%)

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	BGC	A	1001	-	-	0/2/22/22	0/1/1/1
3	BG6	A	1002	-	-	0/6/26/26	0/1/1/1
2	BGC	A	1003	-	-	0/2/22/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	BG6	A	1004	-	-	0/6/26/26	0/1/1/1
5	CIT	A	1007	-	-	0/6/16/16	0/0/0/0
2	BGC	B	1001	-	-	0/2/22/22	0/1/1/1
3	BG6	B	1002	-	-	0/6/26/26	0/1/1/1
2	BGC	B	1003	-	-	0/2/22/22	0/1/1/1
3	BG6	B	1004	-	-	0/6/26/26	0/1/1/1
5	CIT	B	1007	-	-	0/6/16/16	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1004	BG6	P-O3P	2.96	1.60	1.50
3	B	1004	BG6	P-O3P	3.09	1.61	1.50
3	B	1002	BG6	P-O3P	3.17	1.61	1.50
3	A	1002	BG6	P-O3P	3.36	1.62	1.50

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	BGC	C1-O5-C5	-4.43	105.41	113.39
2	B	1001	BGC	C1-O5-C5	-3.73	106.67	113.39
2	B	1001	BGC	O5-C1-C2	-3.69	103.92	110.04
2	A	1003	BGC	O5-C1-C2	-3.31	104.55	110.04
2	B	1003	BGC	C1-O5-C5	-3.25	107.53	113.39

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1002	BG6	2	0
2	A	1003	BGC	2	0
5	A	1007	CIT	1	0
3	B	1004	BG6	1	0
5	B	1007	CIT	5	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	899/917 (98%)	0.21	32 (3%) 43 45	31, 59, 85, 113	0
1	B	899/917 (98%)	0.21	29 (3%) 48 50	32, 58, 85, 112	0
All	All	1798/1834 (98%)	0.21	61 (3%) 46 47	31, 59, 85, 113	0

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	22	ILE	5.4
1	A	16	ASP	4.6
1	B	16	ASP	4.3
1	B	19	VAL	4.3
1	A	564	ILE	4.0

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. 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
5	CIT	B	1007	13/13	0.56	0.27	5.76	106,108,108,108	0
2	BGC	B	1001	12/12	0.94	0.33	2.43	48,49,50,51	0
2	BGC	A	1001	12/12	0.89	0.27	1.49	50,52,53,55	0
2	BGC	A	1003	12/12	0.96	0.21	1.35	36,38,38,41	0
3	BG6	A	1002	16/16	0.93	0.23	0.45	75,78,80,80	0
3	BG6	A	1004	16/16	0.98	0.18	0.35	36,41,42,43	0
2	BGC	B	1003	12/12	0.93	0.18	0.33	38,39,40,42	0
3	BG6	B	1004	16/16	0.96	0.18	0.24	36,43,44,45	0
3	BG6	B	1002	16/16	0.95	0.21	0.09	71,74,75,76	0
4	NA	B	1006	1/1	0.92	0.07	-1.83	52,52,52,52	0
4	NA	B	1005	1/1	0.90	0.09	-1.84	65,65,65,65	0
4	NA	A	1006	1/1	0.91	0.05	-1.96	53,53,53,53	0
4	NA	A	1005	1/1	0.89	0.08	-2.63	63,63,63,63	0
5	CIT	A	1007	13/13	0.72	0.36	-	110,110,111,111	0

6.5 Other polymers [i](#)

There are no such residues in this entry.