



wwPDB X-ray Structure Validation Summary Report

Feb 13, 2017 – 05:38 pm GMT

PDB ID : 2VF5
Title : GLUCOSAMINE-6-PHOSPHATE SYNTHASE IN COMPLEX WITH GLUCOSAMINE-6-PHOSPHATE
Authors : Mouilleron, S.; Golinelli-Pimpaneau, B.
Deposited on : 2007-10-30
Resolution : 2.90 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

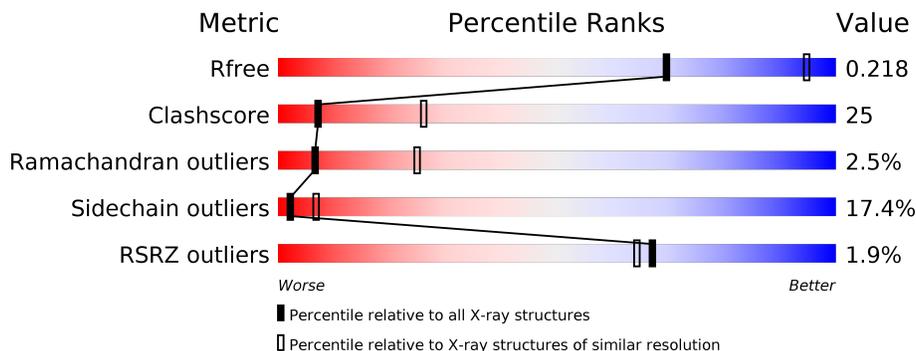
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1586 (2.90-2.90)
Clashscore	112137	1807 (2.90-2.90)
Ramachandran outliers	110173	1768 (2.90-2.90)
Sidechain outliers	110143	1770 (2.90-2.90)
RSRZ outliers	101464	1596 (2.90-2.90)

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	X	608	

2 Entry composition [i](#)

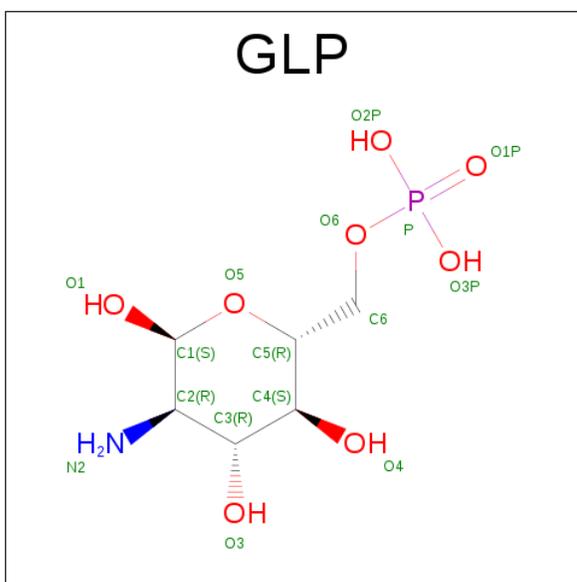
There are 3 unique types of molecules in this entry. The entry contains 2853 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 GLUCOSAMINE--FRUCTOSE-6-PHOSPHATE AMINO-TRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	X	366	2825	1786	483	544	12	0	2	0

- Molecule 2 is SUGAR (GLUCOSAMINE 6-PHOSPHATE) (three-letter code: GLP) (formula: C₆H₁₄NO₈P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	X	1	16	6	1	8	1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	X	12	Total 12 O 12	0	0

4 Data and refinement statistics i

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	144.76Å 144.76Å 171.45Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	15.00 – 2.90 14.95 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (15.00-2.90) 99.8 (14.95-2.90)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.42 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.217 , 0.242 0.180 , 0.218	Depositor DCC
R_{free} test set	767 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	74.8	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 43.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.025 for $-2/3^*h-1/3^*k+2/3^*l,-1/3^*h-2/3^*k-2/3^*l,2/3^*h-2/3^*k+1/3^*l$ 0.012 for $-h,1/3^*h-1/3^*k+2/3^*l,2/3^*h+4/3^*k+1/3^*l$ 0.002 for $-1/3^*h+1/3^*k-2/3^*l,-k,-4/3^*h-2/3^*k+1/3^*l$	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	2853	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% 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: GLP

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	X	0.95	3/2881 (0.1%)	1.08	10/3899 (0.3%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	X	587	VAL	CB-CG1	-7.13	1.37	1.52
1	X	312	TYR	CE1-CZ	6.66	1.47	1.38
1	X	607	VAL	CB-CG2	5.34	1.64	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	X	518	VAL	CB-CA-C	-6.92	98.26	111.40
1	X	288	LEU	CA-CB-CG	6.71	130.72	115.30
1	X	463	ASP	CB-CG-OD2	6.59	124.23	118.30
1	X	511	ASP	CB-CG-OD2	6.36	124.03	118.30
1	X	286	ASP	CB-CG-OD2	6.15	123.83	118.30

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	X	2825	0	2853	144	0
2	X	16	0	11	0	0
3	X	12	0	0	0	0
All	All	2853	0	2864	144	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 25.

The worst 5 of 144 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:X:577:THR:HG22	1:X:581:GLN:HE21	1.08	1.07
1:X:351:GLU:OE2	1:X:379:SER:HB2	1.60	1.00
1:X:443:ARG:HD2	1:X:567:VAL:HG12	1.43	1.00
1:X:466:HIS:HB3	1:X:514:MET:HG3	1.48	0.95
1:X:577:THR:HG22	1:X:581:GLN:NE2	1.88	0.88

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	X	366/608 (60%)	308 (84%)	49 (13%)	9 (2%)	6 25

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	284	ASN
1	X	380	SER
1	X	337	VAL
1	X	390	MET
1	X	422	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	X	307/500 (61%)	254 (83%)	53 (17%)	2 7

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	376	VAL
1	X	419	LEU
1	X	554	VAL
1	X	379	SER
1	X	391	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	X	250	HIS
1	X	259	GLN
1	X	348	GLN
1	X	446	GLN
1	X	581	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 carbohydrates in this entry.

5.6 Ligand geometry

1 ligand is modelled in this entry.

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	GLP	X	609	-	16,16,16	1.49	3 (18%)	22,24,24	3.21	10 (45%)

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	GLP	X	609	-	-	0/6/26/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	X	609	GLP	C1-C2	-2.87	1.49	1.52
2	X	609	GLP	C2-N2	-2.73	1.43	1.47
2	X	609	GLP	P-O1P	2.59	1.59	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	X	609	GLP	O2P-P-O6	-2.20	100.87	106.73
2	X	609	GLP	O3P-P-O2P	2.04	115.84	107.61
2	X	609	GLP	C3-C4-C5	2.44	114.51	110.22
2	X	609	GLP	C3-C2-N2	2.86	116.80	111.00
2	X	609	GLP	O4-C4-C3	3.29	117.51	110.36

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	X	366/608 (60%)	-0.55	7 (1%) 67 64	18, 27, 33, 44	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	X	282	GLY	4.0
1	X	269	GLY	3.6
1	X	243	GLY	3.5
1	X	283	PRO	2.7
1	X	608	GLU	2.7

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(Å ²)	Q<0.9
2	GLP	X	609	16/16	0.98	0.08	-1.39	17,26,28,30	0

6.5 Other polymers [i](#)

There are no such residues in this entry.