wwPDB X-ray Structure Validation Summary Report (i)

Dec 18, 2023 – 08:51 AM EST

PDB ID	:	1O6C
Title	:	Crystal structure of UDP-N-acetylglucosamine 2-epimerase
Authors	:	Structural GenomiX
Deposited on	:	2003-11-03
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
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:

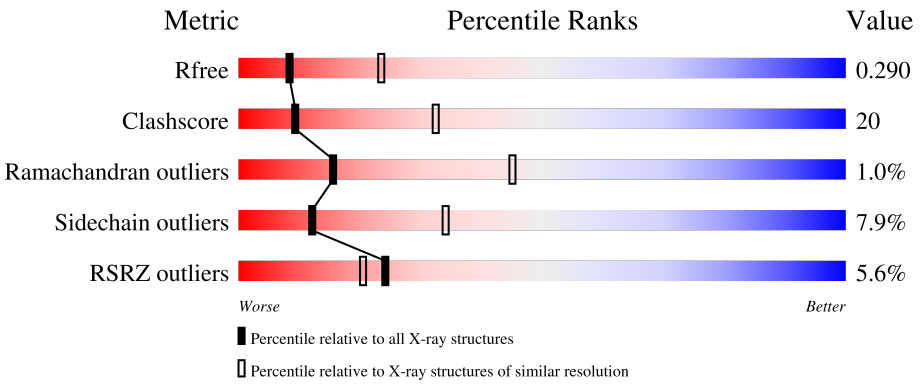
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 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}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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 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	388	9%, 56%, 31%, 8%
1	B	388	2%, 58%, 34%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5507 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 UDP-N-acetylglucosamine 2-epimerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	356	2635	1672	458	499	6	0	0	0
1	B	371	2860	1819	490	541	10	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	cloning artifact	UNP P39131
A	7	MSE	MET	modified residue	UNP P39131
A	19	MSE	MET	modified residue	UNP P39131
A	45	MSE	MET	modified residue	UNP P39131
A	64	MSE	MET	modified residue	UNP P39131
A	140	MSE	MET	modified residue	UNP P39131
A	201	MSE	MET	modified residue	UNP P39131
A	216	MSE	MET	modified residue	UNP P39131
A	219	MSE	MET	modified residue	UNP P39131
A	343	MSE	MET	modified residue	UNP P39131
A	381	GLY	-	cloning artifact	UNP P39131
A	382	SER	-	cloning artifact	UNP P39131
A	383	HIS	-	cloning artifact	UNP P39131
A	384	HIS	-	cloning artifact	UNP P39131
A	385	HIS	-	cloning artifact	UNP P39131
A	386	HIS	-	cloning artifact	UNP P39131
A	387	HIS	-	cloning artifact	UNP P39131
A	388	HIS	-	cloning artifact	UNP P39131
B	1	MSE	-	cloning artifact	UNP P39131
B	7	MSE	MET	modified residue	UNP P39131
B	19	MSE	MET	modified residue	UNP P39131
B	45	MSE	MET	modified residue	UNP P39131
B	64	MSE	MET	modified residue	UNP P39131
B	140	MSE	MET	modified residue	UNP P39131
B	201	MSE	MET	modified residue	UNP P39131

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	216	MSE	MET	modified residue	UNP P39131
B	219	MSE	MET	modified residue	UNP P39131
B	343	MSE	MET	modified residue	UNP P39131
B	381	GLY	-	cloning artifact	UNP P39131
B	382	SER	-	cloning artifact	UNP P39131
B	383	HIS	-	cloning artifact	UNP P39131
B	384	HIS	-	cloning artifact	UNP P39131
B	385	HIS	-	cloning artifact	UNP P39131
B	386	HIS	-	cloning artifact	UNP P39131
B	387	HIS	-	cloning artifact	UNP P39131
B	388	HIS	-	cloning artifact	UNP P39131

- Molecule 2 is water.

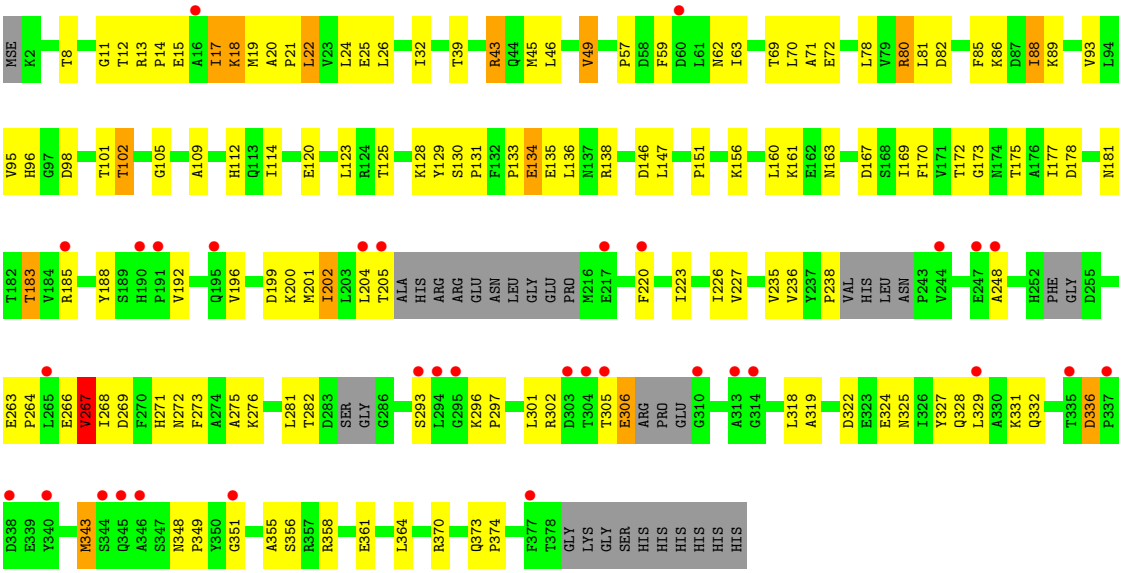
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	3	3	3	0	0
			Total	O		
2	B	9	9	9	0	0

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.

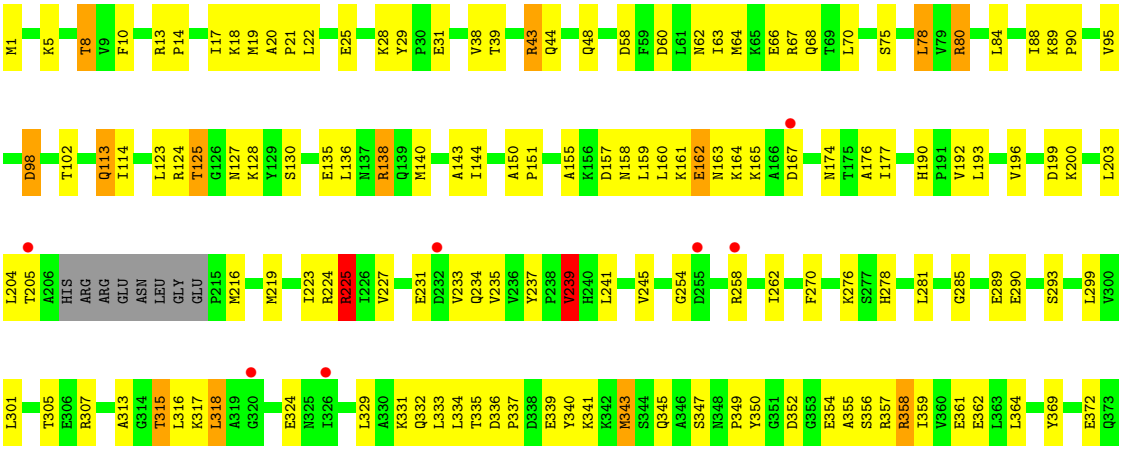
• Molecule 1: UDP-N-acetylglucosamine 2-epimerase

Chain A:



• Molecule 1: UDP-N-acetylglucosamine 2-epimerase

Chain B:



P374
D375
S376
F377
T378
G379
LYS
GLY
SER
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	63.82Å 63.82Å 452.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.71 – 2.90 48.71 – 2.90	Depositor EDS
% Data completeness (in resolution range)	(Not available) (48.71-2.90) 99.3 (48.71-2.90)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.90 (at 2.91Å)	Xtriage
Refinement program	REFMAC 4.0	Depositor
R, R_{free}	0.274 , 0.334 0.236 , 0.290	Depositor DCC
R_{free} test set	2141 reflections (9.76%)	wwPDB-VP
Wilson B-factor (Å^2)	65.1	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 72.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	5507	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% 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)

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	0.54	0/2673	0.94	2/3606 (0.1%)
1	B	0.65	0/2908	1.11	11/3927 (0.3%)
All	All	0.60	0/5581	1.03	13/7533 (0.2%)

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 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	80	ARG	NE-CZ-NH1	-9.59	115.50	120.30
1	A	370	ARG	NE-CZ-NH1	-8.74	115.93	120.30
1	B	138	ARG	NE-CZ-NH2	-7.34	116.63	120.30
1	B	225	ARG	NE-CZ-NH1	6.64	123.62	120.30
1	B	124	ARG	CD-NE-CZ	-5.90	115.33	123.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	58	ASP	Mainchain

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	2635	0	2488	105	0
1	B	2860	0	2807	117	0
2	A	3	0	0	0	0
2	B	9	0	0	2	0
All	All	5507	0	5295	214	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 20.

The worst 5 of 214 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:305:THR:HB	1:B:318:LEU:HD11	1.34	1.04
1:A:43:ARG:HH21	1:A:57:PRO:HG2	1.26	0.98
1:B:333:LEU:HD21	1:B:343:MSE:HE3	1.47	0.95
1:B:39:THR:HB	1:B:63:ILE:HD12	1.51	0.93
1:B:204:LEU:HD11	1:B:219:MSE:SE	2.22	0.90

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	344/388 (89%)	311 (90%)	27 (8%)	6 (2%)	9 31
1	B	367/388 (95%)	346 (94%)	20 (5%)	1 (0%)	41 71

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	711/776 (92%)	657 (92%)	47 (7%)	7 (1%)	15 45

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	18	LYS
1	A	267	VAL
1	B	254	GLY
1	A	11	GLY
1	A	167	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	256/323 (79%)	235 (92%)	21 (8%)	11 32
1	B	300/323 (93%)	277 (92%)	23 (8%)	13 35
All	All	556/646 (86%)	512 (92%)	44 (8%)	12 34

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	67	ARG
1	B	225	ARG
1	B	75	SER
1	B	113	GLN
1	B	293	SER

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

Mol	Chain	Res	Type
1	A	190	HIS
1	A	288	GLN
1	A	373	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	345	GLN
1	B	366	HIS

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)

There are no ligands in this entry.

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	347/388 (89%)	0.75	33 (9%)	8	6	23, 62, 107, 138	0
1	B	361/388 (93%)	0.34	7 (1%)	66	65	16, 36, 66, 103	0
All	All	708/776 (91%)	0.54	40 (5%)	24	20	16, 45, 98, 138	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	344	SER	7.2
1	A	340	TYR	4.9
1	B	255	ASP	4.7
1	B	232	ASP	4.4
1	A	244	VAL	4.2

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.