wwPDB X-ray Structure Validation Summary Report (i)

Jan 30, 2024 – 01:23 PM EST

PDB ID	:	1HOZ
Title	:	CRYSTAL STRUCTURE OF AN INOSINE-ADENOSINE-GUANOSINE-PREFERRING NUCLEOSIDE HYDROLASE FROM TRYPANOSOMA VIVAX
Authors	:	Versees, W.; Decanniere, K.; Pelle, R.; Depoorter, J.; Parkin, D.W.; Steyaert, J.
Deposited on	:	2000-12-12
Resolution	:	1.60 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:

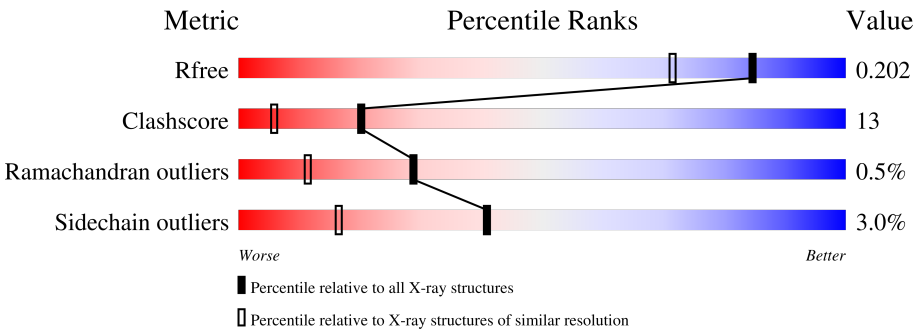
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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%

Mol	Chain	Length	Quality of chain
1	A	339	79% 13% • 7%
1	B	339	78% 13% •• 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5465 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 INOSINE-ADENOSINE-GUANOSINE-PREFERRING NUCLEOSIDE HYDROLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	316	2498	1595	403	475	25	0	20	0
1	B	315	2493	1592	405	471	25	0	18	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-12	MET	-	expression tag	UNP Q9GPQ4
A	-11	ARG	-	expression tag	UNP Q9GPQ4
A	-10	GLY	-	expression tag	UNP Q9GPQ4
A	-9	SER	-	expression tag	UNP Q9GPQ4
A	-8	PRO	-	expression tag	UNP Q9GPQ4
A	-7	HIS	-	expression tag	UNP Q9GPQ4
A	-6	HIS	-	expression tag	UNP Q9GPQ4
A	-5	HIS	-	expression tag	UNP Q9GPQ4
A	-4	HIS	-	expression tag	UNP Q9GPQ4
A	-3	HIS	-	expression tag	UNP Q9GPQ4
A	-2	HIS	-	expression tag	UNP Q9GPQ4
A	-1	GLY	-	expression tag	UNP Q9GPQ4
A	0	SER	-	expression tag	UNP Q9GPQ4
A	301	ASN	LYS	SEE REMARK 999	UNP Q9GPQ4
B	-12	MET	-	expression tag	UNP Q9GPQ4
B	-11	ARG	-	expression tag	UNP Q9GPQ4
B	-10	GLY	-	expression tag	UNP Q9GPQ4
B	-9	SER	-	expression tag	UNP Q9GPQ4
B	-8	PRO	-	expression tag	UNP Q9GPQ4
B	-7	HIS	-	expression tag	UNP Q9GPQ4
B	-6	HIS	-	expression tag	UNP Q9GPQ4
B	-5	HIS	-	expression tag	UNP Q9GPQ4
B	-4	HIS	-	expression tag	UNP Q9GPQ4
B	-3	HIS	-	expression tag	UNP Q9GPQ4

Continued on next page...

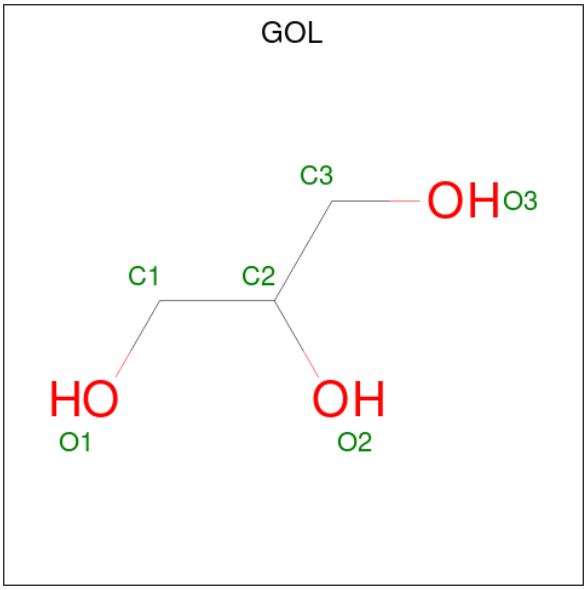
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	HIS	-	expression tag	UNP Q9GPQ4
B	-1	GLY	-	expression tag	UNP Q9GPQ4
B	0	SER	-	expression tag	UNP Q9GPQ4
B	301	ASN	LYS	SEE REMARK 999	UNP Q9GPQ4

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ca 1 1	0	0
2	B	1	Total Ca 1 1	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 4 is water.

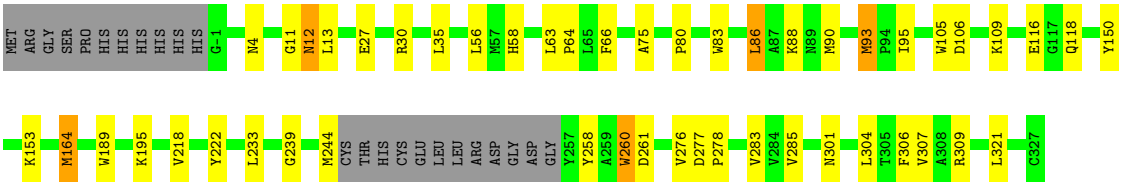
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	214	214	214	0	0
			Total	O		
4	B	210	210	210	0	0

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. 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. 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.

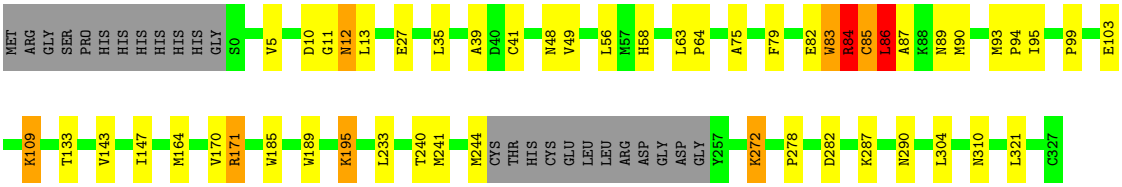
• Molecule 1: INOSINE-ADENOSINE-GUANOSINE-PREFERRING NUCLEOSIDE HYDROLASE

Chain A:



• Molecule 1: INOSINE-ADENOSINE-GUANOSINE-PREFERRING NUCLEOSIDE HYDROLASE

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	52.09Å 74.00Å 82.00Å 90.00° 104.56° 90.00°	Depositor
Resolution (Å)	30.00 – 1.60 29.83 – 1.60	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-1.60) 99.8 (29.83-1.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.44 (at 1.60Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.186 , 0.210 0.180 , 0.202	Depositor DCC
R_{free} test set	3989 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	33.4	Xtriage
Anisotropy	0.293	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 47.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5465	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.67% 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: CA, GOL

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	0.48	0/2635	0.73	1/3590 (0.0%)
1	B	0.53	1/2620 (0.0%)	0.76	3/3568 (0.1%)
All	All	0.50	1/5255 (0.0%)	0.75	4/7158 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	85	CYS	CB-SG	-13.13	1.59	1.82

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	84	ARG	NE-CZ-NH2	-10.55	115.03	120.30
1	B	84	ARG	NE-CZ-NH1	6.41	123.51	120.30
1	B	84	ARG	N-CA-C	6.32	128.05	111.00
1	A	11	GLY	N-CA-C	5.23	126.17	113.10

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	2498	0	2441	59	3

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2493	0	2440	70	2
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	36	0	48	3	0
3	B	12	0	16	0	0
4	A	214	0	0	2	3
4	B	210	0	0	10	4
All	All	5465	0	4945	126	6

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 13.

The worst 5 of 126 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:12[A]:ASN:ND2	1:B:86:LEU:HD13	1.68	1.07
1:A:13:LEU:HD21	1:A:86:LEU:HG	1.30	1.07
1:B:12[A]:ASN:HD21	1:B:86:LEU:HD13	1.13	1.06
1:A:276[A]:VAL:HG11	1:A:306:PHE:HB3	1.42	1.00
1:A:93[A]:MET:HE2	1:A:95:ILE:H	1.29	0.98

The worst 5 of 6 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:301:ASN:OD1	4:B:1375:HOH:O[1_454]	1.55	0.65
4:A:1482:HOH:O	4:B:1478:HOH:O[1_454]	1.96	0.24
1:B:195[A]:LYS:NZ	4:A:1393:HOH:O[1_656]	1.97	0.23
1:A:106:ASP:OD2	1:B:290:ASN:ND2[2_656]	2.02	0.18
1:A:105:TRP:CD1	4:B:1364:HOH:O[2_656]	2.12	0.08

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	332/339 (98%)	325 (98%)	6 (2%)	1 (0%)	41 21
1	B	329/339 (97%)	316 (96%)	11 (3%)	2 (1%)	25 8
All	All	661/678 (98%)	641 (97%)	17 (3%)	3 (0%)	29 11

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	84	ARG
1	B	86	LEU
1	A	260	TRP

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	283/290 (98%)	274 (97%)	9 (3%)	39 15
1	B	280/290 (97%)	269 (96%)	11 (4%)	32 10
All	All	563/580 (97%)	543 (96%)	20 (4%)	41 12

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	109	LYS
1	B	195[A]	LYS
1	B	272	LYS
1	B	195[B]	LYS
1	A	164[A]	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	58	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	119	GLN
1	B	310	ASN
1	B	224	GLN
1	A	119	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 10 ligands modelled in this entry, 2 are monoatomic - leaving 8 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	GOL	A	1273	-	5,5,5	0.51	0	5,5,5	0.66	0
3	GOL	A	1272	-	5,5,5	0.64	0	5,5,5	0.90	0
3	GOL	B	1270	-	5,5,5	0.50	0	5,5,5	1.05	0
3	GOL	A	1271	-	5,5,5	0.81	0	5,5,5	0.91	0
3	GOL	A	1274	-	5,5,5	0.64	0	5,5,5	0.92	0
3	GOL	A	1267	2	5,5,5	0.54	0	5,5,5	0.76	0
3	GOL	A	1268	-	5,5,5	0.84	0	5,5,5	1.10	1 (20%)
3	GOL	B	1269	-	5,5,5	0.65	0	5,5,5	0.96	0

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	GOL	A	1273	-	-	2/4/4/4	-
3	GOL	A	1272	-	-	0/4/4/4	-
3	GOL	B	1270	-	-	2/4/4/4	-
3	GOL	A	1271	-	-	1/4/4/4	-
3	GOL	A	1274	-	-	1/4/4/4	-
3	GOL	A	1267	2	-	2/4/4/4	-
3	GOL	A	1268	-	-	2/4/4/4	-
3	GOL	B	1269	-	-	1/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1268	GOL	O2-C2-C1	-2.02	100.21	109.12

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1267	GOL	O1-C1-C2-C3
3	B	1270	GOL	O1-C1-C2-C3
3	A	1267	GOL	O1-C1-C2-O2
3	A	1268	GOL	O1-C1-C2-C3
3	A	1273	GOL	O1-C1-C2-C3

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1272	GOL	1	0
3	A	1271	GOL	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)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.