wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 02:23 AM GMT

PDB ID	:	2GYI
Title	:	DESIGN, SYNTHESIS, AND CHARACTERIZATION OF A POTENT XYLOSE ISOMERASE INHIBITOR, D-THREONOHYDROXAMIC ACID, AND HIGH-RESOLUTION X-RAY CRYSTALLOGRAPHIC STRUCTURE OF THE ENZYME-INHIBITOR COMPLEX
Authors	:	Allen, K.N.; Lavie, A.; Petsko, G.A.; Ringe, D.
Deposited on	:	1994-09-01
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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

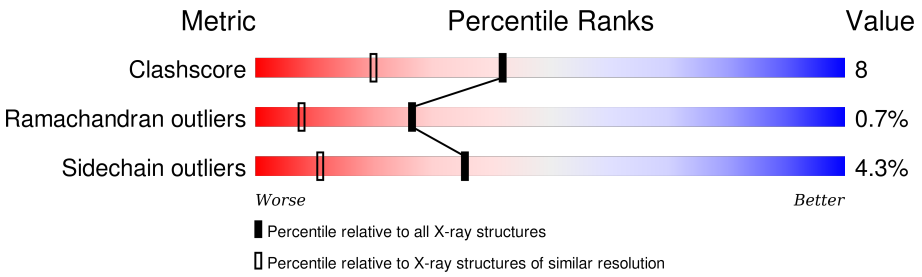
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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))
Clashscore	102246	2732 (1.60-1.60)
Ramachandran outliers	100387	2654 (1.60-1.60)
Sidechain outliers	100360	2653 (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 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 <=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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	386	71% 25% • •
1	B	386	78% 19% • •

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
3	HYA	A	960	X	-	-	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6694 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.

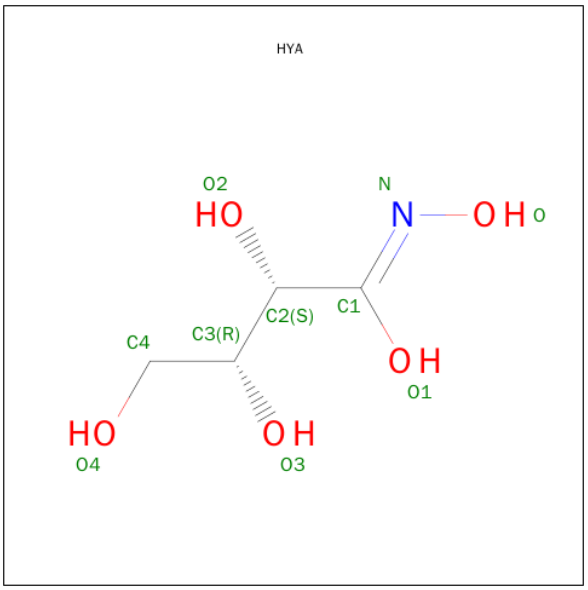
- Molecule 1 is a protein called XYLOSE ISOMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	385	3018	1901	539	570	8	0	0	0
1	B	385	3018	1901	539	570	8	0	0	0

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
2	B	2	2	2	0	0
2	A	2	2	2	0	0

- Molecule 3 is 2,3,4,N-TETRAHYDROXY-BUTYRIMIDIC ACID (three-letter code: HYA) (formula: $C_4H_9NO_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	10	4	1	5	0	0
			Total	C	N	O		
3	B	1	10	4	1	5	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	303	303	303	0	0
			Total	O		
4	B	331	331	331	0	0

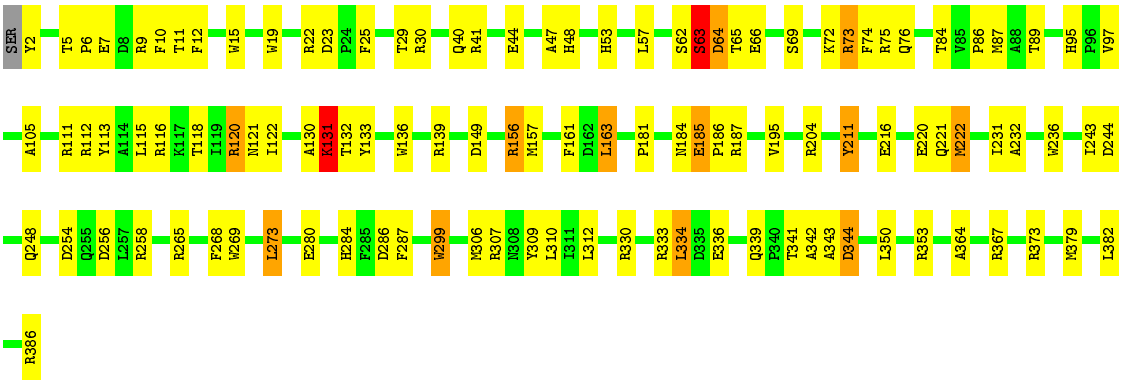
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

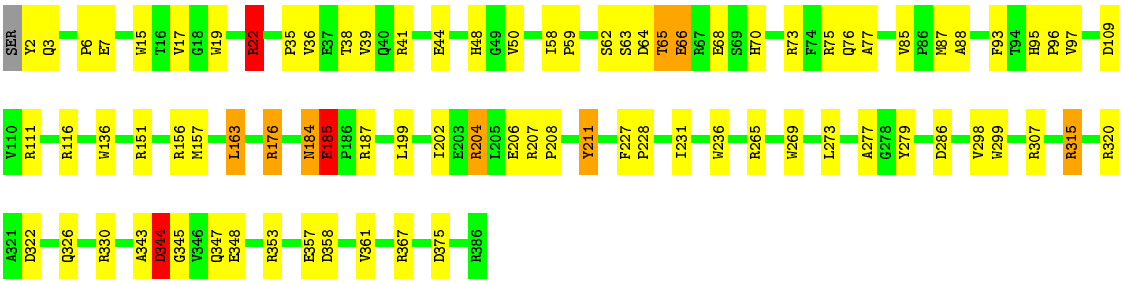
- Molecule 1: XYLOSE ISOMERASE

Chain A:



- Molecule 1: XYLOSE ISOMERASE

Chain B:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	86.90Å 99.05Å 93.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.60	Depositor
% Data completeness (in resolution range)	(Not available) (10.00-1.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.216 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6694	wwPDB-VP
Average B, all atoms ($Å^2$)	17.0	wwPDB-VP

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: HYA, MG

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	1.06	6/3090 (0.2%)	1.61	63/4189 (1.5%)
1	B	0.93	2/3090 (0.1%)	1.57	52/4189 (1.2%)
All	All	1.00	8/6180 (0.1%)	1.59	115/8378 (1.4%)

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	3
1	B	0	2
All	All	0	5

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	64	ASP	C-N	24.36	1.90	1.34
1	A	63	SER	C-N	-19.08	0.90	1.34
1	B	66	GLU	CG-CD	11.42	1.69	1.51
1	B	66	GLU	CB-CG	9.44	1.70	1.52
1	A	149	ASP	C-N	-9.28	1.12	1.34

The worst 5 of 115 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	64	ASP	O-C-N	-23.17	85.63	122.70
1	B	265	ARG	NE-CZ-NH2	-19.68	110.46	120.30
1	B	265	ARG	NE-CZ-NH1	16.80	128.70	120.30
1	A	63	SER	C-N-CA	15.95	161.57	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	112	ARG	NE-CZ-NH1	11.21	125.91	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	131	LYS	Mainchain
1	A	364	ALA	Mainchain
1	A	63	SER	Peptide
1	B	22	ARG	Sidechain
1	B	63	SER	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	3018	0	2898	65	0
1	B	3018	0	2901	38	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	10	0	8	1	0
3	B	10	0	8	0	0
4	A	303	0	0	16	0
4	B	331	0	0	20	0
All	All	6694	0	5815	100	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 8.

The worst 5 of 100 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:5:THR:HG23	1:A:7:GLU:CG	1.61	1.29
1:A:64:ASP:C	1:A:65:THR:N	1.90	1.25
1:A:69:SER:OG	1:A:73:ARG:NH2	1.81	1.13

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:5:THR:CG2	1:A:7:GLU:CG	2.28	1.11
1:A:5:THR:HG23	1:A:7:GLU:HG2	1.32	1.10

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	383/386 (99%)	367 (96%)	13 (3%)	3 (1%)	24 6
1	B	383/386 (99%)	367 (96%)	14 (4%)	2 (0%)	34 12
All	All	766/772 (99%)	734 (96%)	27 (4%)	5 (1%)	26 8

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	344	ASP
1	B	344	ASP
1	A	185	GLU
1	B	185	GLU
1	A	23	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	301/302 (100%)	289 (96%)	12 (4%)	38	12
1	B	301/302 (100%)	287 (95%)	14 (5%)	32	9
All	All	602/604 (100%)	576 (96%)	26 (4%)	35	11

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	350	LEU
1	B	41	ARG
1	B	347	GLN
1	B	3	GLN
1	B	7	GLU

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

Mol	Chain	Res	Type
1	A	339	GLN
1	B	76	GLN
1	B	221	GLN
1	A	326	GLN
1	B	184	ASN

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 (i)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	HYA	A	960	2	8,9,9	1.34	0	5,11,11	2.72	3 (60%)
3	HYA	B	970	2	8,9,9	1.40	1 (12%)	5,11,11	1.45	1 (20%)

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	HYA	A	960	2	1/1/3/3	0/8/12/12	0/0/0/0
3	HYA	B	970	2	-	0/8/12/12	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	970	HYA	C1-N	3.31	1.35	1.27

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	960	HYA	O3-C3-C4	-2.32	103.80	109.22
3	B	970	HYA	C4-C3-C2	2.58	116.24	111.94
3	A	960	HYA	O2-C2-C1	3.61	118.07	110.84
3	A	960	HYA	C4-C3-C2	4.09	118.75	111.94

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	960	HYA	C2

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	960	HYA	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)

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.