



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 23, 2024 – 01:31 AM EDT

PDB ID : 6HA3  
Title : Human transketolase variant E160Q in covalent complex with donor ketose D-fructose-6-phosphate  
Authors : Dai, S.; Sautner, V.; Tittmann, K.  
Deposited on : 2018-08-07  
Resolution : 1.08 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](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	FAILED
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	FAILED
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)	:	2.37.1

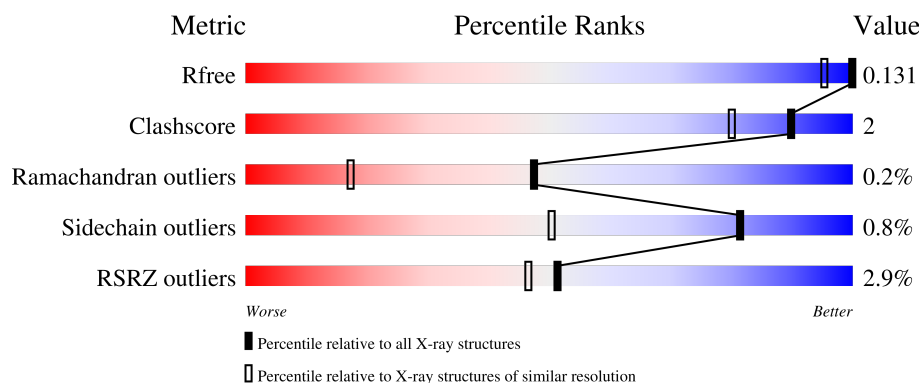
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

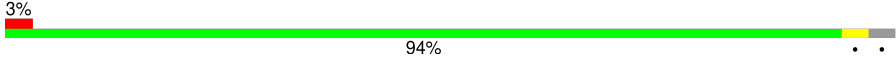
The reported resolution of this entry is 1.08 Å.

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	1386 (1.12-1.04)
Clashscore	141614	1021 (1.10-1.06)
Ramachandran outliers	138981	1381 (1.12-1.04)
Sidechain outliers	138945	1379 (1.12-1.04)
RSRZ outliers	127900	1359 (1.12-1.04)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	637	

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	EDO	A	714	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 11405 atoms, of which 5353 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.

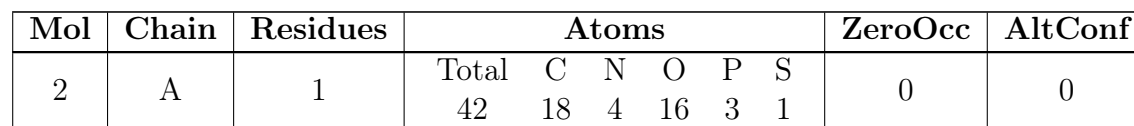
- Molecule 1 is a protein called Transketolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	620	Total	C	H	N	O	S	0	99	0
			10568	3313	5299	926	1001	29			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	160	GLN	GLU	engineered mutation	UNP P29401
A	624	LEU	-	expression tag	UNP P29401
A	625	VAL	-	expression tag	UNP P29401
A	626	PRO	-	expression tag	UNP P29401
A	627	ARG	-	expression tag	UNP P29401
A	628	GLY	-	expression tag	UNP P29401
A	629	SER	-	expression tag	UNP P29401
A	630	LEU	-	expression tag	UNP P29401
A	631	GLU	-	expression tag	UNP P29401
A	632	HIS	-	expression tag	UNP P29401
A	633	HIS	-	expression tag	UNP P29401
A	634	HIS	-	expression tag	UNP P29401
A	635	HIS	-	expression tag	UNP P29401
A	636	HIS	-	expression tag	UNP P29401
A	637	HIS	-	expression tag	UNP P29401

- Molecule 2 is 2-C-{3-[(4-amino-2-methylpyrimidin-5-yl)methyl]-5-(2-[(R)-hydroxy(phosphonooxy)phosphoryl]oxy)ethyl)-4-methyl-1,3-thiazol-3-ium-2-yl}-6-O-phosphono-D-glucitol (three-letter code: T6F) (formula: C<sub>18</sub>H<sub>32</sub>N<sub>4</sub>O<sub>16</sub>P<sub>3</sub>S) (labeled as "Ligand of Interest" by depositor).



- 
- EDO
- Chemical structure of EDO (Ethane-1,2-diol) showing a zigzag conformation. The carbon atoms are labeled C1 and C2 in green. The hydroxyl groups are labeled HO and OH in red, with their respective oxygen atoms labeled O1 and O2 in green.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H O 10 2 6 2	0	0
3	A	1	Total C O 8 4 4	0	1
3	A	1	Total C H O 10 2 6 2	0	0



*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H O 10 2 6 2	0	0
3	A	1	Total C H O 10 2 6 2	0	1
3	A	1	Total C H O 10 2 6 2	0	0
3	A	1	Total C H O 20 4 12 4	0	1
3	A	1	Total C H O 10 2 6 2	0	0
3	A	1	Total C H O 10 2 6 2	0	1
3	A	1	Total C O 4 2 2	0	1
3	A	1	Total C O 4 2 2	0	1
3	A	1	Total C O 4 2 2	0	0

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Na 1 1	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Ca 1 1	0	1

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Mg 1 1	0	1

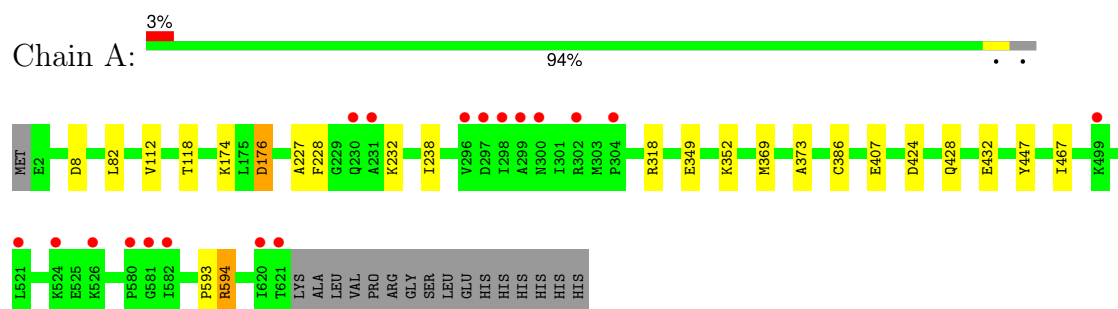
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	663	Total O 682 682	0	41

### 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: Transketolase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.07Å 86.09Å 73.21Å 90.00° 125.22° 90.00°	Depositor
Resolution (Å)	24.56 – 1.08 24.56 – 1.08	Depositor EDS
% Data completeness (in resolution range)	96.7 (24.56-1.08) 96.7 (24.56-1.08)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.68 (at 1.08Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.111 , 0.130 0.112 , 0.131	Depositor DCC
$R_{free}$ test set	11913 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	11.4	Xtriage
Anisotropy	0.060	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 56.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.99	EDS
Total number of atoms	11405	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.99% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: EDO, MG, CA, NA, T6F

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	0.42	0/5638	0.67	4/7622 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	594[A]	ARG	NE-CZ-NH2	5.35	122.97	120.30
1	A	594[B]	ARG	NE-CZ-NH2	5.35	122.97	120.30
1	A	176[A]	ASP	CB-CG-OD1	5.22	123.00	118.30
1	A	176[B]	ASP	CB-CG-OD1	5.22	123.00	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	A	5269	5299	5002	16	0
2	A	42	0	26	1	0
3	A	56	54	83	6	0
4	A	1	0	0	0	0
5	A	1	0	0	0	0
6	A	1	0	0	0	0
7	A	682	0	0	6	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	6052	5353	5111	19	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 2.

The worst 5 of 19 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:352:LYS:HE3	7:A:1292:HOH:O	1.61	1.01
3:A:714:EDO:H21	7:A:1284:HOH:O	1.63	0.95
3:A:714:EDO:C2	7:A:1284:HOH:O	2.25	0.78
1:A:407:GLU:OE2	7:A:1103[B]:HOH:O	2.15	0.65
1:A:82:LEU:HD21	3:A:703[B]:EDO:H12	1.82	0.62

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	719/637 (113%)	699 (97%)	19 (3%)	1 (0%)	51 18

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	593	PRO

### 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	590/518 (114%)	585 (99%)	5 (1%)	81	50

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	112	VAL
1	A	432	GLU
1	A	447	TYR
1	A	594[A]	ARG
1	A	594[B]	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

Mogul failed to run properly - this section is therefore empty.

## 5.5 Carbohydrates [i](#)

Mogul failed to run properly - this section is therefore empty.

## 5.6 Ligand geometry [i](#)

Mogul failed to run properly - this section is therefore empty.

## 5.7 Other polymers [i](#)

Mogul failed to run properly - this section is therefore empty.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	620/637 (97%)	-0.38	18 (2%) 51 46	8, 12, 27, 42	10 (1%)

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	298	ILE	7.0
1	A	231[A]	ALA	4.3
1	A	526	LYS	4.0
1	A	300[A]	ASN	3.7
1	A	580	PRO	3.6

### 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](#)

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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	B-factors(Å <sup>2</sup> )	Q<0.9
3	EDO	A	710	4/4	0.74	0.22	58,69,70,71	0
3	EDO	A	705	4/4	0.87	0.19	22,26,28,28	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	EDO	A	714	4/4	0.88	0.11	12,18,21,22	4
3	EDO	A	713[A]	4/4	0.90	0.19	14,15,16,19	4
3	EDO	A	711[A]	4/4	0.91	0.11	29,35,39,39	10
3	EDO	A	707	4/4	0.92	0.12	21,26,30,31	10
3	EDO	A	709[A]	4/4	0.94	0.08	28,34,39,41	10
3	EDO	A	709[B]	4/4	0.94	0.08	31,37,40,41	10
3	EDO	A	702	4/4	0.94	0.15	11,21,26,26	10
3	EDO	A	706[A]	4/4	0.95	0.08	12,18,22,22	0
3	EDO	A	704	4/4	0.96	0.08	11,16,19,19	10
3	EDO	A	703[A]	4/4	0.98	0.10	8,10,12,12	4
3	EDO	A	703[B]	4/4	0.98	0.10	13,14,18,19	4
3	EDO	A	712[B]	4/4	0.99	0.10	21,22,22,23	4
2	T6F	A	701	42/42	1.00	0.05	8,10,13,14	17
4	NA	A	708	1/1	1.00	0.02	11,11,11,11	0
5	CA	A	715[A]	1/1	1.00	0.04	10,10,10,10	1
6	MG	A	716[B]	1/1	1.00	0.04	6,6,6,6	1

## 6.5 Other polymers

There are no such residues in this entry.