



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 21, 2024 – 06:03 PM EDT

PDB ID : 3C3J  
Title : Crystal structure of tagatose-6-phosphate ketose/aldose isomerase from Escherichia coli  
Authors : Zhang, R.; Skarina, T.; Egorova, O.; Savchenko, A.; Edwards, A.M.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2008-01-28  
Resolution : 1.80 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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

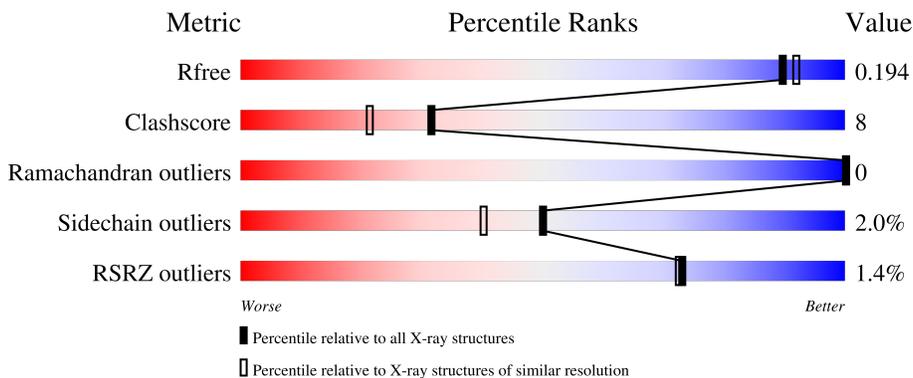
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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))
$R_{free}$	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

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	384	 2% 87% 9% . .
1	B	384	 % 87% 9% .
1	C	384	 % 83% 12% .
1	D	384	 % 83% 11% . .
1	E	384	 3% 80% 14% . .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	384	 <p>A horizontal bar chart representing the quality of the chain. The bar is divided into three segments: a large green segment (87%), a smaller yellow segment (8%), and a very small grey segment (5%). A small red square is at the beginning of the bar. The percentage values are printed below the bar.</p>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 19285 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.

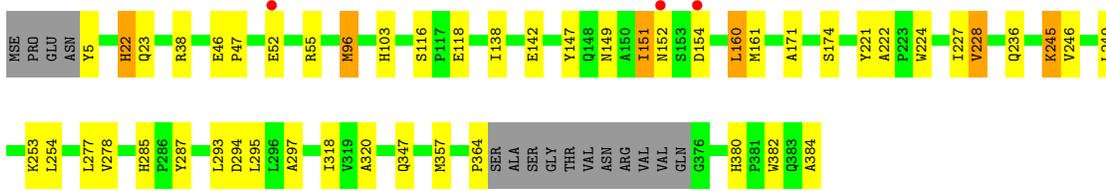
- Molecule 1 is a protein called Putative tagatose-6-phosphate ketose/aldose isomerase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	370	Total 2840	C 1807	N 486	O 533	S 5	Se 9	0	0	0
1	B	369	Total 2832	C 1803	N 484	O 531	S 5	Se 9	0	0	0
1	C	367	Total 2820	C 1795	N 482	O 529	S 5	Se 9	0	0	0
1	D	369	Total 2832	C 1803	N 484	O 531	S 5	Se 9	0	0	0
1	E	369	Total 2835	C 1804	N 485	O 532	S 5	Se 9	0	0	0
1	F	369	Total 2832	C 1803	N 484	O 531	S 5	Se 9	0	0	0

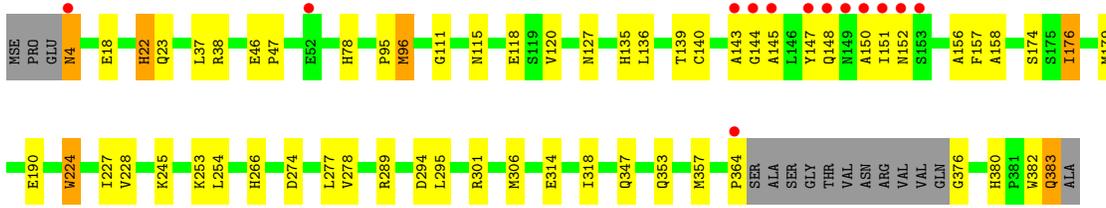
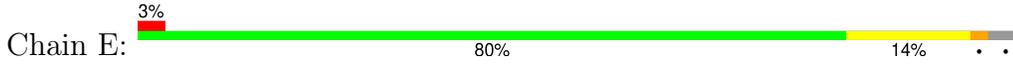
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	382	Total 382	O 382	0	0
2	B	386	Total 386	O 386	0	0
2	C	375	Total 375	O 375	0	0
2	D	375	Total 375	O 375	0	0
2	E	391	Total 391	O 391	0	0
2	F	385	Total 385	O 385	0	0

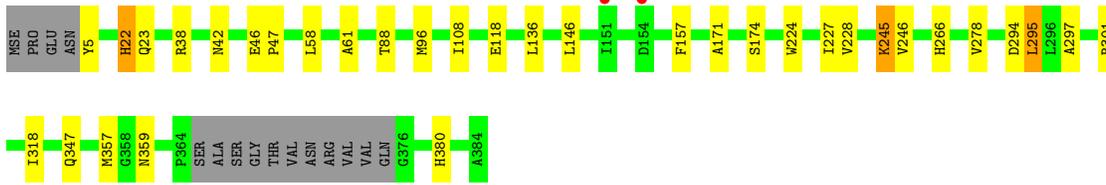
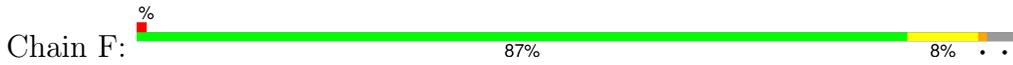




• Molecule 1: Putative tagatose-6-phosphate ketose/aldose isomerase



• Molecule 1: Putative tagatose-6-phosphate ketose/aldose isomerase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.94Å 81.21Å 138.32Å 90.00° 90.21° 90.00°	Depositor
Resolution (Å)	47.46 – 1.80 47.46 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (47.46-1.80) 99.5 (47.46-1.80)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.96 (at 1.81Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.136 , 0.185 0.146 , 0.194	Depositor DCC
$R_{free}$ test set	10357 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.8	Xtrriage
Anisotropy	0.257	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 60.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.018 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	19285	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 49.75 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.1382e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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](#)

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	0.65	0/2901	0.67	1/3945 (0.0%)
1	B	0.63	0/2893	0.70	2/3934 (0.1%)
1	C	0.64	0/2880	0.69	0/3915
1	D	0.64	1/2893 (0.0%)	0.69	1/3934 (0.0%)
1	E	0.63	0/2896	0.72	3/3938 (0.1%)
1	F	0.61	0/2893	0.68	4/3934 (0.1%)
All	All	0.63	1/17356 (0.0%)	0.69	11/23600 (0.0%)

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	E	1	0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	228	VAL	CB-CG1	-5.34	1.41	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	38	ARG	NE-CZ-NH1	8.03	124.31	120.30
1	E	38	ARG	NE-CZ-NH1	7.45	124.02	120.30
1	F	295	LEU	CA-CB-CG	-7.02	99.15	115.30
1	B	38	ARG	NE-CZ-NH2	-6.91	116.85	120.30
1	E	38	ARG	NE-CZ-NH2	-6.53	117.03	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	E	383	GLN	CA

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	2840	0	2784	40	0
1	B	2832	0	2778	36	0
1	C	2820	0	2766	55	0
1	D	2832	0	2778	53	0
1	E	2835	0	2779	78	0
1	F	2832	0	2778	28	0
2	A	382	0	0	9	1
2	B	386	0	0	10	2
2	C	375	0	0	17	1
2	D	375	0	0	20	1
2	E	391	0	0	26	0
2	F	385	0	0	6	1
All	All	19285	0	16663	269	3

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 269 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:E:96:MSE:HE2	1:E:382:TRP:CH2	1.67	1.28
1:A:352:MSE:CE	1:A:355:LEU:HD12	1.78	1.12
1:A:351:LEU:HG	1:A:352:MSE:HE3	1.12	1.09
1:A:352:MSE:HE1	1:A:355:LEU:HD12	1.09	1.07
1:B:351:LEU:HD23	1:B:352:MSE:HE2	1.33	1.04

All (3) 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)
2:B:750:HOH:O	2:D:636:HOH:O[1_545]	2.11	0.09
2:A:545:HOH:O	2:C:746:HOH:O[2_646]	2.12	0.08
2:B:764:HOH:O	2:F:520:HOH:O[1_655]	2.19	0.01

### 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	366/384 (95%)	358 (98%)	8 (2%)	0	100	100
1	B	365/384 (95%)	358 (98%)	7 (2%)	0	100	100
1	C	363/384 (94%)	358 (99%)	5 (1%)	0	100	100
1	D	365/384 (95%)	360 (99%)	5 (1%)	0	100	100
1	E	365/384 (95%)	360 (99%)	5 (1%)	0	100	100
1	F	365/384 (95%)	357 (98%)	8 (2%)	0	100	100
All	All	2189/2304 (95%)	2151 (98%)	38 (2%)	0	100	100

There are no Ramachandran outliers to report.

#### 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	302/304 (99%)	297 (98%)	5 (2%)	56	47
1	B	301/304 (99%)	296 (98%)	5 (2%)	56	47
1	C	300/304 (99%)	296 (99%)	4 (1%)	65	59

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	301/304 (99%)	292 (97%)	9 (3%)	36	24
1	E	302/304 (99%)	294 (97%)	8 (3%)	41	29
1	F	301/304 (99%)	296 (98%)	5 (2%)	56	47
All	All	1807/1824 (99%)	1771 (98%)	36 (2%)	50	40

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	224	TRP
1	F	245	LYS
1	E	245	LYS
1	F	96	MSE
1	C	224	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	103	HIS
1	E	22	HIS
1	F	285	HIS
1	D	149	ASN
1	D	285	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 oligosaccharides 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, 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	361/384 (94%)	-0.46	6 (1%) 69 67	12, 18, 28, 39	0
1	B	360/384 (93%)	-0.52	4 (1%) 77 77	11, 17, 28, 40	0
1	C	358/384 (93%)	-0.50	3 (0%) 82 82	11, 16, 27, 42	0
1	D	360/384 (93%)	-0.50	3 (0%) 82 82	11, 17, 29, 39	0
1	E	360/384 (93%)	-0.41	13 (3%) 46 44	11, 17, 29, 50	0
1	F	360/384 (93%)	-0.44	2 (0%) 85 85	12, 18, 29, 40	0
All	All	2159/2304 (93%)	-0.47	31 (1%) 73 72	11, 17, 28, 50	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	151	ILE	3.8
1	E	147	TYR	3.5
1	A	4	ASN	3.4
1	D	152	ASN	3.3
1	E	143	ALA	3.3

### 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.