



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 07:24 pm BST

PDB ID : 1J5S  
Title : Crystal structure of uronate isomerase (TM0064) from *Thermotoga maritima*  
at 2.85 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2002-07-02  
Resolution : 2.85 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

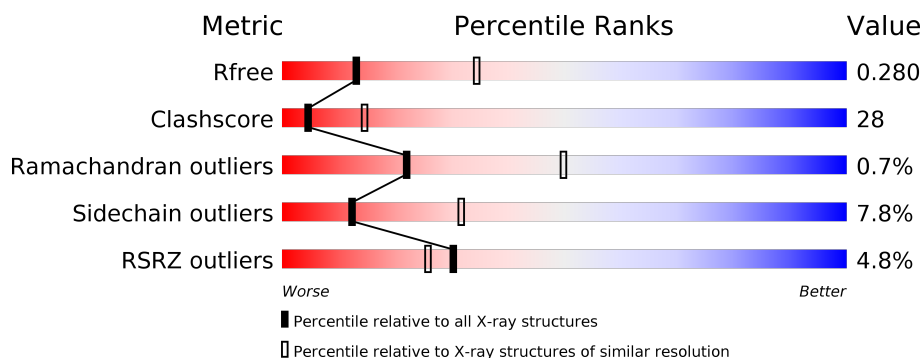
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.85 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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  $\geq 3$ , 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	463	<div> <div>4%</div> <div> <div></div> <div>52%</div> <div>40%</div> <div>5%</div> </div> </div>
1	B	463	<div> <div>5%</div> <div> <div></div> <div>52%</div> <div>40%</div> <div>5%</div> </div> </div>
1	C	463	<div> <div>5%</div> <div> <div></div> <div>55%</div> <div>37%</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11175 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 URONATE ISOMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	451	Total	C	N	O	S	81	0	0
			3696	2367	632	680	17			
1	B	450	Total	C	N	O	S	86	0	0
			3686	2361	629	679	17			
1	C	450	Total	C	N	O	S	83	0	0
			3686	2361	629	679	17			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	LEADER SEQUENCE	UNP Q9WXR9
A	-10	GLY	-	LEADER SEQUENCE	UNP Q9WXR9
A	-9	SER	-	LEADER SEQUENCE	UNP Q9WXR9
A	-8	ASP	-	LEADER SEQUENCE	UNP Q9WXR9
A	-7	LYS	-	LEADER SEQUENCE	UNP Q9WXR9
A	-6	ILE	-	LEADER SEQUENCE	UNP Q9WXR9
A	-5	HIS	-	EXPRESSION TAG	UNP Q9WXR9
A	-4	HIS	-	EXPRESSION TAG	UNP Q9WXR9
A	-3	HIS	-	EXPRESSION TAG	UNP Q9WXR9
A	-2	HIS	-	EXPRESSION TAG	UNP Q9WXR9
A	-1	HIS	-	EXPRESSION TAG	UNP Q9WXR9
A	0	HIS	-	EXPRESSION TAG	UNP Q9WXR9
B	-11	MET	-	LEADER SEQUENCE	UNP Q9WXR9
B	-10	GLY	-	LEADER SEQUENCE	UNP Q9WXR9
B	-9	SER	-	LEADER SEQUENCE	UNP Q9WXR9
B	-8	ASP	-	LEADER SEQUENCE	UNP Q9WXR9
B	-7	LYS	-	LEADER SEQUENCE	UNP Q9WXR9
B	-6	ILE	-	LEADER SEQUENCE	UNP Q9WXR9
B	-5	HIS	-	EXPRESSION TAG	UNP Q9WXR9
B	-4	HIS	-	EXPRESSION TAG	UNP Q9WXR9
B	-3	HIS	-	EXPRESSION TAG	UNP Q9WXR9
B	-2	HIS	-	EXPRESSION TAG	UNP Q9WXR9
B	-1	HIS	-	EXPRESSION TAG	UNP Q9WXR9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	HIS	-	EXPRESSION TAG	UNP Q9WXR9
C	-11	MET	-	LEADER SEQUENCE	UNP Q9WXR9
C	-10	GLY	-	LEADER SEQUENCE	UNP Q9WXR9
C	-9	SER	-	LEADER SEQUENCE	UNP Q9WXR9
C	-8	ASP	-	LEADER SEQUENCE	UNP Q9WXR9
C	-7	LYS	-	LEADER SEQUENCE	UNP Q9WXR9
C	-6	ILE	-	LEADER SEQUENCE	UNP Q9WXR9
C	-5	HIS	-	EXPRESSION TAG	UNP Q9WXR9
C	-4	HIS	-	EXPRESSION TAG	UNP Q9WXR9
C	-3	HIS	-	EXPRESSION TAG	UNP Q9WXR9
C	-2	HIS	-	EXPRESSION TAG	UNP Q9WXR9
C	-1	HIS	-	EXPRESSION TAG	UNP Q9WXR9
C	0	HIS	-	EXPRESSION TAG	UNP Q9WXR9

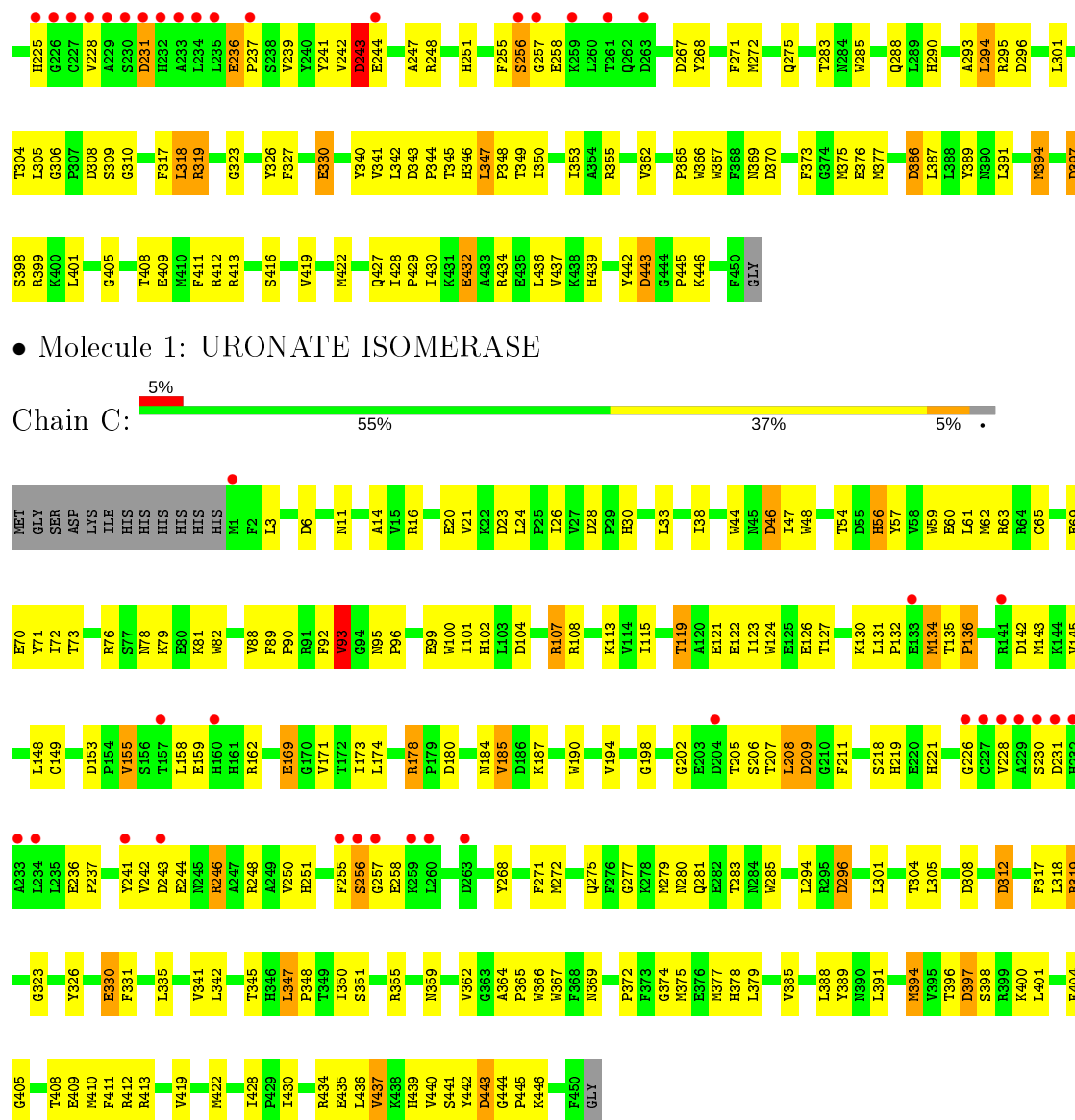
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	40	Total O 40 40	0	0
2	B	33	Total O 33 33	0	0
2	C	34	Total O 34 34	0	0



● Molecule 1: URONATE ISOMERASE





• Molecule 1: URONATE ISOMERASE

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.41 Å   79.96 Å   89.43 Å 115.73°   97.57°   110.44°	Depositor
Resolution (Å)	24.99 – 2.85 24.99 – 2.75	Depositor EDS
% Data completeness (in resolution range)	96.6 (24.99-2.85) 96.7 (24.99-2.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.91 (at 2.76 Å)	Xtriage
Refinement program	REFMAC 5.0	Depositor
R, $R_{free}$	0.234 , 0.274 0.248 , 0.280	Depositor DCC
$R_{free}$ test set	4272 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.1	Xtriage
Anisotropy	0.297	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 63.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.021 for k,h,-h-k-l	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	11175	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.61% 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](#)

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	0.52	0/3793	0.88	12/5136 (0.2%)
1	B	0.54	0/3782	0.90	12/5121 (0.2%)
1	C	0.54	0/3782	0.89	12/5121 (0.2%)
All	All	0.54	0/11357	0.89	36/15378 (0.2%)

There are no bond length outliers.

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	243	ASP	N-CA-CB	-9.25	93.94	110.60
1	B	243	ASP	CB-CG-OD2	8.96	126.37	118.30
1	A	104	ASP	CB-CG-OD2	7.03	124.62	118.30
1	A	132	PRO	CA-N-CD	-6.98	101.72	111.50
1	B	23	ASP	CB-CG-OD2	6.65	124.29	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	3696	0	3621	217	7
1	B	3686	0	3614	202	11
1	C	3686	0	3614	223	4
2	A	40	0	0	3	0
2	B	33	0	0	4	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	34	0	0	5	0
All	All	11175	0	10849	603	11

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 28.

The worst 5 of 603 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:194:VAL:HG12	1:A:205:THR:CG2	1.53	1.34
1:A:251:HIS:ND1	2:A:454:HOH:O	1.65	1.21
1:B:62:MET:CE	1:B:72:ILE:HG12	1.73	1.19
1:A:194:VAL:CG1	1:A:205:THR:CG2	2.22	1.17
1:C:400:LYS:HD3	2:C:458:HOH:O	1.40	1.17

The worst 5 of 11 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:204:ASP:OD2	1:B:112:LYS:NZ[1_655]	0.84	1.36
1:A:128:LYS:NZ	1:B:169:GLU:OE2[1_665]	1.20	1.00
1:B:128:LYS:NZ	1:C:169:GLU:CD[1_556]	1.49	0.71
1:B:128:LYS:NZ	1:C:169:GLU:OE2[1_556]	1.51	0.69
1:A:255:PHE:O	1:B:5:GLU:OE2[1_655]	1.58	0.62

## 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	449/463 (97%)	425 (95%)	20 (4%)	4 (1%)	17 43
1	B	448/463 (97%)	420 (94%)	25 (6%)	3 (1%)	22 50

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	448/463 (97%)	422 (94%)	23 (5%)	3 (1%)	22	50
All	All	1345/1389 (97%)	1267 (94%)	68 (5%)	10 (1%)	22	50

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	257	GLY
1	B	257	GLY
1	C	257	GLY
1	B	169	GLU
1	B	243	ASP

### 5.3.2 Protein sidechains ⓘ

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	399/409 (98%)	367 (92%)	32 (8%)	12	31
1	B	398/409 (97%)	367 (92%)	31 (8%)	12	32
1	C	398/409 (97%)	368 (92%)	30 (8%)	13	34
All	All	1195/1227 (97%)	1102 (92%)	93 (8%)	12	32

5 of 93 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	155	VAL
1	B	330	GLU
1	C	330	GLU
1	B	171	VAL
1	B	256	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	288	GLN
1	C	56	HIS
1	C	251	HIS
1	B	221	HIS
1	C	288	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 carbohydrates 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	451/463 (97%)	-0.03	18 (3%) 38 32	8, 28, 84, 122	18 (3%)
1	B	450/463 (97%)	-0.02	24 (5%) 26 22	4, 27, 83, 122	19 (4%)
1	C	450/463 (97%)	-0.02	23 (5%) 28 23	8, 28, 83, 122	19 (4%)
All	All	1351/1389 (97%)	-0.03	65 (4%) 30 26	4, 28, 84, 122	56 (4%)

The worst 5 of 65 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	259	LYS	5.7
1	B	232	HIS	5.3
1	A	232	HIS	5.3
1	A	227	CYS	5.0
1	C	232	HIS	4.9

### 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 carbohydrates 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.