



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 16, 2018 – 09:40 am GMT

PDB ID : 1I3M  
Title : MOLECULAR BASIS FOR SEVERE EPIMERASE-DEFICIENCY GALACTOSEMIA: X-RAY STRUCTURE OF THE HUMAN V94M-SUBSTITUTED UDP-GALACTOSE 4-EPIMERASE  
Authors : Thoden, J.B.; Wohlers, T.M.; Fridovich-Keil, J.L.; Holden, H.M.  
Deposited on : 2001-02-15  
Resolution : 1.50 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  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

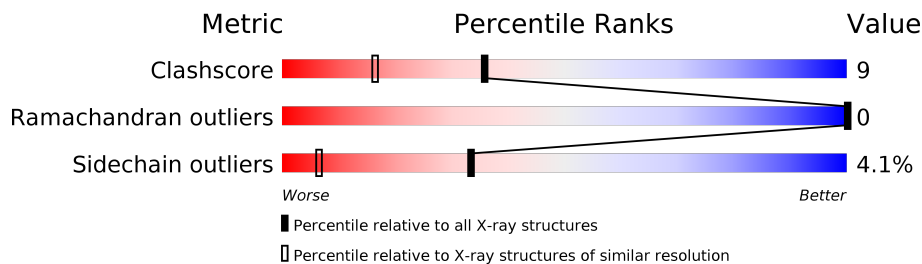
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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	122078	2726 (1.50-1.50)
Ramachandran outliers	120005	2660 (1.50-1.50)
Sidechain outliers	119972	2658 (1.50-1.50)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	348	 75% 21% ..
1	B	348	 76% 20% ..

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6402 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 UDP-GLUCOSE 4-EPIMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	347	Total	C	N	O	S	0	3	0
			2699	1707	468	508	16			
1	B	345	Total	C	N	O	S	0	5	0
			2689	1705	464	504	16			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	94	MET	VAL	ENGINEERED	UNP Q14376
B	94	MET	VAL	ENGINEERED	UNP Q14376

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	3	Total	Cl	0	0
			3	3		
2	A	1	Total	Cl	0	0
			1	1		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

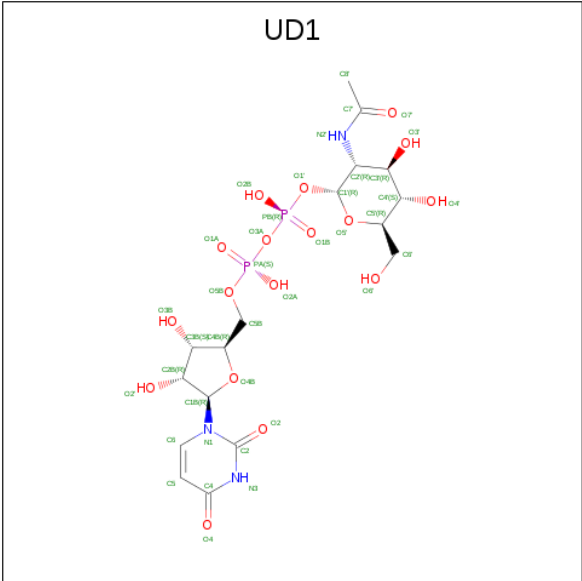
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Mg	0	0
			1	1		

- Molecule 4 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: C<sub>21</sub>H<sub>27</sub>N<sub>7</sub>O<sub>14</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
4	B	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

- Molecule 5 is URIDINE-DIPHOSPHATE-N-ACETYLGLUCOSAMINE (three-letter code: UD1) (formula: C<sub>17</sub>H<sub>27</sub>N<sub>3</sub>O<sub>17</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	1
			40	17	3	18	2		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	N	O	P	0	0
			39	17	3	17	2		

- Molecule 6 is water.

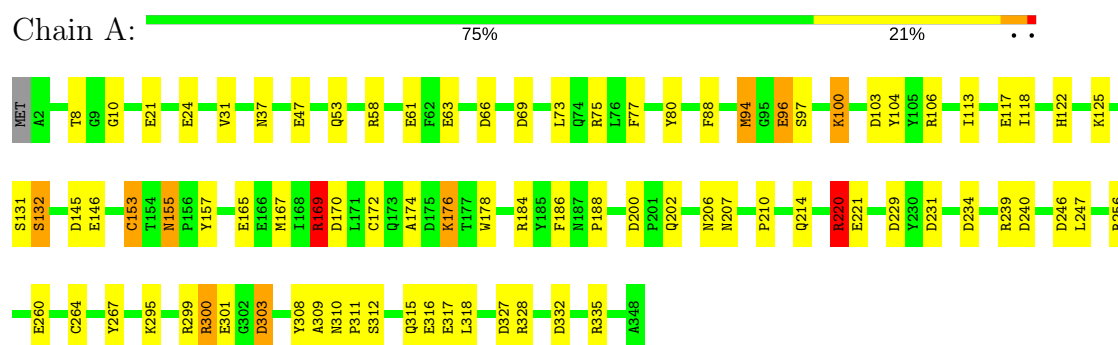
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	418	Total	O	0	0
			418	418		
6	B	424	Total	O	0	0
			424	424		

### 3 Residue-property plots [i](#)

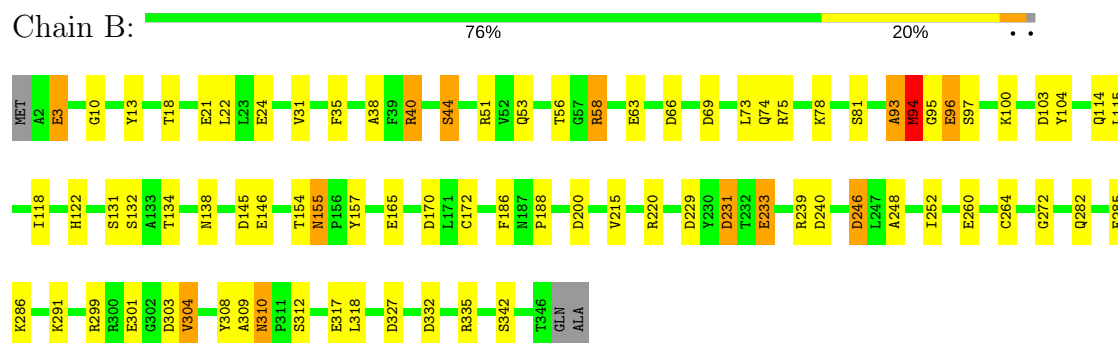
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

#### • Molecule 1: UDP-GLUCOSE 4-EPIMERASE



#### • Molecule 1: UDP-GLUCOSE 4-EPIMERASE



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.90Å 89.90Å 96.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.50	Depositor
% Data completeness (in resolution range)	98.4 (50.00-1.50)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	TNT	Depositor
R, $R_{free}$	0.184 , 0.214	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6402	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	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: MG, UD1, NAD, CL

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	0.94	16/2768 (0.6%)	1.26	35/3744 (0.9%)
1	B	0.95	12/2770 (0.4%)	1.25	30/3748 (0.8%)
All	All	0.95	28/5538 (0.5%)	1.26	65/7492 (0.9%)

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	317	GLU	CD-OE2	7.54	1.33	1.25
1	B	146	GLU	CD-OE2	7.33	1.33	1.25
1	B	233	GLU	CD-OE2	7.00	1.33	1.25
1	A	316	GLU	CD-OE2	6.68	1.32	1.25
1	A	21	GLU	CD-OE2	6.60	1.32	1.25

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	145	ASP	CB-CG-OD2	-9.92	109.37	118.30
1	A	220	ARG	NE-CZ-NH1	9.14	124.87	120.30
1	A	234	ASP	CB-CG-OD2	-8.71	110.46	118.30
1	B	145	ASP	CB-CG-OD1	8.02	125.52	118.30
1	B	335	ARG	NE-CZ-NH1	7.80	124.20	120.30

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	2699	0	2649	44	0
1	B	2689	0	2644	46	0
2	A	1	0	0	0	0
2	B	3	0	0	1	0
3	A	1	0	0	0	0
4	A	44	0	26	4	0
4	B	44	0	26	3	0
5	A	40	0	6	0	0
5	B	39	0	25	4	0
6	A	418	0	0	3	1
6	B	424	0	0	8	1
All	All	6402	0	5376	95	1

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 9.

The worst 5 of 95 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:132:SER:HB2	4:A:400:NAD:H5N	1.49	0.95
1:B:73:LEU:HB3	1:B:118[A]:ILE:HD13	1.54	0.89
1:B:58:ARG:HG3	1:B:58:ARG:HH11	1.53	0.74
1:A:100:LYS:HD3	1:A:103:ASP:HB2	1.70	0.73
1:A:155:ASN:OD1	1:A:157:TYR:HB3	1.92	0.70

All (1) 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)
6:A:1189:HOH:O	6:B:1638:HOH:O[4_456]	2.18	0.02

## 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	348/348 (100%)	340 (98%)	8 (2%)	0	100	100
1	B	349/348 (100%)	342 (98%)	7 (2%)	0	100	100
All	All	697/696 (100%)	682 (98%)	15 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	283/282 (100%)	273 (96%)	10 (4%)	39	10
1	B	284/282 (101%)	271 (95%)	13 (5%)	29	5
All	All	567/564 (100%)	544 (96%)	23 (4%)	33	6

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	3	GLU
1	B	44	SER
1	B	310	ASN
1	B	40	ARG
1	B	58	ARG

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	337	GLN
1	A	339	GLN
1	B	138	ASN
1	A	315	GLN
1	B	114	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 10 ligands modelled in this entry, 5 are monoatomic - leaving 5 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$
4	NAD	A	400	-	40,48,48	1.01	3 (7%)	44,73,73	1.67	5 (11%)
5	UD1	A	401[A]	-	33,41,41	1.21	4 (12%)	45,62,62	2.60	5 (11%)
5	UD1	A	401[B]	-	33,41,41	1.21	4 (12%)	45,62,62	2.60	5 (11%)
4	NAD	B	900	-	40,48,48	1.22	4 (10%)	44,73,73	1.91	6 (13%)
5	UD1	B	901	-	33,41,41	1.28	4 (12%)	45,62,62	2.58	8 (17%)

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
4	NAD	A	400	-	-	0/22/62/62	0/5/5/5
5	UD1	A	401[A]	-	-	0/22/63/63	0/3/3/3
5	UD1	A	401[B]	-	-	0/22/63/63	0/3/3/3
4	NAD	B	900	-	-	0/22/62/62	0/5/5/5

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	UD1	B	901	-	-	0/22/63/63	0/3/3/3

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	901	UD1	C6-C5	-2.70	1.32	1.38
5	A	401[B]	UD1	C6-C5	-2.65	1.32	1.38
5	A	401[A]	UD1	C6-C5	-2.65	1.32	1.38
4	B	900	NAD	PA-O2A	-2.37	1.43	1.55
4	A	400	NAD	PA-O2A	-2.09	1.45	1.55

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	900	NAD	C5N-C6N-N1N	-6.04	111.17	120.39
4	A	400	NAD	C5N-C4N-C3N	-5.45	113.95	120.35
4	A	400	NAD	C5N-C6N-N1N	-5.36	112.20	120.39
4	B	900	NAD	C5N-C4N-C3N	-5.00	114.47	120.35
5	A	401[B]	UD1	C5-C4-N3	-3.93	114.03	123.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	400	NAD	4	0
4	B	900	NAD	3	0
5	B	901	UD1	4	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 is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.