



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 08:04 am BST

PDB ID : 5MNI  
Title : Escherichia coli AGPase mutant R130A apo form  
Authors : Cifuentes, J.O.; Comino, N.; Marina, A.; Orrantia, A.; Eguskiza, A.; Guerin, M.E.  
Deposited on : 2016-12-13  
Resolution : 3.09 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

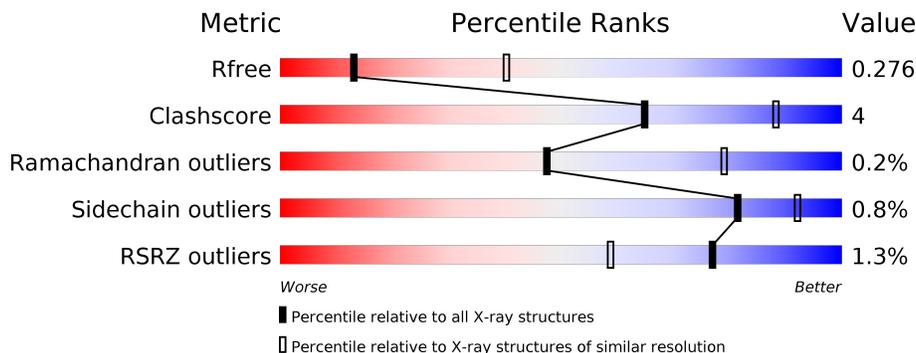
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.09 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	431	 % 83% 8% 8%
1	B	431	 84% 9% 6%
1	C	431	 % 81% 10% 9%
1	D	431	 2% 86% 8% 5%
1	E	431	 87% 6% 7%
1	F	431	 2% 83% 7% 10%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	431	
1	H	431	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 45964 atoms, of which 22263 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 Glucose-1-phosphate adenylyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	E	401	5912	1922	2886	521	562	21	0	3	0
1	G	393	5661	1862	2738	499	543	19	0	0	0
1	H	393	5304	1774	2512	485	516	17	0	2	0
1	F	390	5786	1881	2825	505	555	20	0	2	0
1	A	396	5737	1880	2778	510	551	18	0	2	0
1	B	406	5945	1944	2887	518	576	20	0	0	0
1	D	409	5795	1909	2788	521	557	20	0	0	0
1	C	394	5824	1891	2849	514	550	20	0	0	0

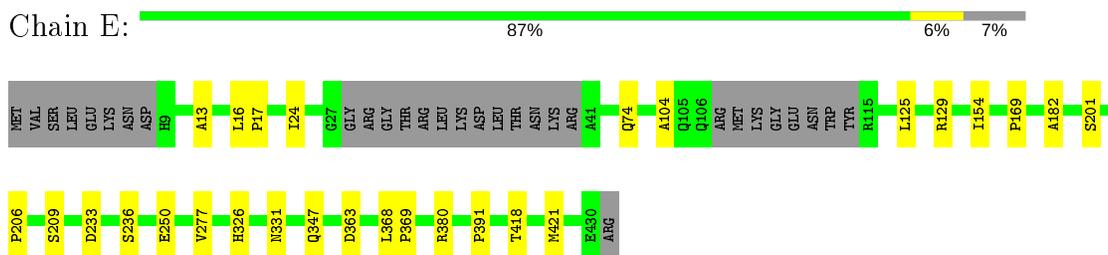
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	130	ALA	ARG	engineered mutation	UNP P0A6V1
G	130	ALA	ARG	engineered mutation	UNP P0A6V1
H	130	ALA	ARG	engineered mutation	UNP P0A6V1
F	130	ALA	ARG	engineered mutation	UNP P0A6V1
A	130	ALA	ARG	engineered mutation	UNP P0A6V1
B	130	ALA	ARG	engineered mutation	UNP P0A6V1
D	130	ALA	ARG	engineered mutation	UNP P0A6V1
C	130	ALA	ARG	engineered mutation	UNP P0A6V1

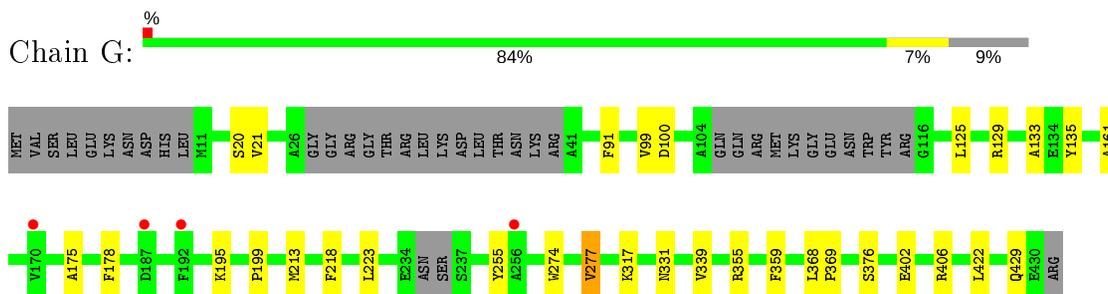
### 3 Residue-property plots [i](#)

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.

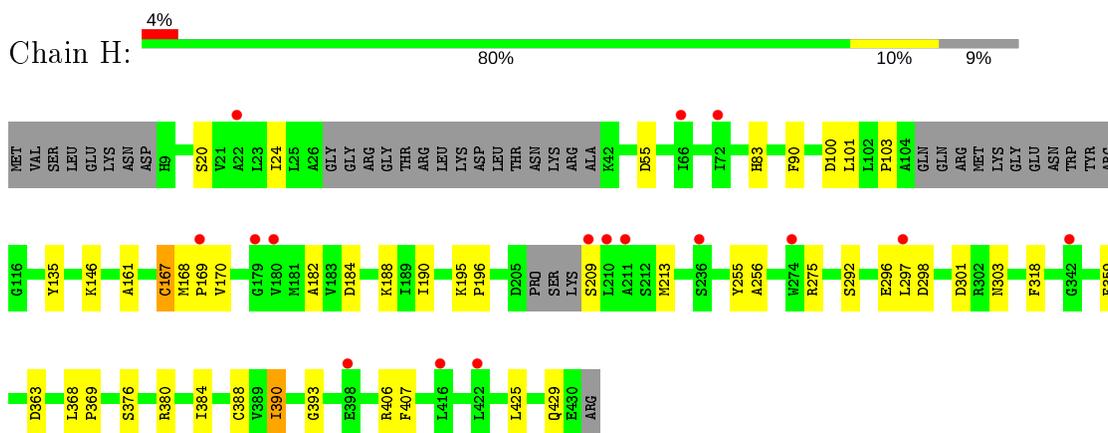
- Molecule 1: Glucose-1-phosphate adenylyltransferase



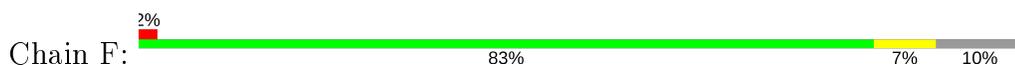
- Molecule 1: Glucose-1-phosphate adenylyltransferase



- Molecule 1: Glucose-1-phosphate adenylyltransferase

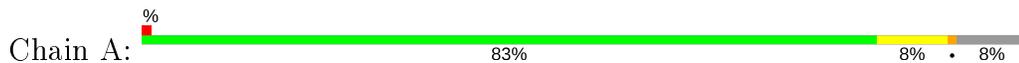


- Molecule 1: Glucose-1-phosphate adenylyltransferase

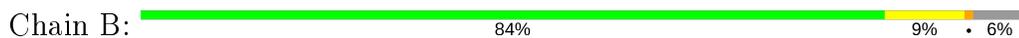




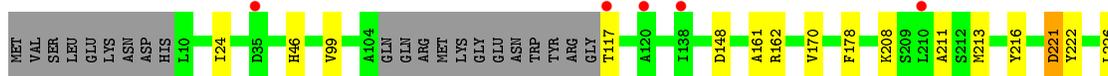
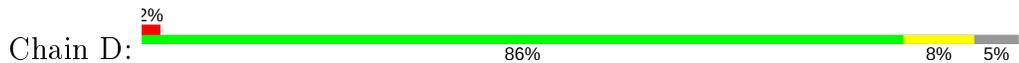
• Molecule 1: Glucose-1-phosphate adenylyltransferase



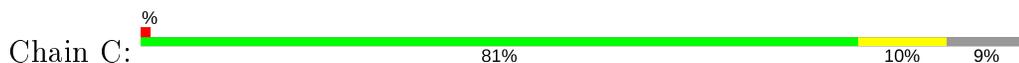
• Molecule 1: Glucose-1-phosphate adenylyltransferase

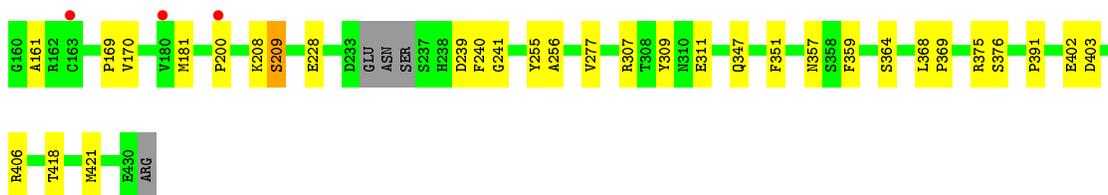


• Molecule 1: Glucose-1-phosphate adenylyltransferase



• Molecule 1: Glucose-1-phosphate adenylyltransferase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.22Å 147.60Å 125.57Å 90.00° 91.52° 90.00°	Depositor
Resolution (Å)	73.80 – 3.09 95.62 – 3.09	Depositor EDS
% Data completeness (in resolution range)	99.5 (73.80-3.09) 88.9 (95.62-3.09)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.00 (at 3.07Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.235 , 0.273 0.238 , 0.276	Depositor DCC
$R_{free}$ test set	3174 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	72.0	Xtrriage
Anisotropy	0.234	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 44.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.028 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	45964	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	92.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.35% 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	0.33	0/3023	0.50	0/4120
1	B	0.34	0/3126	0.48	0/4261
1	C	0.32	0/3039	0.48	0/4136
1	D	0.36	0/3073	0.51	0/4197
1	E	0.33	0/3091	0.48	0/4209
1	F	0.32	0/3024	0.48	0/4117
1	G	0.31	0/2988	0.47	0/4080
1	H	0.32	0/2851	0.50	0/3908
All	All	0.33	0/24215	0.49	0/33028

There are no bond length outliers.

There are no bond angle outliers.

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	2959	2778	2774	25	0
1	B	3058	2887	2886	30	0
1	C	2975	2849	2849	24	0
1	D	3007	2788	2788	33	2
1	E	3026	2886	2877	17	0
1	F	2961	2825	2821	19	2
1	G	2923	2738	2738	20	0
1	H	2792	2512	2510	32	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	23701	22263	22243	190	2

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 4.

The worst 5 of 190 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:D:226:LEU:HG	1:D:244:LEU:HD11	1.72	0.71
1:A:397:GLY:N	1:A:403:ASP:OD2	2.28	0.67
1:H:146:LYS:NZ	1:H:298:ASP:O	2.28	0.67
1:B:243:ASP:N	1:B:243:ASP:OD1	2.28	0.66
1:C:200:PRO:O	1:C:209:SER:HB3	1.95	0.65

All (2) 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:F:322:ARG:HH22	1:D:363:ASP:OD1[2_657]	1.35	0.25
1:F:322:ARG:NH2	1:D:363:ASP:OD1[2_657]	2.10	0.10

## 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	388/431 (90%)	369 (95%)	18 (5%)	1 (0%)	41 73
1	B	400/431 (93%)	385 (96%)	14 (4%)	1 (0%)	41 73
1	C	386/431 (90%)	373 (97%)	12 (3%)	1 (0%)	41 73
1	D	405/431 (94%)	380 (94%)	25 (6%)	0	100 100
1	E	395/431 (92%)	379 (96%)	14 (4%)	2 (0%)	29 64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	380/431 (88%)	365 (96%)	14 (4%)	1 (0%)	41	73
1	G	385/431 (89%)	364 (94%)	20 (5%)	1 (0%)	41	73
1	H	385/431 (89%)	366 (95%)	19 (5%)	0	100	100
All	All	3124/3448 (91%)	2981 (95%)	136 (4%)	7 (0%)	47	79

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	391	PRO
1	A	277	VAL
1	B	277	VAL
1	G	277	VAL
1	E	277	VAL

### 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	297/372 (80%)	294 (99%)	3 (1%)	76	90
1	B	314/372 (84%)	310 (99%)	4 (1%)	69	87
1	C	307/372 (82%)	305 (99%)	2 (1%)	84	93
1	D	294/372 (79%)	292 (99%)	2 (1%)	84	93
1	E	311/372 (84%)	311 (100%)	0	100	100
1	F	309/372 (83%)	306 (99%)	3 (1%)	76	90
1	G	294/372 (79%)	293 (100%)	1 (0%)	92	96
1	H	258/372 (69%)	254 (98%)	4 (2%)	62	84
All	All	2384/2976 (80%)	2365 (99%)	19 (1%)	81	92

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	276	ASP

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	423	ARG
1	D	221	ASP
1	F	322	ARG
1	D	380	ARG

Some 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](#)

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	396/431 (91%)	0.11	3 (0%) 86 72	49, 84, 128, 166	0
1	B	406/431 (94%)	0.08	2 (0%) 91 81	48, 77, 113, 164	0
1	C	394/431 (91%)	0.16	3 (0%) 86 72	47, 76, 114, 139	0
1	D	409/431 (94%)	0.15	7 (1%) 70 49	47, 84, 120, 161	0
1	E	401/431 (93%)	0.05	0 100 100	50, 76, 113, 149	0
1	F	390/431 (90%)	0.14	7 (1%) 68 47	53, 83, 119, 162	0
1	G	393/431 (91%)	0.18	4 (1%) 82 67	55, 89, 139, 157	0
1	H	393/431 (91%)	0.35	16 (4%) 37 18	56, 105, 148, 172	0
All	All	3182/3448 (92%)	0.15	42 (1%) 77 59	47, 84, 130, 172	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	211	ALA	8.0
1	H	180	VAL	4.1
1	H	210	LEU	4.1
1	H	209	SER	3.3
1	G	256	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 carbohydrates in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.