



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 25, 2023 – 12:22 pm GMT

PDB ID : 8CAG  
Title : Hypoxanthine-guanine phosphoribosyltransferase from E. coli  
Authors : Timofeev, V.I.; Shevtsov, M.B.; Abramchik, Y.A.; Kostromina, M.A.; Zayats, E.A.; Kuranova, I.P.; Esipov, R.S.  
Deposited on : 2023-01-24  
Resolution : 2.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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  
Xtriage (Phenix) : 1.13  
EDS : 2.31.3  
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) : 2.31.3

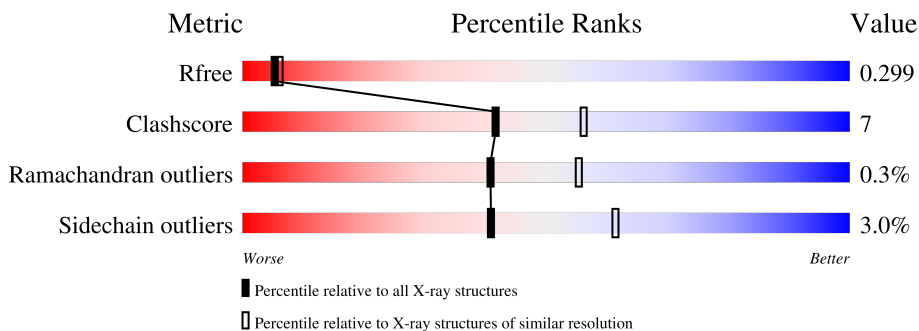
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	176	
1	B	176	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2724 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 Hypoxanthine phosphoribosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	172	1375	873	235	260	7	0	2	0
1	B	166	1314	835	225	247	7	0	0	0

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

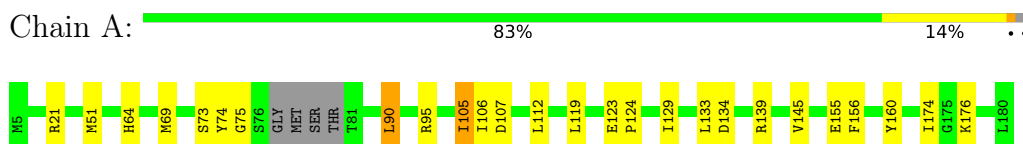
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	15	Total 15	O 15	0	0
3	B	18	Total 18	O 18	0	0

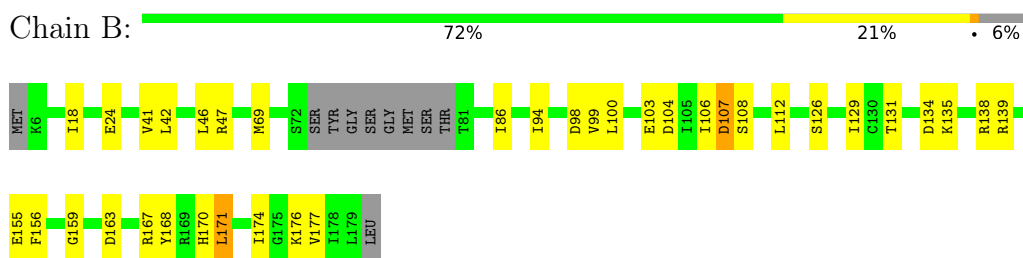
### 3 Residue-property plots [i](#)

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. 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. 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: Hypoxanthine phosphoribosyltransferase



- Molecule 1: Hypoxanthine phosphoribosyltransferase



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	84.65Å 84.65Å 167.24Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.74 – 2.40 29.74 – 2.40	Depositor EDS
% Data completeness (in resolution range)	98.3 (29.74-2.40) 98.4 (29.74-2.40)	Depositor EDS
$R_{merge}$	0.23	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.00 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.0	Depositor
R, $R_{free}$	0.254 , 0.293 0.258 , 0.299	Depositor DCC
$R_{free}$ test set	1396 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.8	Xtrriage
Anisotropy	0.042	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 17.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	0.226 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2724	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.44% 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: MG

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	0.76	0/1397	0.95	0/1882
1	B	0.76	0/1332	0.94	0/1795
All	All	0.76	0/2729	0.94	0/3677

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	1375	0	1396	17	0
1	B	1314	0	1342	22	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	15	0	0	0	0
3	B	18	0	0	1	0
All	All	2724	0	2738	38	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 7.

All (38) 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:69:MET:HG2	1:A:90:LEU:HD13	1.75	0.69
1:A:155:GLU:OE2	1:A:176:LYS:NZ	2.22	0.69
1:A:69:MET:CG	1:A:90:LEU:HD13	2.24	0.67
1:B:159:GLY:HA2	1:B:168:TYR:O	1.96	0.65
1:B:134:ASP:OD1	1:B:139:ARG:NH1	2.31	0.63
1:B:135:LYS:NZ	1:B:155:GLU:O	2.30	0.60
1:B:135:LYS:HD3	1:B:154:ASP:OD1	2.03	0.58
1:A:112:LEU:HG	1:A:129:ILE:HD13	1.88	0.56
1:B:176:LYS:NZ	3:B:301:HOH:O	2.32	0.55
1:B:168:TYR:HB3	1:B:171:LEU:HD22	1.88	0.54
1:B:104:ASP:N	1:B:104:ASP:OD1	2.40	0.52
1:B:106:ILE:HG12	1:B:112:LEU:CD2	2.40	0.52
1:A:105:ILE:HG23	1:A:133:LEU:HB2	1.92	0.51
1:B:106:ILE:CD1	1:B:112:LEU:HD23	2.42	0.50
1:A:106:ILE:CD1	1:A:145:VAL:HG21	2.43	0.48
1:A:133:LEU:HD11	1:A:174:ILE:HD12	1.96	0.47
1:B:42:LEU:HD22	1:B:100:LEU:HB3	1.98	0.46
1:A:51:MET:SD	1:A:160:TYR:HA	2.55	0.46
1:B:98:ASP:OD1	1:B:126:SER:OG	2.24	0.45
1:A:69:MET:SD	1:A:119:LEU:HD21	2.56	0.45
1:B:163:ASP:HA	1:B:167:ARG:O	2.16	0.44
1:A:95:ARG:HB2	1:A:95:ARG:NH1	2.31	0.44
1:B:69:MET:CE	1:B:86:ILE:HD12	2.47	0.44
1:A:64:HIS:CE1	1:B:170:HIS:CG	3.06	0.44
1:A:73:SER:O	1:A:75:GLY:N	2.51	0.43
1:B:143:VAL:O	1:B:145:VAL:HG23	2.19	0.43
1:B:103:GLU:O	1:B:131:THR:HA	2.19	0.43
1:A:134:ASP:OD1	1:A:139:ARG:NH2	2.52	0.42
1:A:106:ILE:HG12	1:A:112:LEU:CD2	2.49	0.42
1:B:46:LEU:HB3	1:B:47:ARG:HA	2.01	0.42
1:B:156:PHE:CD2	1:B:177:VAL:HB	2.55	0.41
1:A:21:ARG:HD2	1:A:21:ARG:HA	1.91	0.41
1:A:106:ILE:HD12	1:A:145:VAL:HG11	2.01	0.41
1:A:123:GLU:N	1:A:124:PRO:CD	2.84	0.41
1:B:107:ASP:HB3	1:B:108:SER:H	1.62	0.41
1:B:18:ILE:HD11	1:B:174:ILE:HG12	2.02	0.41
1:B:94:ILE:HB	1:B:99:VAL:HG21	2.03	0.40
1:B:112:LEU:HG	1:B:129:ILE:HD13	2.04	0.40

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	170/176 (97%)	163 (96%)	6 (4%)	1 (1%)	25	36
1	B	162/176 (92%)	156 (96%)	6 (4%)	0	100	100
All	All	332/352 (94%)	319 (96%)	12 (4%)	1 (0%)	41	55

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	74	TYR

### 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	153/157 (98%)	149 (97%)	4 (3%)	46	66
1	B	147/157 (94%)	142 (97%)	5 (3%)	37	56
All	All	300/314 (96%)	291 (97%)	9 (3%)	41	61

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	90	LEU
1	A	105	ILE
1	A	107	ASP
1	A	156	PHE
1	B	24	GLU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	41	VAL
1	B	107	ASP
1	B	138	ARG
1	B	171	LEU

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](#)

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.