wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 04:40 pm BST

PDB ID	:	1GKM
Title	:	HISTIDINE AMMONIA-LYASE (HAL) FROM PSEUDOMONAS PUTIDA INHIBITED WITH L-CYSTEINE
Authors	:	Baedeker, M.; Schulz, G.E.
Deposited on	:	2001-08-16
Resolution	:	1.00 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

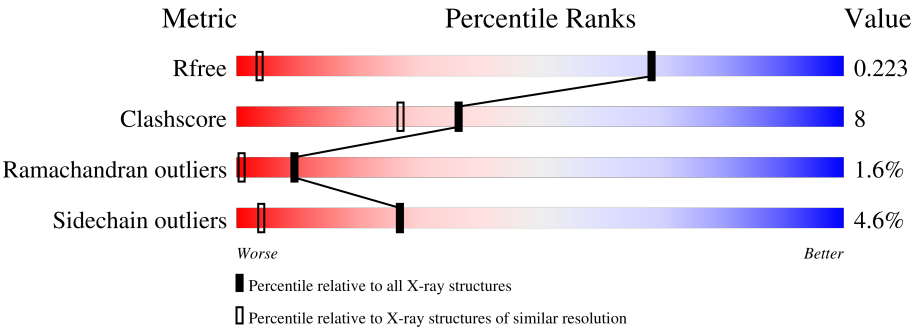
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.00 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	1050 (1.06-0.94)
Clashscore	141614	1117 (1.06-0.94)
Ramachandran outliers	138981	1043 (1.06-0.94)
Sidechain outliers	138945	1045 (1.06-0.94)

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	B	507	84% 11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	CYS	B	601	X	X	-	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4360 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.

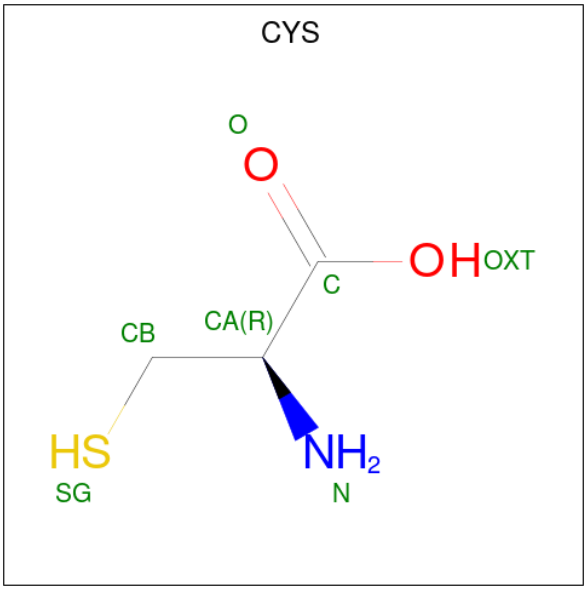
- Molecule 1 is a protein called Histidine ammonia-lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	507	3797	2376	674	729	18	0	12	0

There are 4 discrepancies between the modelled and reference sequences:

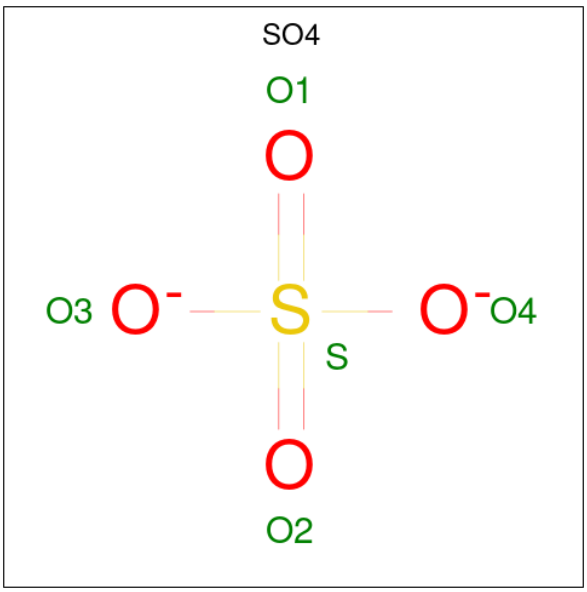
Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	ALA	deletion	UNP P21310
B	?	-	SER	deletion	UNP P21310
B	142	ZRF	GLY	conflict	UNP P21310
B	271	ALA	CYS	conflict	UNP P21310

- Molecule 2 is CYSTEINE (three-letter code: CYS) (formula: $C_3H_7NO_2S$).



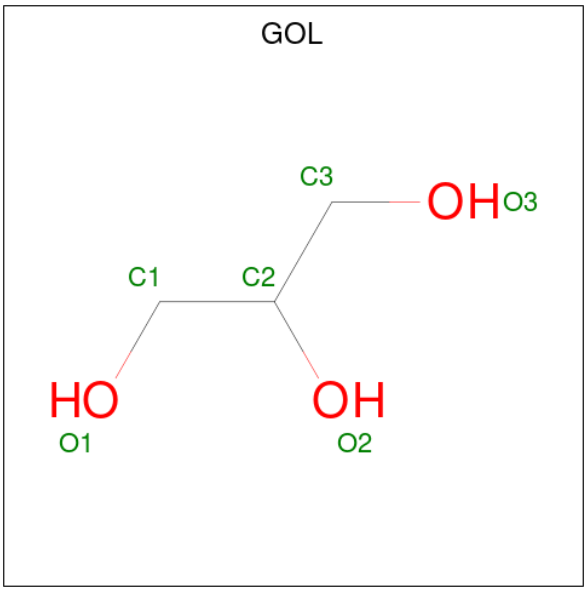
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	B	1	7	3	1	2	1	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total O S 5 4 1	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 6 3 3	0	0

- Molecule 5 is water.

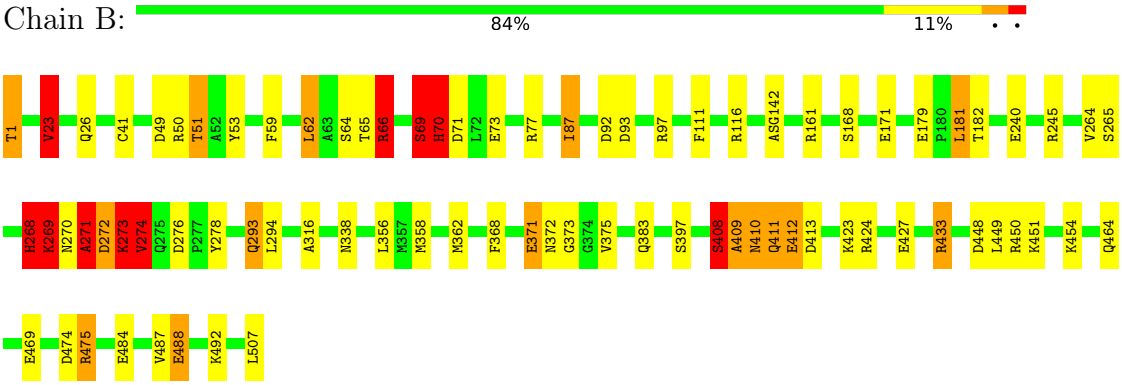
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	545	Total O 545 545	0	0

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: Histidine ammonia-lyase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	79.27Å 116.79Å 129.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 1.00 29.26 – 1.01	Depositor EDS
% Data completeness (in resolution range)	96.0 (40.00-1.00) 95.8 (29.26-1.01)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.39 (at 1.01Å)	Xtriage
Refinement program	SHELX	Depositor
R, R_{free}	0.119 , 0.135 0.227 , 0.223	Depositor DCC
R_{free} test set	15248 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	6.9	Xtriage
Anisotropy	0.312	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 72.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4360	wwPDB-VP
Average B, all atoms (Å^2)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.26% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, GOL, ZRF

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	B	0.89	5/3899 (0.1%)	1.36	55/5284 (1.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	507	LEU	C-OXT	9.98	1.42	1.23
1	B	411	GLN	C-O	6.41	1.35	1.23
1	B	271	ALA	N-CA	5.91	1.58	1.46
1	B	397[A]	SER	CB-OG	-5.86	1.34	1.42
1	B	397[B]	SER	CB-OG	-5.86	1.34	1.42

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	53	TYR	CB-CG-CD2	17.73	131.64	121.00
1	B	97	ARG	NE-CZ-NH2	-14.53	113.04	120.30
1	B	87	ILE	CG1-CB-CG2	-12.93	82.95	111.40
1	B	97	ARG	NE-CZ-NH1	11.46	126.03	120.30
1	B	368	PHE	CB-CG-CD2	-10.82	113.23	120.80

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	B	3797	0	3836	60	0
2	B	7	0	3	1	0
3	B	5	0	0	1	0
4	B	6	0	8	0	0
5	B	545	0	0	12	0
All	All	4360	0	3847	61	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 8.

The worst 5 of 61 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)
2:B:601:CYS:O	3:B:602:SO4:O3	1.83	0.94
1:B:274:VAL:HG22	1:B:274:VAL:O	1.80	0.81
1:B:268:HIS:ND1	1:B:271:ALA:HB2	1.98	0.78
1:B:1:THR:OG1	1:B:23:VAL:HG23	1.89	0.72
1:B:265:SER:HA	1:B:268:HIS:CD2	2.25	0.72

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	B	514/507 (101%)	496 (96%)	10 (2%)	8 (2%)	9 1

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	271	ALA
1	B	272	ASP
1	B	410	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	270	ASN
1	B	273	LYS

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	B	399/387 (103%)	382 (96%)	17 (4%)	29 4

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	408	SER
1	B	475	ARG
1	B	181	LEU
1	B	268	HIS
1	B	269	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	193	ASN
1	B	464	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)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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	$\#\|Z\| > 2$
1	ZRF	B	142	1,2	13,14,15	1.79	3 (23%)	15,19,21	2.04	5 (33%)

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
1	ZRF	B	142	1,2	-	2/4/25/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	142	ZRF	CA2-C2	-4.44	1.44	1.48
1	B	142	ZRF	O1-CB2	-2.53	1.21	1.32
1	B	142	ZRF	CA1-C1	-2.19	1.48	1.51

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	142	ZRF	C2-N3-C1	5.01	110.50	107.97
1	B	142	ZRF	O2-C2-CA2	3.68	133.03	130.96
1	B	142	ZRF	CA2-C2-N3	-2.27	102.30	103.37
1	B	142	ZRF	O3-C3-CA3	-2.20	119.74	126.39
1	B	142	ZRF	C2-CA2-N2	2.15	110.44	108.93

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	142	ZRF	N2-C1-CA1-CB
1	B	142	ZRF	N3-C1-CA1-CB

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

3 ligands are modelled in this entry.

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	$\#\|Z\| > 2$
3	SO4	B	602	2	4,4,4	0.44	0	6,6,6	0.85	0
2	CYS	B	601	1,3	5,6,6	2.43	3 (60%)	5,7,7	3.15	3 (60%)
4	GOL	B	603	-	5,5,5	0.86	0	5,5,5	0.58	0

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
2	CYS	B	601	1,3	1/1/2/2	6/6/6/6	-
4	GOL	B	603	-	-	0/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	601	CYS	CA-N	3.40	1.66	1.48
2	B	601	CYS	OXT-C	-3.21	1.20	1.30
2	B	601	CYS	CB-CA	-2.33	1.50	1.53

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	601	CYS	CB-CA-C	-4.90	105.01	109.89
2	B	601	CYS	OXT-C-CA	3.52	125.39	113.38
2	B	601	CYS	OXT-C-O	-3.16	116.91	124.09

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	601	CYS	CA

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	601	CYS	O-C-CA-N
2	B	601	CYS	N-CA-CB-SG
2	B	601	CYS	C-CA-CB-SG
2	B	601	CYS	OXT-C-CA-N
2	B	601	CYS	O-C-CA-CB

There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	602	SO4	1	0
2	B	601	CYS	1	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

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.