



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 13, 2020 – 10:32 AM BST

PDB ID : 2QGH
Title : Crystal structure of diaminopimelate decarboxylase from *Helicobacter pylori* complexed with L-lysine
Authors : Hu, T.; Wu, D.; Jiang, H.; Shen, X.
Deposited on : 2007-06-28
Resolution : 2.30 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.14.4.dev1
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) : 2.14.4.dev1

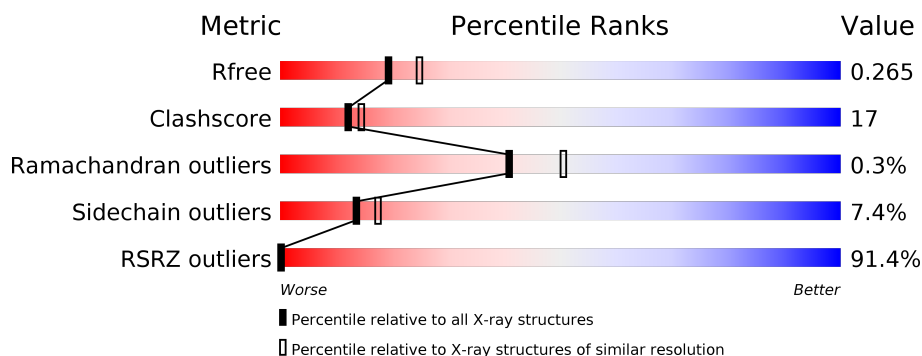
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 ≥ 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	425	<div> <div>85%</div> <div>61% 28% • 7%</div> </div>

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	PLP	A	406	-	-	-	X
3	LYS	A	407	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GOL	A	1003	-	-	-	X

2 Entry composition [i](#)

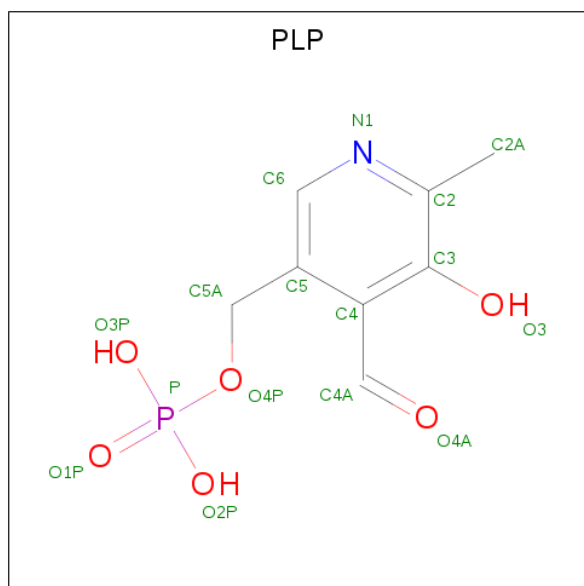
There are 5 unique types of molecules in this entry. The entry contains 3254 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Diaminopimelate decarboxylase.

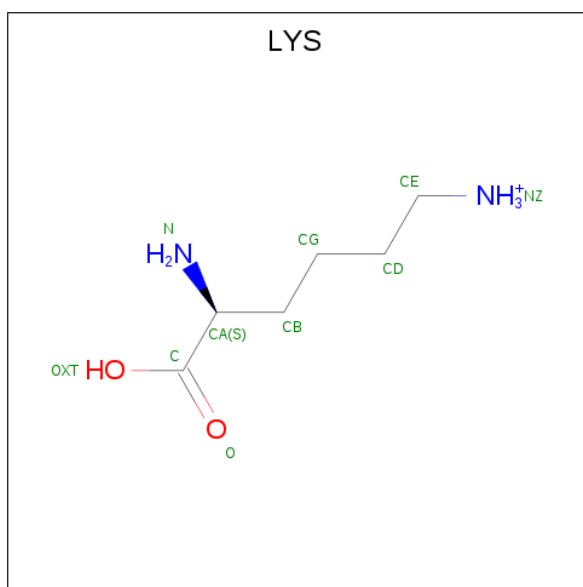
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	394	Total	C	N	O	S	0	0	0
			3102	1999	515	579	9			

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			15	8	1	5	1		

- Molecule 3 is LYSINE (three-letter code: LYS) (formula: $C_6H_{15}N_2O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			10	6	2	2		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	121	Total 121	O 121	0	0

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 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.

- Chain A:**

61% 85% 28% 7%

Residue	Chain A	Count	Category
L402	GLY	1	L402
L403	GLY	1	
L404	GLY	1	
L405	GLY	1	
L406	GLY	1	
P342	LVS	1	P342
P343	LVS	1	
L344	LVS	1	
P345	LVS	1	
P346	LVS	1	
G347	LVS	1	
D348	LVS	1	
K349	LVS	1	
L350	LVS	1	
A351	LVS	1	
L352	LVS	1	L352
E353	LVS	1	
K354	LVS	1	
V355	LVS	1	
G356	LVS	1	
A357	LVS	1	
V358	LVS	1	
G359	LVS	1	
S360	LVS	1	
S361	LVS	1	
M362	LVS	1	M362
A363	LVS	1	
S364	LVS	1	
Q365	LVS	1	
V366	LVS	1	
S367	LVS	1	
S368	LVS	1	
R369	LVS	1	
P370	LVS	1	
K371	LVS	1	
L372	LVS	1	L372
L373	LVS	1	
P374	LVS	1	
L375	LVS	1	
L376	LVS	1	
A376	LVS	1	
L377	LVS	1	
E378	LVS	1	
D379	LVS	1	
HIS	LVS	1	
K381	LVS	1	K381
L382	LVS	1	
R383	LVS	1	
V384	LVS	1	
L385	LVS	1	
R386	LVS	1	
K387	LVS	1	
R388	LVS	1	
E389	LVS	1	
A390	LVS	1	
L391	LVS	1	L391
E392	LVS	1	
L393	LVS	1	
K394	LVS	1	
V395	LVS	1	
R396	LVS	1	
L397	LVS	1	
E398	LVS	1	
E399	LVS	1	
E400	LVS	1	
E401	LVS	1	E401
E402	LVS	1	
E403	LVS	1	
E404	LVS	1	
E405	LVS	1	
E406	LVS	1	
E407	LVS	1	
E408	LVS	1	
E409	LVS	1	
E410	LVS	1	
E411	LVS	1	E411
E412	LVS	1	
E413	LVS	1	
E414	LVS	1	
E415	LVS	1	
E416	LVS	1	
E417	LVS	1	
E418	LVS	1	
E419	LVS	1	
E420	LVS	1	
E421	LVS	1	E421
E422	LVS	1	
E423	LVS	1	
E424	LVS	1	
E425	LVS	1	
E426	LVS	1	
E427	LVS	1	
E428	LVS	1	
E429	LVS	1	
E430	LVS	1	
E431	LVS	1	E431
E432	LVS	1	
E433	LVS	1	
E434	LVS	1	
E435	LVS	1	
E436	LVS	1	
E437	LVS	1	
E438	LVS	1	
E439	LVS	1	
E440	LVS	1	
E441	LVS	1	E441
E442	LVS	1	
E443	LVS	1	
E444	LVS	1	
E445	LVS	1	
E446	LVS	1	
E447	LVS	1	
E448	LVS	1	
E449	LVS	1	
E450	LVS	1	
E451	LVS	1	E451
E452	LVS	1	
E453	LVS	1	
E454	LVS	1	
E455	LVS	1	
E456	LVS	1	
E457	LVS	1	
E458	LVS	1	
E459	LVS	1	
E460	LVS	1	
E461	LVS	1	E461
E462	LVS	1	
E463	LVS	1	
E464	LVS	1	
E465	LVS	1	
E466	LVS	1	
E467	LVS	1	
E468	LVS	1	
E469	LVS	1	
E470	LVS	1	
E471	LVS	1	E471
E472	LVS	1	
E473	LVS	1	
E474	LVS	1	
E475	LVS	1	
E476	LVS	1	
E477	LVS	1	
E478	LVS	1	
E479	LVS	1	
E480	LVS	1	
E481	LVS	1	E481
E482	LVS	1	
E483	LVS	1	
E484	LVS	1	
E485	LVS</		

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	79.25Å 79.25Å 134.28Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	15.00 – 2.30 19.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	96.8 (15.00-2.30) 96.6 (19.00-2.30)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.28 (at 2.30Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.206 , 0.267 0.204 , 0.265	Depositor DCC
R_{free} test set	2146 reflections (9.94%)	wwPDB-VP
Wilson B-factor (Å ²)	14.2	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 44.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.038 for -h,-k,l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3254	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: GOL, PLP

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	0.36	0/3156	0.60	0/4251

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	3102	0	3174	108	0
2	A	15	0	7	1	0
3	A	10	0	9	1	0
4	A	6	0	8	2	0
5	A	121	0	0	9	0
All	All	3254	0	3198	108	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 17.

The worst 5 of 108 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:159:VAL:HG11	1:A:163:GLU:HG2	1.34	1.04
1:A:159:VAL:HB	5:A:1107:HOH:O	1.67	0.94
1:A:387:LYS:HA	1:A:387:LYS:HE2	1.54	0.88
1:A:325:VAL:HG12	1:A:334:THR:HB	1.57	0.86
1:A:53:ILE:HG21	5:A:1122:HOH:O	1.79	0.81

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	386/425 (91%)	369 (96%)	16 (4%)	1 (0%)	41	50

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	329	CYS

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	336/361 (93%)	311 (93%)	25 (7%)	13	17

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	162	LYS
1	A	182	VAL
1	A	391	LEU
1	A	170	TRP
1	A	202	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	137	ASN
1	A	367	ASN
1	A	274	GLN
1	A	99	GLN
1	A	243	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 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	LYS	A	407	2	5,9,9	0.68	0	4,10,10	0.34	0
2	PLP	A	406	3	15,15,16	2.44	5 (33%)	20,22,23	1.94	7 (35%)
4	GOL	A	1003	-	5,5,5	0.18	0	5,5,5	0.31	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
3	LYS	A	407	2	-	0/5/9/9	-
2	PLP	A	406	3	-	2/6/6/8	0/1/1/1
4	GOL	A	1003	-	-	0/4/4/4	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	406	PLP	C3-C2	4.72	1.45	1.40
2	A	406	PLP	P-O4P	-4.15	1.46	1.60
2	A	406	PLP	C2-N1	4.06	1.41	1.33
2	A	406	PLP	C6-C5	3.58	1.45	1.37
2	A	406	PLP	C6-N1	2.95	1.40	1.34

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	406	PLP	O4P-P-O1P	4.59	119.34	106.47
2	A	406	PLP	C5A-C5-C6	-3.41	113.77	119.37
2	A	406	PLP	C4A-C4-C3	-3.20	115.07	120.50
2	A	406	PLP	C5-C6-N1	-2.55	119.57	123.82
2	A	406	PLP	C4A-C4-C5	2.14	123.14	120.94

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	406	PLP	C5A-O4P-P-O3P
2	A	406	PLP	C5A-O4P-P-O2P

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	407	LYS	1	0
2	A	406	PLP	1	0
4	A	1003	GOL	2	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](#)

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, 95th 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 ²)	Q<0.9
1	A	394/425 (92%)	6.40	360 (91%) 0 0	13, 25, 45, 57	0

The worst 5 of 360 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	196	ILE	23.9
1	A	146	PRO	23.5
1	A	147	TYR	23.3
1	A	86	VAL	20.1
1	A	394	LEU	19.6

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 monosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	GOL	A	1003	6/6	0.01	0.71	36,36,38,42	0
3	LYS	A	407	10/10	0.35	0.69	18,20,23,24	0
2	PLP	A	406	15/16	0.42	0.42	16,17,19,21	0

6.5 Other polymers [i](#)

There are no such residues in this entry.