wwPDB X-ray Structure Validation Summary Report (i)

Dec 17, 2023 – 06:23 PM EST

PDB ID	:	1GHD
Title	:	Crystal structure of the glutaryl-7-aminocephalosporanic acid acylase by mad phasing
Authors	:	Ding, Y.; Jiang, W.; Mao, X.; He, H.; Zhang, S.; Tang, H.; Bartlam, M.; Ye, S.; Jiang, F.; Liu, Y.; Zhao, G.; Rao, Z.
Deposited on	:	2000-12-07
Resolution	:	2.40 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 (i)) were used in the production of this report:

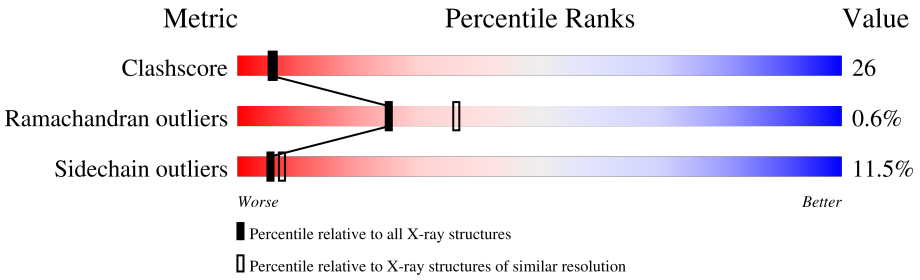
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 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))
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 >=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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	171	62% 23% • • 11%
2	B	522	59% 31% 8% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5504 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 GLUTARYL-7-AMINOCEPHALOSPORANIC ACID ACYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	153	1201	764	212	224	1	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	147	MSE	MET	modified residue	UNP O86089

- Molecule 2 is a protein called GLUTARYL-7-AMINOCEPHALOSPORANIC ACID ACYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
2	B	520	4105	2595	724	775	11	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	64	MSE	MET	modified residue	UNP O86089
B	73	MSE	MET	modified residue	UNP O86089
B	149	MSE	MET	modified residue	UNP O86089
B	156	MSE	MET	modified residue	UNP O86089
B	172	MSE	MET	modified residue	UNP O86089
B	282	MSE	MET	modified residue	UNP O86089
B	294	MSE	MET	modified residue	UNP O86089
B	304	MSE	MET	modified residue	UNP O86089
B	416	MSE	MET	modified residue	UNP O86089
B	460	MSE	MET	modified residue	UNP O86089
B	473	MSE	MET	modified residue	UNP O86089

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	47	Total O 47 47	0	0
3	B	151	Total O 151 151	0	0

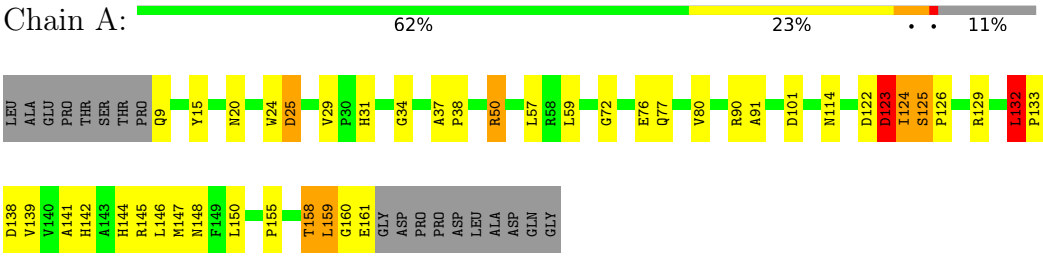
3 Residue-property plots (i)

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.

Note EDS was not executed.

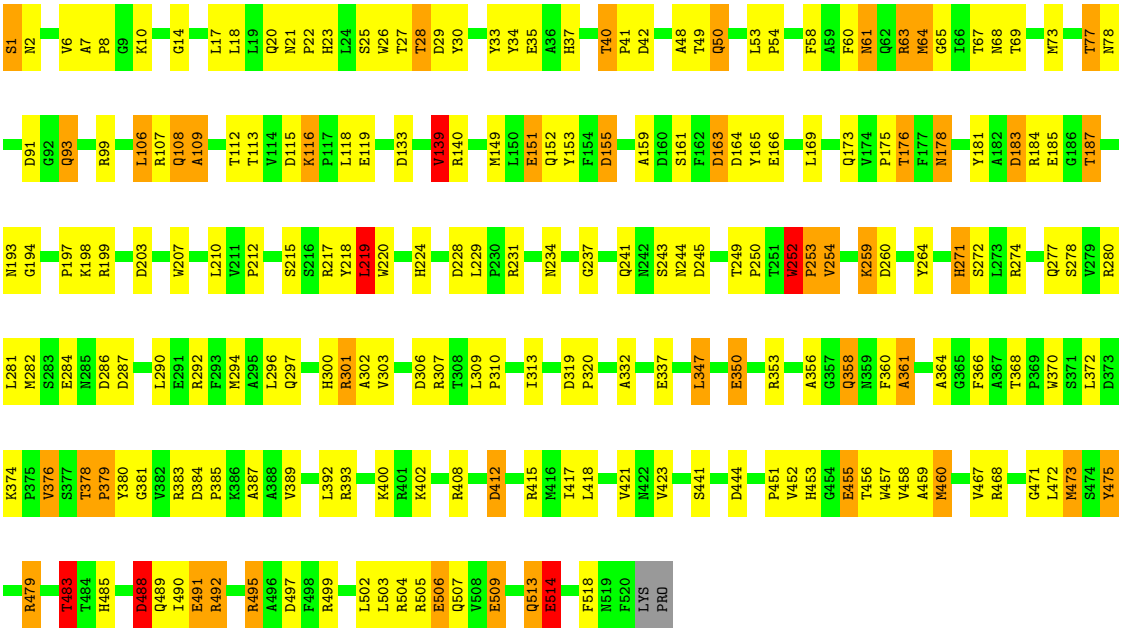
• Molecule 1: GLUTARYL-7-AMINOCEPHALOSPORANIC ACID ACYLASE

Chain A:



• Molecule 2: GLUTARYL-7-AMINOCEPHALOSPORANIC ACID ACYLASE

Chain B:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	73.48Å 73.48Å 382.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.40	Depositor
% Data completeness (in resolution range)	97.8 (30.00-2.40)	Depositor
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.218 , 0.265	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5504	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	0.92	3/1239 (0.2%)	1.09	8/1694 (0.5%)
2	B	0.96	11/4208 (0.3%)	1.12	31/5727 (0.5%)
All	All	0.95	14/5447 (0.3%)	1.11	39/7421 (0.5%)

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	514	GLU	CG-CD	15.97	1.75	1.51
2	B	514	GLU	CB-CG	15.20	1.81	1.52
1	A	123	ASP	CB-CG	11.22	1.75	1.51
2	B	509	GLU	CG-CD	10.54	1.67	1.51
2	B	509	GLU	CB-CG	10.52	1.72	1.52

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	514	GLU	OE1-CD-OE2	-21.37	97.66	123.30
1	A	123	ASP	CB-CG-OD2	16.25	132.92	118.30
1	A	123	ASP	OD1-CG-OD2	-12.78	99.02	123.30
2	B	155	ASP	CB-CG-OD2	10.67	127.90	118.30
2	B	203	ASP	CB-CG-OD1	-9.14	110.08	118.30

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	A	1201	0	1118	49	0
2	B	4105	0	3936	229	0
3	A	47	0	0	2	0
3	B	151	0	0	8	0
All	All	5504	0	5054	265	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 26.

The worst 5 of 265 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:514:GLU:CG	2:B:514:GLU:CB	1.81	1.54
2:B:514:GLU:CG	2:B:514:GLU:CD	1.75	1.54
1:A:123:ASP:CB	1:A:123:ASP:CG	1.75	1.50
2:B:485:HIS:HE1	2:B:514:GLU:OE1	1.30	1.13
2:B:63:ARG:NH2	2:B:64:MSE:HE3	1.69	1.06

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	151/171 (88%)	139 (92%)	10 (7%)	2 (1%)	12 17
2	B	518/522 (99%)	477 (92%)	39 (8%)	2 (0%)	34 48
All	All	669/693 (96%)	616 (92%)	49 (7%)	4 (1%)	25 36

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	132	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	109	ALA
2	B	514	GLU
1	A	123	ASP

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	117/130 (90%)	109 (93%)	8 (7%)	16 25
2	B	429/420 (102%)	374 (87%)	55 (13%)	4 5
All	All	546/550 (99%)	483 (88%)	63 (12%)	5 7

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	187	THR
2	B	473	MSE
2	B	286	ASP
2	B	468	ARG
2	B	488	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	108	GLN
2	B	224	HIS
2	B	189	ASN
2	B	244	ASN
1	A	142	HIS

5.3.3 RNA

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)

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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.