



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 27, 2020 – 01:32 PM BST

PDB ID : 6KJ8  
Title : E. coli ATCase holoenzyme mutant - G166P (catalytic chain)  
Authors : Lei, Z.; Zheng, J.; Jia, Z.C.  
Deposited on : 2019-07-22  
Resolution : 3.01 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	:	2.13
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)	:	2.13

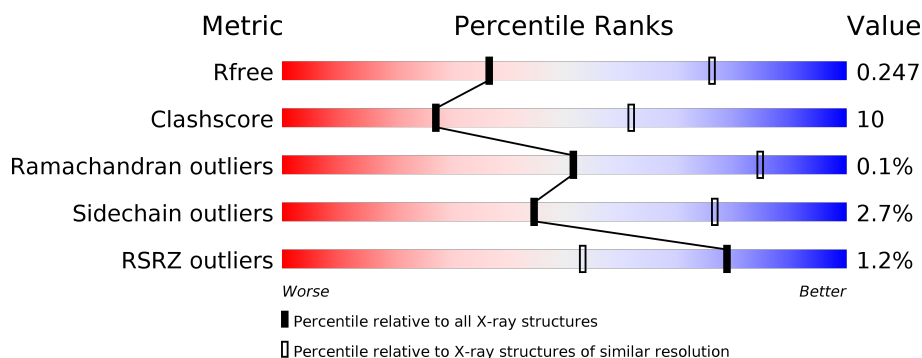
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.01 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	2399 (3.04-3.00)
Clashscore	141614	2734 (3.04-3.00)
Ramachandran outliers	138981	2640 (3.04-3.00)
Sidechain outliers	138945	2643 (3.04-3.00)
RSRZ outliers	127900	2287 (3.04-3.00)

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	310	<div> <div>75%</div> <div>21%</div> <div>••</div> </div>
1	C	310	<div> <div>73%</div> <div>23%</div> <div>••</div> </div>
1	E	310	<div> <div>74%</div> <div>23%</div> <div>•</div> </div>
2	B	153	<div> <div>3%</div> <div>77%</div> <div>10%</div> <div>•</div> <div>12%</div> </div>
2	D	153	<div> <div>5%</div> <div>58%</div> <div>25%</div> <div>•</div> <div>16%</div> </div>
2	F	153	<div> <div>2%</div> <div>78%</div> <div>16%</div> <div>7%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10058 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 Aspartate carbamoyltransferase catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	301	Total	C	N	O	S	0	0	0
			2320	1470	401	440	9			
1	C	303	Total	C	N	O	S	0	0	0
			2359	1494	412	444	9			
1	E	301	Total	C	N	O	S	0	0	0
			2351	1491	411	440	9			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	166	PRO	GLY	engineered mutation	UNP P0A786
C	166	PRO	GLY	engineered mutation	UNP P0A786
E	166	PRO	GLY	engineered mutation	UNP P0A786

- Molecule 2 is a protein called Aspartate carbamoyltransferase regulatory chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	135	Total	C	N	O	S	0	0	0
			1000	633	174	189	4			
2	D	129	Total	C	N	O	S	0	0	0
			919	579	155	181	4			
2	F	143	Total	C	N	O	S	0	0	0
			1078	679	190	205	4			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		
3	D	1	Total	Zn	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	F	1	Total	Zn	0	0
			1	1		

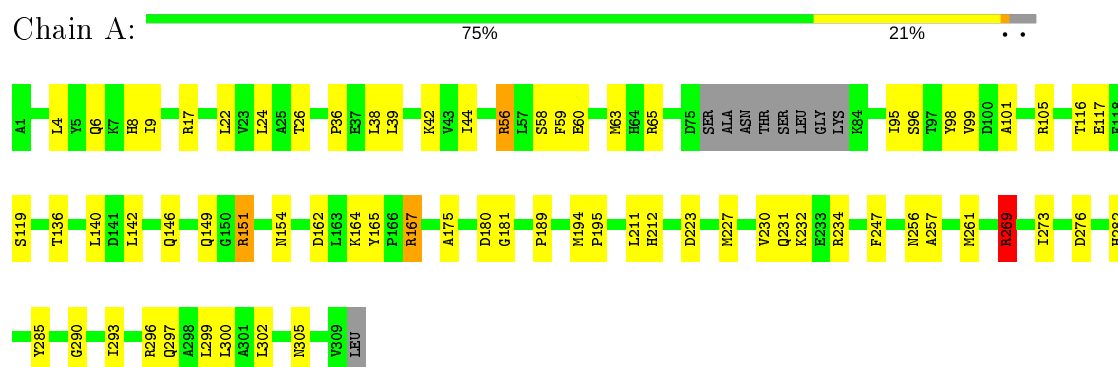
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	4	Total	O	0	0
			4	4		
4	B	5	Total	O	0	0
			5	5		
4	C	10	Total	O	0	0
			10	10		
4	D	1	Total	O	0	0
			1	1		
4	E	6	Total	O	0	0
			6	6		
4	F	2	Total	O	0	0
			2	2		

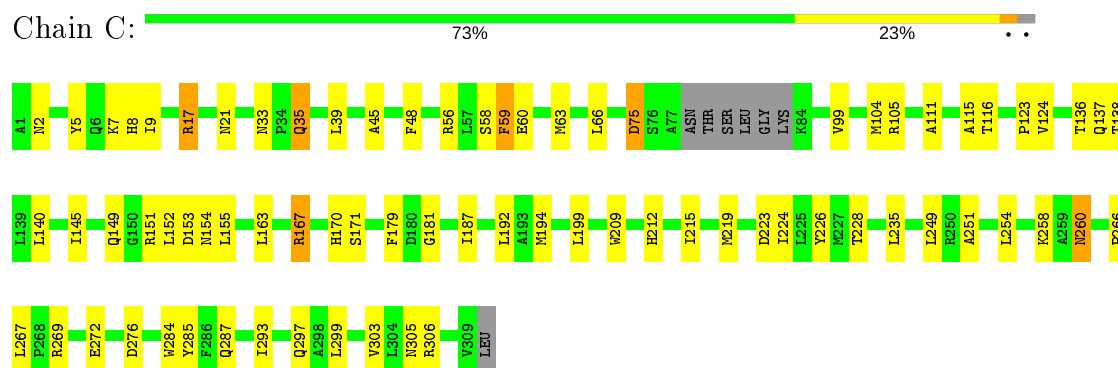
### 3 Residue-property plots

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.

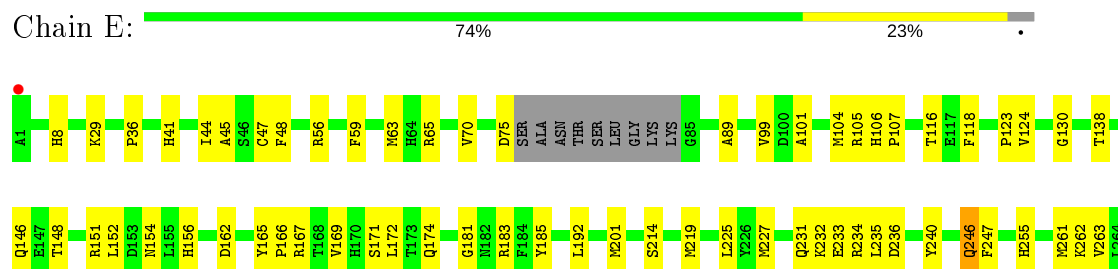
- Molecule 1: Aspartate carbamoyltransferase catalytic subunit



- Molecule 1: Aspartate carbamoyltransferase catalytic subunit

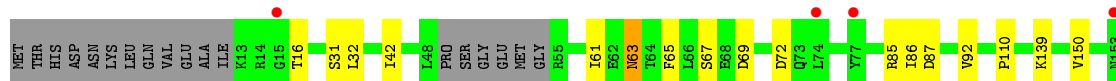
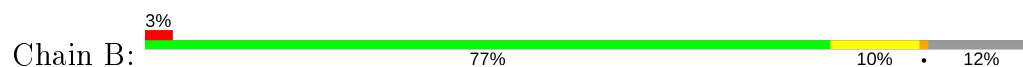


- Molecule 1: Aspartate carbamoyltransferase catalytic subunit

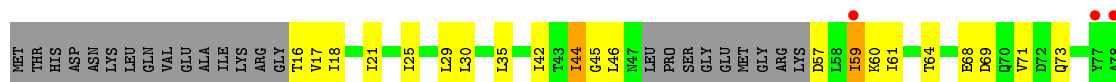




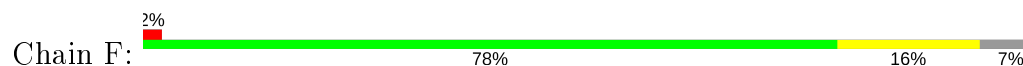
- Molecule 2: Aspartate carbamoyltransferase regulatory chain



- Molecule 2: Aspartate carbamoyltransferase regulatory chain



- Molecule 2: Aspartate carbamoyltransferase regulatory chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	126.59Å 126.59Å 196.49Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.87 – 3.01 47.87 – 3.01	Depositor EDS
% Data completeness (in resolution range)	99.6 (47.87-3.01) 99.6 (47.87-3.01)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.30 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, $R_{free}$	0.211 , 0.247 0.210 , 0.247	Depositor DCC
$R_{free}$ test set	1988 reflections (5.44%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.7	Xtriage
Anisotropy	0.240	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 41.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.018 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	10058	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 42.32 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.0750e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.59	0/2366	0.80	3/3222 (0.1%)
1	C	0.65	0/2405	0.77	2/3267 (0.1%)
1	E	0.61	0/2397	0.77	0/3256
2	B	0.50	0/1014	0.72	0/1378
2	D	0.52	1/932 (0.1%)	0.72	2/1273 (0.2%)
2	F	0.54	0/1094	0.70	0/1484
All	All	0.59	1/10208 (0.0%)	0.76	7/13880 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	141	CYS	CB-SG	-6.74	1.70	1.82

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	167	ARG	NE-CZ-NH2	-8.77	115.91	120.30
1	C	167	ARG	NE-CZ-NH1	6.08	123.34	120.30
1	A	167	ARG	NE-CZ-NH2	-6.08	117.26	120.30
2	D	44	ILE	CG1-CB-CG2	-6.01	98.17	111.40
2	D	59	ILE	CG1-CB-CG2	-5.92	98.38	111.40

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	2320	0	2278	44	1
1	C	2359	0	2352	52	0
1	E	2351	0	2353	52	0
2	B	1000	0	978	10	0
2	D	919	0	842	33	0
2	F	1078	0	1075	13	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
3	F	1	0	0	0	0
4	A	4	0	0	1	0
4	B	5	0	0	0	0
4	C	10	0	0	0	0
4	D	1	0	0	0	0
4	E	6	0	0	1	0
4	F	2	0	0	0	0
All	All	10058	0	9878	200	1

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 10.

The worst 5 of 200 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:E:262:LYS:HD2	1:E:284:TRP:CD1	1.92	1.03
2:D:18:ILE:HG12	2:D:83:VAL:HG22	1.40	0.99
1:E:231:GLN:OE1	1:E:234:ARG:NH1	2.06	0.88
2:D:18:ILE:HD12	2:D:59:ILE:HD11	1.58	0.85
1:E:261:MET:HE1	1:E:263:VAL:HG23	1.59	0.83

All (1) 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:A:17:ARG:NH2	1:A:180:ASP:OD2[6_554]	2.13	0.07

## 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	297/310 (96%)	289 (97%)	8 (3%)	0	100	100
1	C	299/310 (96%)	291 (97%)	7 (2%)	1 (0%)	41	75
1	E	297/310 (96%)	288 (97%)	9 (3%)	0	100	100
2	B	131/153 (86%)	118 (90%)	13 (10%)	0	100	100
2	D	125/153 (82%)	114 (91%)	11 (9%)	0	100	100
2	F	141/153 (92%)	130 (92%)	11 (8%)	0	100	100
All	All	1290/1389 (93%)	1230 (95%)	59 (5%)	1 (0%)	51	85

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	75	ASP

### 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	247/262 (94%)	241 (98%)	6 (2%)	49	79
1	C	254/262 (97%)	246 (97%)	8 (3%)	40	74
1	E	254/262 (97%)	245 (96%)	9 (4%)	36	70
2	B	107/137 (78%)	104 (97%)	3 (3%)	43	76
2	D	92/137 (67%)	90 (98%)	2 (2%)	52	80
2	F	118/137 (86%)	117 (99%)	1 (1%)	81	93

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1072/1197 (90%)	1043 (97%)	29 (3%)	44 76

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	75	ASP
1	C	285	TYR
1	E	269	ARG
1	C	260	ASN
2	D	84	ASN

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

Mol	Chain	Res	Type
1	C	149	GLN
1	C	260	ASN

### 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 3 ligands modelled in this entry, 3 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

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 [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	301/310 (97%)	-0.50	0 <span>100</span> <span>100</span>	34, 51, 70, 104	0
1	C	303/310 (97%)	-0.47	0 <span>100</span> <span>100</span>	35, 47, 66, 94	0
1	E	301/310 (97%)	-0.46	1 (0%) <span>94</span> <span>83</span>	35, 49, 68, 107	0
2	B	135/153 (88%)	0.14	4 (2%) <span>50</span> <span>22</span>	42, 82, 123, 132	0
2	D	129/153 (84%)	0.25	8 (6%) <span>20</span> <span>6</span>	45, 107, 153, 159	0
2	F	143/153 (93%)	-0.10	3 (2%) <span>63</span> <span>34</span>	32, 73, 111, 122	0
All	All	1312/1389 (94%)	-0.30	16 (1%) <span>79</span> <span>53</span>	32, 53, 118, 159	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	80	GLN	3.8
2	D	94	LYS	3.7
2	D	78	ALA	3.2
2	D	153	ASN	3.0
2	B	74	LEU	2.8

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	ZN	D	201	1/1	0.99	0.16	56,56,56,56	0
3	ZN	F	201	1/1	0.99	0.15	39,39,39,39	0
3	ZN	B	201	1/1	0.99	0.14	47,47,47,47	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.