



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 03:48 pm BST

PDB ID : 5ILQ  
Title : Crystal structure of truncated unliganded Aspartate Transcarbamoylase from Plasmodium falciparum  
Authors : Lunev, S.; Bosch, S.S.; Batista, F.D.A.; Wrenger, C.; Groves, M.R.  
Deposited on : 2016-03-04  
Resolution : 2.50 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
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

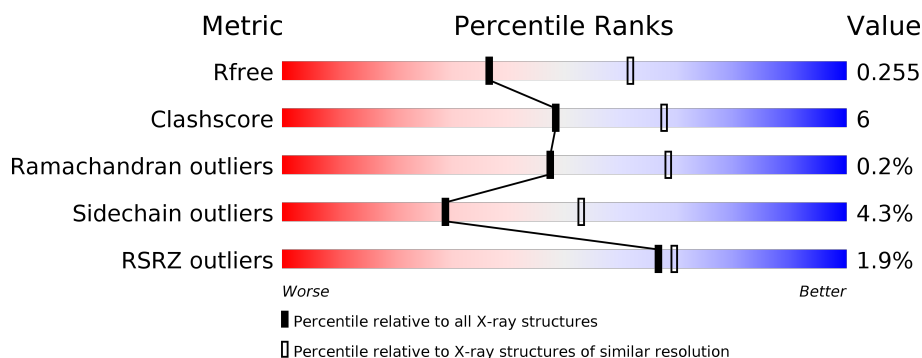
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	349	<div> <div>2%</div> <div> <div></div> <div>77%</div> <div>10%</div> <div>•</div> <div>9%</div> </div> </div>
1	B	349	<div> <div>%</div> <div> <div></div> <div>77%</div> <div>12%</div> <div>•</div> <div>7%</div> </div> </div>
1	C	349	<div> <div>2%</div> <div> <div></div> <div>79%</div> <div>12%</div> <div>•</div> <div>7%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7907 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	318	Total	C	N	O	S	0	0	0
			2563	1637	422	496	8			
1	B	324	Total	C	N	O	S	0	0	0
			2612	1669	430	505	8			
1	C	325	Total	C	N	O	S	0	0	0
			2623	1678	431	506	8			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	376	SER	-	expression tag	UNP Q8IDP8
A	377	ALA	-	expression tag	UNP Q8IDP8
A	378	TRP	-	expression tag	UNP Q8IDP8
A	379	SER	-	expression tag	UNP Q8IDP8
A	380	HIS	-	expression tag	UNP Q8IDP8
A	381	PRO	-	expression tag	UNP Q8IDP8
A	382	GLN	-	expression tag	UNP Q8IDP8
A	383	PHE	-	expression tag	UNP Q8IDP8
A	384	GLU	-	expression tag	UNP Q8IDP8
A	385	LYS	-	expression tag	UNP Q8IDP8
B	376	SER	-	expression tag	UNP Q8IDP8
B	377	ALA	-	expression tag	UNP Q8IDP8
B	378	TRP	-	expression tag	UNP Q8IDP8
B	379	SER	-	expression tag	UNP Q8IDP8
B	380	HIS	-	expression tag	UNP Q8IDP8
B	381	PRO	-	expression tag	UNP Q8IDP8
B	382	GLN	-	expression tag	UNP Q8IDP8
B	383	PHE	-	expression tag	UNP Q8IDP8
B	384	GLU	-	expression tag	UNP Q8IDP8
B	385	LYS	-	expression tag	UNP Q8IDP8
C	376	SER	-	expression tag	UNP Q8IDP8
C	377	ALA	-	expression tag	UNP Q8IDP8
C	378	TRP	-	expression tag	UNP Q8IDP8

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	379	SER	-	expression tag	UNP Q8IDP8
C	380	HIS	-	expression tag	UNP Q8IDP8
C	381	PRO	-	expression tag	UNP Q8IDP8
C	382	GLN	-	expression tag	UNP Q8IDP8
C	383	PHE	-	expression tag	UNP Q8IDP8
C	384	GLU	-	expression tag	UNP Q8IDP8
C	385	LYS	-	expression tag	UNP Q8IDP8

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	23	Total	O	0	0
			23	23		
4	B	38	Total	O	0	0
			38	38		
4	C	25	Total	O	0	0
			25	25		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.98Å 103.80Å 87.11Å 90.00° 117.68° 90.00°	Depositor
Resolution (Å)	45.03 – 2.50 20.00 – 2.50	Depositor EDS
% Data completeness (in resolution range)	95.1 (45.03-2.50) 95.3 (20.00-2.50)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.24 (at 2.50Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.208 , 0.250 0.211 , 0.255	Depositor DCC
$R_{free}$ test set	2354 reflections (5.21%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.6	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 35.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.044 for -h-l,k,h 0.044 for l,k,-h-l 0.048 for h,-k,-h-l 0.049 for -h-l,-k,l 0.078 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7907	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.62% 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

### 5.1 Standard geometry

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

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	0.93	0/2609	1.03	13/3524 (0.4%)
1	B	0.97	2/2661 (0.1%)	1.04	11/3597 (0.3%)
1	C	0.91	0/2673	1.00	9/3613 (0.2%)
All	All	0.93	2/7943 (0.0%)	1.03	33/10734 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	372	SER	CB-OG	-5.17	1.35	1.42
1	B	282	GLU	CG-CD	5.07	1.59	1.51

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	126	LEU	CA-CB-CG	12.11	143.16	115.30
1	C	126	LEU	CA-CB-CG	11.73	142.28	115.30
1	A	111	ARG	NE-CZ-NH1	-9.51	115.55	120.30
1	C	237	ARG	NE-CZ-NH1	8.48	124.54	120.30
1	B	237	ARG	NE-CZ-NH1	8.03	124.32	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:



Mol	Chain	Res	Type	Group
1	C	375	SER	Peptide

## 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	2563	0	2570	31	0
1	B	2612	0	2620	41	0
1	C	2623	0	2629	26	0
2	A	12	0	16	4	0
2	B	6	0	8	3	0
3	C	5	0	0	0	0
4	A	23	0	0	0	0
4	B	38	0	0	1	0
4	C	25	0	0	1	0
All	All	7907	0	7843	94	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 6.

The worst 5 of 94 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:401:GOL:O1	4:B:501:HOH:O	1.58	1.19
1:B:125:VAL:O	2:B:401:GOL:H31	1.81	0.81
1:B:146:PHE:O	1:B:150:SER:HB3	1.85	0.77
1:A:67:ASP:HB2	2:A:402:GOL:H11	1.67	0.76
1:A:146:PHE:O	1:A:150:SER:HB3	1.87	0.75

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	312/349 (89%)	300 (96%)	12 (4%)	0	100	100
1	B	320/349 (92%)	309 (97%)	10 (3%)	1 (0%)	41	61
1	C	321/349 (92%)	309 (96%)	11 (3%)	1 (0%)	41	61
All	All	953/1047 (91%)	918 (96%)	33 (4%)	2 (0%)	47	68

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	338	ASN
1	B	337	VAL

### 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	298/328 (91%)	285 (96%)	13 (4%)	28	52
1	B	303/328 (92%)	289 (95%)	14 (5%)	27	50
1	C	304/328 (93%)	292 (96%)	12 (4%)	32	57
All	All	905/984 (92%)	866 (96%)	39 (4%)	29	53

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	163	LYS
1	B	227	THR
1	C	352	TYR
1	B	164	LYS
1	B	172	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	243	ASN
1	B	289	HIS
1	C	224	ASN
1	B	165	ASN
1	B	179	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 carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

4 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	C	401	-	4,4,4	0.62	0	6,6,6	0.82	0
2	GOL	A	401	-	5,5,5	1.03	0	5,5,5	1.21	1 (20%)
2	GOL	B	401	-	5,5,5	0.66	0	5,5,5	1.64	1 (20%)
2	GOL	A	402	-	5,5,5	0.89	0	5,5,5	1.38	1 (20%)

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	GOL	A	401	-	-	0/4/4/4	-
2	GOL	B	401	-	-	2/4/4/4	-
2	GOL	A	402	-	-	4/4/4/4	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	401	GOL	O3-C3-C2	3.21	125.57	110.20
2	A	402	GOL	O1-C1-C2	2.66	122.98	110.20
2	A	401	GOL	C3-C2-C1	-2.14	103.40	111.70

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	401	GOL	C1-C2-C3-O3
2	B	401	GOL	O2-C2-C3-O3
2	A	402	GOL	O1-C1-C2-C3
2	A	402	GOL	O2-C2-C3-O3
2	A	402	GOL	O1-C1-C2-O2

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	GOL	1	0
2	B	401	GOL	3	0
2	A	402	GOL	3	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, 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	318/349 (91%)	-0.05	7 (2%) 62 65	37, 62, 107, 143	0
1	B	324/349 (92%)	-0.16	4 (1%) 79 80	34, 57, 103, 137	0
1	C	325/349 (93%)	-0.08	7 (2%) 62 65	37, 66, 104, 124	0
All	All	967/1047 (92%)	-0.10	18 (1%) 66 69	34, 62, 105, 143	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	138	LYS	6.5
1	C	139	GLY	4.3
1	B	378	TRP	3.8
1	C	323	ASN	3.4
1	A	135	SER	3.3

### 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 carbohydrates 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( $\text{\AA}^2$ )	Q<0.9
2	GOL	A	401	6/6	0.84	0.15	51,70,74,76	0
2	GOL	B	401	6/6	0.89	0.19	42,53,59,59	0
2	GOL	A	402	6/6	0.89	0.12	52,59,66,68	0
3	SO4	C	401	5/5	0.99	0.14	42,43,56,64	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.