



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 27, 2020 – 10:43 pm BST

PDB ID : 3AQL  
Title : Structure of bacterial protein (apo form II)  
Authors : Toh, Y.; Takeshita, D.; Tomita , K.  
Deposited on : 2010-11-09  
Resolution : 3.00 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](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

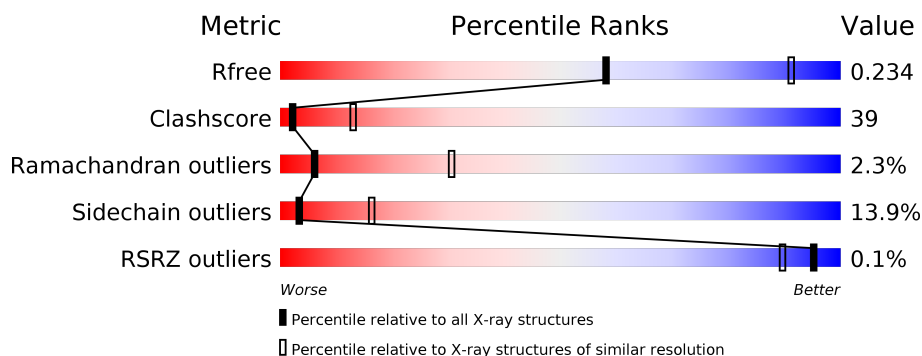
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	415	<div> <div>37%</div> <div>48%</div> <div>8%</div> <div>7%</div> </div>
1	B	415	<div> <div>38%</div> <div>46%</div> <div>10%</div> <div>7%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6342 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 Poly(A) polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	388	Total	C	N	O	S	0	0	0
			3164	2022	568	560	14			
1	B	388	Total	C	N	O	S	0	0	0
			3164	2022	568	560	14			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	234	HIS	ARG	ENGINEERED MUTATION	UNP C9QS13
B	234	HIS	ARG	ENGINEERED MUTATION	UNP C9QS13

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		

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

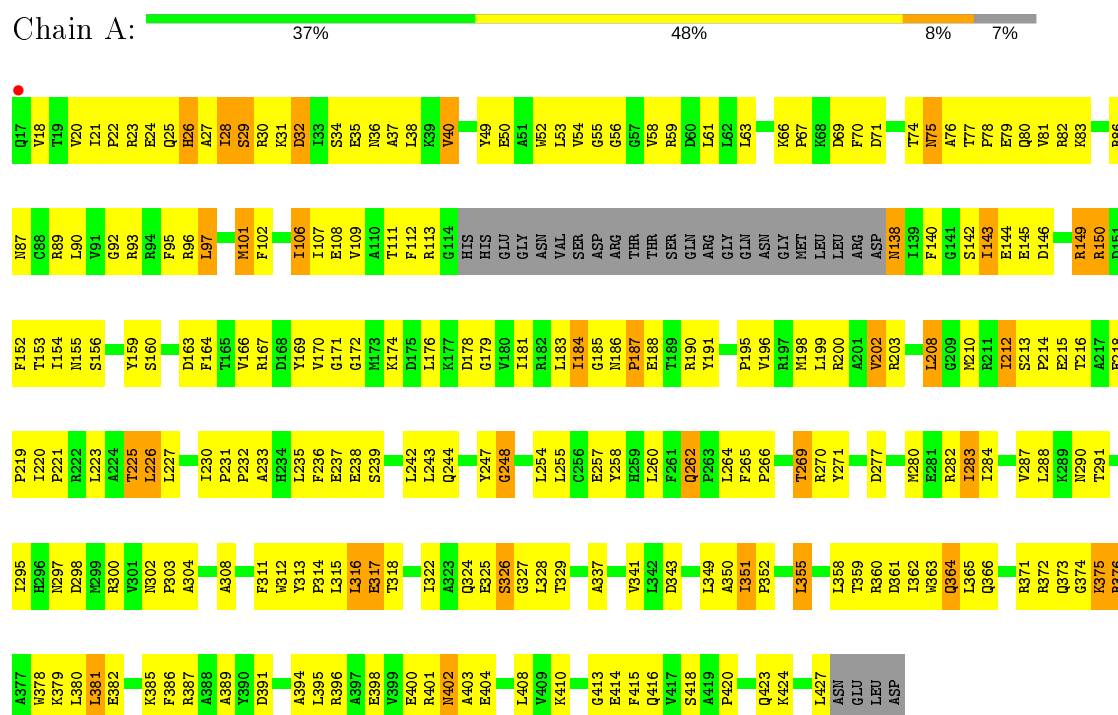


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: Poly(A) polymerase



L243	Q244	Y247	G248	Y249	K253	L254	L255	C256	E257	L260	F261	Q262	P263	L264	F265	T269	R270	Y271	N275	G276	D277	S278	P279	M280	E281	R282	L283	I284	E285	T291	L295	D298	M299	R300	V301	N302	P303	A304	F305	L306	F307	A308	A309	M310	F311	W312	Y313	P314	L315	L316				
K321	I322	A323	Q324	E325	S326	G327	L328	T329	F334	A335	L336	A337	M338	N339	D340	V341	L342	D343	E344	A345	C346	R347	S348	L349	A350	I351	P352	K353	R354	L355	L358	T359	R360	D361	I362	W363	Q364	L365	Q366	M369	S370	R371	R372	Q373	G374	K375	R376	A377	W378	K379	L380	L381	E382	K385
A389	T390	D391	A394	L395	R396	A397	R401	N402	A403	E404	L405	Q406	R407	L408	V409	K410	W411	W412	G413	E414	F415	Q416	V417	S418	P421	L427	ASN	GLU	LEU	ASP																								

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	133.01Å 133.01Å 176.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.55 – 3.00 49.36 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.1 (29.55-3.00) 98.6 (49.36-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.73 (at 2.81Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, $R_{free}$	0.237 , 0.251 0.223 , 0.234	Depositor DCC
$R_{free}$ test set	1958 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	69.7	Xtriage
Anisotropy	0.435	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 43.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6342	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.94% 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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.47	0/3234	0.67	1/4371 (0.0%)
1	B	0.46	0/3234	0.65	1/4371 (0.0%)
All	All	0.46	0/6468	0.66	2/8742 (0.0%)

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	A	0	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	29	SER	N-CA-C	-8.07	89.22	111.00
1	B	93	ARG	N-CA-C	5.53	125.92	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	36	ASN	Peptide
1	B	36	ASN	Peptide
1	B	92	GLY	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	3164	0	3197	259	0
1	B	3164	0	3197	248	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	6	0	8	1	0
3	B	6	0	8	0	0
All	All	6342	0	6410	500	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 39.

The worst 5 of 500 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:B:93:ARG:CG	1:B:94:ARG:H	1.22	1.42
1:A:183:LEU:HD23	1:A:184:ILE:N	1.48	1.28
1:A:372:ARG:NH1	1:A:372:ARG:HA	1.52	1.22
1:B:93:ARG:CG	1:B:94:ARG:N	1.93	1.16
1:B:93:ARG:HG2	1:B:94:ARG:N	1.54	1.13

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	384/415 (92%)	329 (86%)	47 (12%)	8 (2%)	7 33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	384/415 (92%)	335 (87%)	39 (10%)	10 (3%)	5	27
All	All	768/830 (92%)	664 (86%)	86 (11%)	18 (2%)	6	30

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	248	GLY
1	A	327	GLY
1	B	28	ILE
1	B	97	LEU
1	B	104	PRO

### 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	334/358 (93%)	290 (87%)	44 (13%)	4	18
1	B	334/358 (93%)	285 (85%)	49 (15%)	3	15
All	All	668/716 (93%)	575 (86%)	93 (14%)	3	16

5 of 93 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	387	ARG
1	B	84	LEU
1	B	346	CYS
1	A	402	ASN
1	B	38	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	402	ASN
1	B	26	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	402	ASN
1	A	406	GLN
1	B	17	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

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	GOL	B	601	-	5,5,5	0.40	0	5,5,5	0.52	0
3	GOL	A	602	-	5,5,5	0.40	0	5,5,5	0.35	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	GOL	B	601	-	-	2/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	602	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	601	GOL	O1-C1-C2-C3
3	A	602	GOL	C1-C2-C3-O3
3	A	602	GOL	O2-C2-C3-O3
3	B	601	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	602	GOL	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 [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	388/415 (93%)	-0.29	1 (0%) 94 84	60, 92, 126, 145	0
1	B	388/415 (93%)	-0.35	0 100 100	64, 94, 126, 154	0
All	All	776/830 (93%)	-0.32	1 (0%) 95 89	60, 92, 126, 154	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	17	GLN	2.0

### 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(Å <sup>2</sup> )	Q<0.9
3	GOL	A	602	6/6	0.85	0.20	68,85,94,95	0
2	MG	B	501	1/1	0.87	0.12	83,83,83,83	0
3	GOL	B	601	6/6	0.93	0.22	96,103,106,108	0
2	MG	A	500	1/1	0.94	0.27	80,80,80,80	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.