



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 7, 2016 – 02:42 PM EDT

PDB ID : 5KUD  
Title : Crystal structure of full length Cry6Aa  
Authors : Kelker, M.S.; Xu, X.; Lee, M.; Chan, M.; Hung, S.; Dementiev, K.; Chikwana, V.M.; Hey, T.; Narva, K.  
Deposited on : 2016-07-13  
Resolution : 2.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

<http://wwpdb.org/validation/2016/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	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

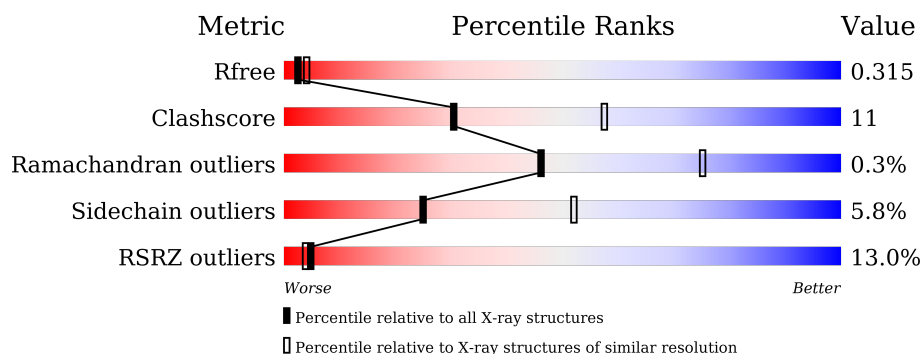
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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. 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	475	<div> <div>11%</div> <div>62%</div> <div>19%</div> <div>•</div> <div>17%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3027 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 Pesticidal crystal protein Cry6Aa.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	393	Total	C	N	O	S	0	0	0
			3006	1899	498	603	6			

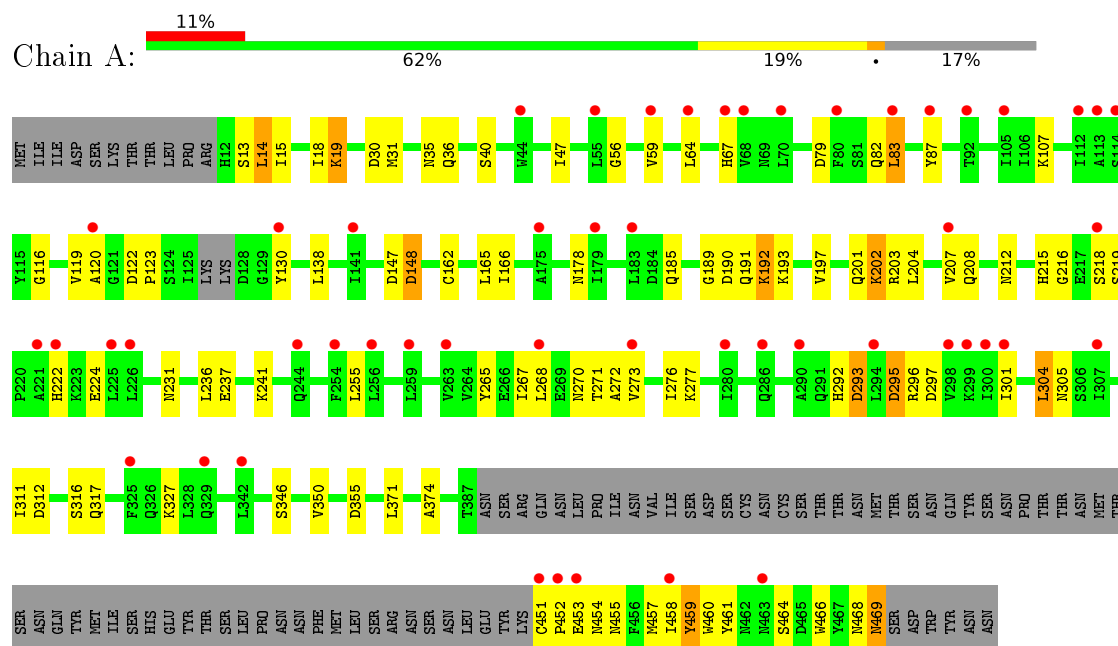
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	21	Total	O	0	0
			21	21		

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Pesticidal crystal protein Cry6Aa



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	50.37Å 71.73Å 142.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.14 – 2.70 47.50 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.4 (29.14-2.70) 89.0 (47.50-2.70)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.80 (at 2.69Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.274 , 0.325 0.270 , 0.315	Depositor DCC
$R_{free}$ test set	1307 reflections (9.93%)	DCC
Wilson B-factor (Å <sup>2</sup> )	54.0	Xtriage
Anisotropy	0.717	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 51.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	3027	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.05% 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](#)

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	0.64	0/3058	0.79	6/4161 (0.1%)

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	371	LEU	CB-CG-CD2	-7.80	97.73	111.00
1	A	83	LEU	CA-CB-CG	6.95	131.29	115.30
1	A	120	ALA	C-N-CA	-6.42	108.82	122.30
1	A	147	ASP	CB-CG-OD1	5.26	123.03	118.30
1	A	138	LEU	CB-CG-CD2	-5.10	102.33	111.00
1	A	31	MET	CG-SD-CE	-5.01	92.19	100.20

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	3006	0	2809	64	0
2	A	21	0	0	3	0
All	All	3027	0	2809	64	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 11.

All (64) 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:455:ASN:O	1:A:458:ILE:HG22	1.30	1.30
1:A:191:GLN:HA	1:A:191:GLN:HE21	1.31	0.94
1:A:455:ASN:O	1:A:458:ILE:CG2	2.19	0.91
1:A:36:GLN:NE2	1:A:40:SER:OG	2.08	0.87
1:A:346:SER:O	1:A:350:VAL:HG23	1.74	0.86
1:A:191:GLN:HA	1:A:191:GLN:NE2	1.92	0.83
1:A:208:GLN:HE21	1:A:305:ASN:HD22	1.25	0.82
1:A:469:ASN:ND2	1:A:469:ASN:N	2.30	0.76
1:A:236:LEU:O	1:A:236:LEU:HD12	1.88	0.73
1:A:469:ASN:HD22	1:A:469:ASN:N	1.85	0.71
1:A:59:VAL:HG12	1:A:59:VAL:O	1.92	0.69
1:A:208:GLN:HE21	1:A:305:ASN:ND2	1.91	0.68
1:A:59:VAL:CG1	1:A:59:VAL:O	2.42	0.68
1:A:241:LYS:NZ	2:A:501:HOH:O	2.18	0.66
1:A:270:ASN:C	1:A:270:ASN:OD1	2.33	0.65
1:A:270:ASN:OD1	1:A:271:THR:N	2.32	0.62
1:A:201:GLN:NE2	1:A:312:ASP:OD1	2.33	0.60
1:A:67:HIS:ND1	1:A:461:TYR:OH	2.33	0.58
1:A:107:LYS:NZ	2:A:502:HOH:O	2.37	0.57
1:A:216:GLY:O	1:A:222:HIS:HB3	2.05	0.57
1:A:255:LEU:HD23	1:A:277:LYS:HE3	1.86	0.57
1:A:178:ASN:HB2	2:A:511:HOH:O	2.03	0.56
1:A:204:LEU:O	1:A:207:VAL:HG12	2.06	0.56
1:A:350:VAL:HG13	1:A:355:ASP:O	2.06	0.55
1:A:295:ASP:OD1	1:A:295:ASP:C	2.47	0.53
1:A:237:GLU:HB3	1:A:241:LYS:NZ	2.24	0.53
1:A:197:VAL:HG13	1:A:311:ILE:HD11	1.90	0.53
1:A:273:VAL:O	1:A:276:ILE:HB	2.10	0.52
1:A:219:SER:O	1:A:222:HIS:HB2	2.09	0.52
1:A:270:ASN:OD1	1:A:272:ALA:N	2.41	0.51
1:A:148:ASP:OD2	1:A:148:ASP:N	2.40	0.51
1:A:56:GLY:HA3	1:A:457:MET:SD	2.51	0.50
1:A:293:ASP:O	1:A:297:ASP:OD2	2.30	0.50
1:A:30:ASP:O	1:A:327:LYS:HE2	2.11	0.50
1:A:83:LEU:HD13	1:A:87:TYR:CE2	2.48	0.49
1:A:453:GLU:CB	1:A:454:ASN:HA	2.42	0.48
1:A:191:GLN:CA	1:A:191:GLN:HE21	2.01	0.48
1:A:185:GLN:O	1:A:189:GLY:N	2.45	0.48
1:A:15:ILE:HG22	1:A:19:LYS:HZ2	1.79	0.48
1:A:236:LEU:C	1:A:236:LEU:HD12	2.35	0.47
1:A:162:CYS:O	1:A:166:ILE:HG13	2.15	0.46

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:218:SER:HB2	1:A:222:HIS:CE1	2.51	0.45
1:A:116:GLY:HA3	1:A:374:ALA:HB2	1.98	0.45
1:A:192:LYS:HB3	1:A:193:LYS:H	1.37	0.45
1:A:459:TYR:HE1	1:A:460:TRP:CZ3	2.34	0.45
1:A:453:GLU:CB	1:A:454:ASN:CA	2.95	0.44
1:A:469:ASN:H	1:A:469:ASN:HD22	1.60	0.44
1:A:47:ILE:HD11	1:A:317:GLN:CB	2.47	0.44
1:A:464:SER:C	1:A:466:TRP:N	2.71	0.44
1:A:202:LYS:CG	1:A:203:ARG:N	2.80	0.44
1:A:64:LEU:HD22	1:A:83:LEU:HD11	2.00	0.44
1:A:130:TYR:HD1	1:A:165:LEU:HD21	1.83	0.43
1:A:468:ASN:C	1:A:469:ASN:ND2	2.72	0.42
1:A:119:VAL:HA	1:A:130:TYR:HB2	2.01	0.42
1:A:122:ASP:HA	1:A:123:PRO:HD3	1.81	0.42
1:A:215:HIS:HB2	1:A:301:ILE:HD11	2.02	0.41
1:A:208:GLN:HA	1:A:304:LEU:HD21	2.02	0.41
1:A:265:TYR:O	1:A:268:LEU:HB2	2.21	0.41
1:A:267:ILE:HG22	1:A:273:VAL:HG21	2.02	0.41
1:A:295:ASP:OD1	1:A:296:ARG:N	2.54	0.41
1:A:190:ASP:O	1:A:191:GLN:HB2	2.20	0.41
1:A:79:ASP:OD2	1:A:215:HIS:NE2	2.51	0.41
1:A:212:ASN:OD1	1:A:216:GLY:HA3	2.21	0.40
1:A:14:LEU:O	1:A:18:ILE:HG13	2.22	0.40

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	387/475 (82%)	377 (97%)	9 (2%)	1 (0%)	46 75

All (1) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	A	452	PRO

### 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	310/428 (72%)	292 (94%)	18 (6%)	25 52

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	13	SER
1	A	14	LEU
1	A	19	LYS
1	A	35	ASN
1	A	82	GLN
1	A	148	ASP
1	A	192	LYS
1	A	202	LYS
1	A	224	GLU
1	A	231	ASN
1	A	292	HIS
1	A	293	ASP
1	A	295	ASP
1	A	304	LEU
1	A	316	SER
1	A	451	CYS
1	A	459	TYR
1	A	469	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	23	ASN
1	A	35	ASN
1	A	36	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	188	HIS
1	A	191	GLN
1	A	305	ASN
1	A	326	GLN
1	A	469	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](#)

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 ⓘ

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	393/475 (82%)	0.79	51 (12%) 5 4	46, 78, 130, 182	0

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	298	VAL	7.3
1	A	83	LEU	7.3
1	A	226	LEU	5.3
1	A	70	LEU	5.3
1	A	80	PHE	4.8
1	A	105	ILE	4.5
1	A	280	ILE	4.3
1	A	300	ILE	4.1
1	A	307	ILE	3.8
1	A	452	PRO	3.6
1	A	112	ILE	3.5
1	A	225	LEU	3.5
1	A	244	GLN	3.4
1	A	87	TYR	3.4
1	A	286	GLN	3.3
1	A	453	GLU	3.2
1	A	463	ASN	3.2
1	A	254	PHE	3.1
1	A	221	ALA	3.0
1	A	92	THR	2.9
1	A	183	LEU	2.9
1	A	268	LEU	2.8
1	A	55	LEU	2.8
1	A	68	VAL	2.7
1	A	59	VAL	2.6
1	A	458	ILE	2.6
1	A	175	ALA	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	218	SER	2.6
1	A	259	LEU	2.5
1	A	325	PHE	2.5
1	A	130	TYR	2.5
1	A	273	VAL	2.5
1	A	290	ALA	2.5
1	A	67	HIS	2.5
1	A	301	ILE	2.5
1	A	294	LEU	2.5
1	A	113	ALA	2.4
1	A	207	VAL	2.3
1	A	222	HIS	2.2
1	A	342	LEU	2.2
1	A	64	LEU	2.2
1	A	256	LEU	2.2
1	A	141	ILE	2.2
1	A	299	LYS	2.2
1	A	120	ALA	2.2
1	A	451	CYS	2.2
1	A	114	SER	2.1
1	A	263	VAL	2.1
1	A	44	TRP	2.1
1	A	179	ILE	2.0
1	A	329	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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.