



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:53 AM GMT

PDB ID : 2J5S  
Title : STRUCTURAL OF ABDH, A BETA-DIKETONE HYDROLASE FROM THE CYANOBACTERIUM ANABAENA SP. PCC 7120 BOUND TO (S)-3-OXOCYCLOHEXYL ACETIC ACID  
Authors : Bennett, J.P.; Whittingham, J.L.; Brzozowski, A.M.; Leonard, P.M.; Grogan, G.  
Deposited on : 2006-09-19  
Resolution : 1.57 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

<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	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
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)	:	trunk26865

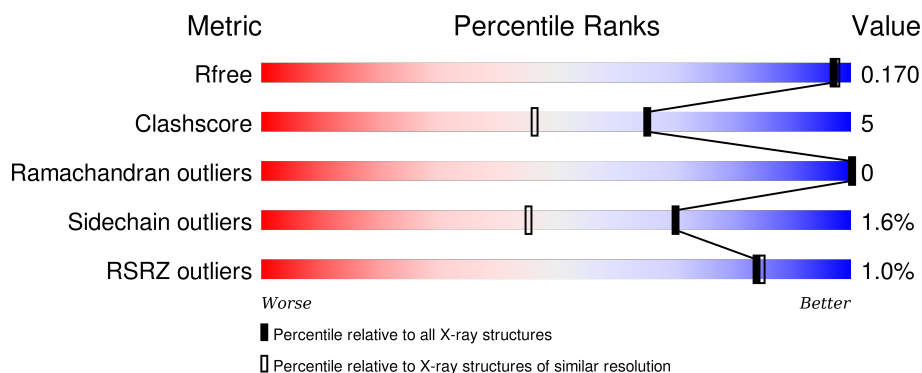
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.57 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	3815 (1.60-1.56)
Clashscore	102246	4131 (1.60-1.56)
Ramachandran outliers	100387	4021 (1.60-1.56)
Sidechain outliers	100360	4018 (1.60-1.56)
RSRZ outliers	91569	3823 (1.60-1.56)

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	263	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 88%, grey 6%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>88%</span> <span>6% • 5%</span> </div> </div>
1	B	263	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 10%, green 83%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>83%</span> <span>10% • 5%</span> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	KTA	A	1253	-	-	-	X
2	KTA	B	1253	-	-	-	X

## 2 Entry composition [i](#)

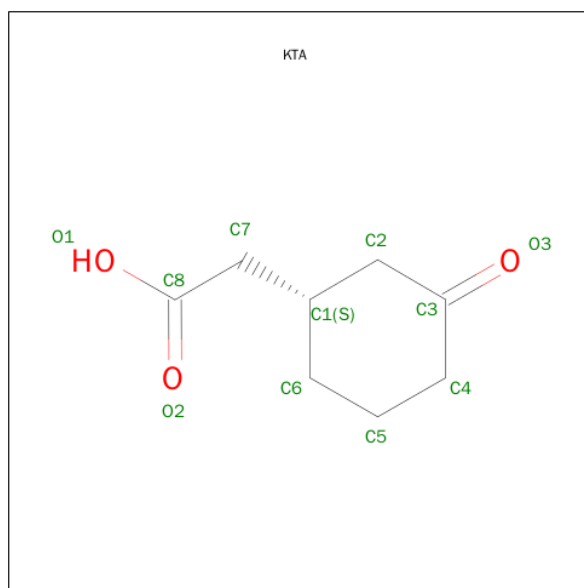
There are 4 unique types of molecules in this entry. The entry contains 4687 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 BETA-DIKETONE HYDROLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	251	Total	C	N	O	S	1	9	1
			2081	1328	363	385	5			
1	B	249	Total	C	N	O	S	7	13	1
			2092	1341	362	384	5			

- Molecule 2 is (S)-CYCLOHEXANONE-2-ACETATE (three-letter code: KTA) (formula:  $C_8H_{12}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			11	8	3		
2	B	1	Total	C	O	0	0
			11	8	3		

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	Ni 1	0	0
3	A	1	Total 1	Ni 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	247	Total 247	O 247	1	0
4	B	243	Total 243	O 243	0	0



- Molecule 1: BETA-DIKETONE HYDROLASE



V232	E13
N233	K41
E234	P60
G235	G66
	F77
R251	P78
N252	S79
T253	L80
	L103
	E107
	V108
	P109
	N115
	S134
	E135
	N136
	P143
	H156
	L161
	Y166
	L193
	P194
	Q195
	S196
	K197
	R218
	I232

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.39 Å 80.39 Å 125.88 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.19 – 1.57 46.69 – 1.57	Depositor EDS
% Data completeness (in resolution range)	98.8 (40.19-1.57) 98.9 (46.69-1.57)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.66 (at 1.57 Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.138 , 0.170 0.137 , 0.170	Depositor DCC
$R_{free}$ test set	3224 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	13.4	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 62.5	EDS
Estimated twinning fraction	0.043 for h,-h-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 63768 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	4687	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.57% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: NI, KTA

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	1.09	2/2136 (0.1%)	0.81	1/2905 (0.0%)
1	B	0.72	2/2164 (0.1%)	0.81	4/2942 (0.1%)
All	All	0.92	4/4300 (0.1%)	0.81	5/5847 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	13	GLU	CD-OE2	35.97	1.65	1.25
1	A	13	GLU	CG-CD	14.68	1.74	1.51
1	B	251	ARG	CZ-NH2	10.36	1.46	1.33
1	B	7	GLU	CB-CG	-5.99	1.40	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	197	LYS	CD-CE-NZ	-7.77	93.84	111.70
1	B	251	ARG	NE-CZ-NH2	-7.04	116.78	120.30
1	A	13	GLU	CB-CG-CD	-7.03	95.23	114.20
1	B	166	TYR	CA-CB-CG	-5.45	103.04	113.40
1	B	193	LEU	CA-CB-CG	5.02	126.85	115.30

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	2081	0	2066	18	0
1	B	2092	0	2087	24	0
2	A	11	0	11	0	0
2	B	11	0	11	3	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	247	0	0	6	0
4	B	243	0	0	4	0
All	All	4687	0	4175	42	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 5.

The worst 5 of 42 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:4:ASN:H	1:A:4:ASN:HD22	1.03	0.98
1:A:4:ASN:N	1:A:4:ASN:HD22	1.68	0.90
1:B:77[B]:PHE:HB3	1:B:78:PRO:HD3	1.61	0.82
1:A:196[A]:SER:OG	4:A:2190:HOH:O	1.99	0.80
1:B:233[A]:ASN:ND2	4:B:2229:HOH:O	2.13	0.79

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	258/263 (98%)	257 (100%)	1 (0%)	0	100	100
1	B	260/263 (99%)	259 (100%)	1 (0%)	0	100	100
All	All	518/526 (98%)	516 (100%)	2 (0%)	0	100	100

There are no Ramachandran outliers to report.

### 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	225/228 (99%)	220 (98%)	5 (2%)	60	30
1	B	227/228 (100%)	225 (99%)	2 (1%)	84	69
All	All	452/456 (99%)	445 (98%)	7 (2%)	70	48

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	107	GLU
1	B	136	ASN
1	A	115	ASN
1	A	4	ASN
1	B	107	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	252	ASN
1	B	101	GLN
1	B	140	GLN
1	A	140	GLN
1	B	136	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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	$\# Z  > 2$
2	KTA	A	1253	-	8,11,11	1.05	0	8,14,14	1.96	2 (25%)
2	KTA	B	1253	-	8,11,11	1.52	2 (25%)	8,14,14	3.37	3 (37%)

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	KTA	A	1253	-	-	0/2/14/14	0/1/1/1
2	KTA	B	1253	-	-	0/2/14/14	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1253	KTA	C4-C3	2.55	1.55	1.50
2	B	1253	KTA	C2-C3	3.18	1.56	1.50

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1253	KTA	C1-C2-C3	-7.62	104.66	113.72
2	A	1253	KTA	C1-C2-C3	-3.99	108.98	113.72
2	B	1253	KTA	O3-C3-C4	-2.02	118.26	122.02
2	A	1253	KTA	C4-C3-C2	3.17	120.25	115.89
2	B	1253	KTA	C4-C3-C2	4.64	122.28	115.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1253	KTA	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	251/263 (95%)	-0.57	2 (0%) 87 88	8, 12, 19, 37	19 (7%)
1	B	249/263 (94%)	-0.58	3 (1%) 81 82	9, 12, 18, 28	18 (7%)
All	All	500/526 (95%)	-0.58	5 (1%) 84 85	8, 12, 19, 37	37 (7%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	77[A]	PHE	4.0
1	B	161[A]	LEU	2.3
1	A	161[A]	LEU	2.1
1	B	253	THR	2.1
1	A	4	ASN	2.1

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	KTA	A	1253	11/11	0.94	0.10	2.64	18,21,23,24	3
2	KTA	B	1253	11/11	0.86	0.20	2.38	25,31,31,32	4
3	NI	B	1254	1/1	0.99	0.02	-	26,26,26,26	1
3	NI	A	1254	1/1	0.99	0.02	-	23,23,23,23	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.