



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 11:24 PM GMT

PDB ID : 4UBP  
Title : STRUCTURE OF BACILLUS PASTEURII UREASE INHIBITED WITH  
ACETOHYDROXAMIC ACID AT 1.55 Å RESOLUTION  
Authors : Benini, S.; Rypniewski, W.R.; Wilson, K.S.; Ciarli, S.; Mangani, S.  
Deposited on : 1999-02-25  
Resolution : 1.55 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

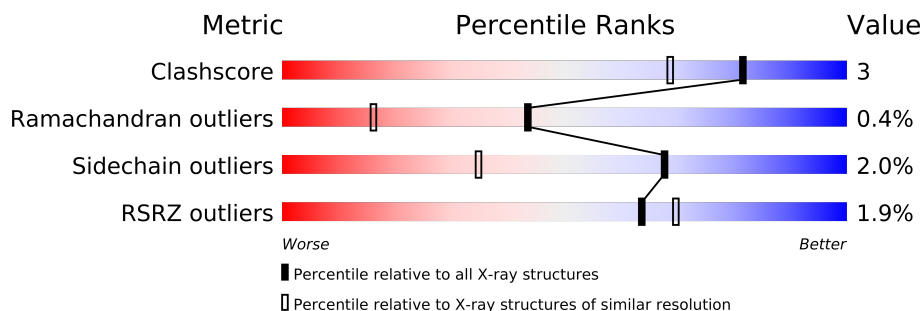
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.55 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))
Clashscore	79885	1249 (1.58-1.54)
Ramachandran outliers	78287	1212 (1.58-1.54)
Sidechain outliers	78261	1210 (1.58-1.54)
RSRZ outliers	66119	1117 (1.58-1.54)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	101	
2	B	126	
3	C	570	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6809 atoms, of which 0 are hydrogen and 0 are deuterium.

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 PROTEIN (UREASE (CHAIN A)).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	101	Total	C	N	O	S	7	0	0
			782	494	133	149	6			

- Molecule 2 is a protein called PROTEIN (UREASE (CHAIN B)).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	122	Total	C	N	O	S	20	0	0
			951	589	171	190	1			

- Molecule 3 is a protein called PROTEIN (UREASE (CHAIN C)).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	570	Total	C	N	O	S	26	0	0
			4323	2714	743	843	23			

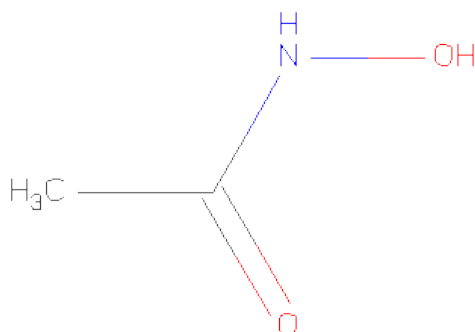
There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	19	GLU	ARG	VARIANT	UNP P41020
C	28	TRP	GLY	VARIANT	UNP P41020
C	29	ILE	-	INSERTION	UNP P41020
C	36	THR	TYR	VARIANT	UNP P41020
C	37	THR	TYR	VARIANT	UNP P41020
C	38	TYR	LEU	VARIANT	UNP P41020
C	220	KCX	LYS	MODIFIED RESIDUE	UNP P41020
C	263	LEU	VAL	VARIANT	UNP P41020
C	420	ILE	MET	VARIANT	UNP P41020

- Molecule 4 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	2	Total	Ni	0	0
			2	2		

- Molecule 5 is ACETOHYDROXAMIC ACID (three-letter code: HAE) (formula:  $C_2H_5NO_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	C	1	Total	C	N	O	0	0
			5	2	1	2		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	99	Total	O	0	0
			99	99		
6	B	146	Total	O	0	0
			146	146		
6	C	501	Total	O	0	0
			501	501		

### 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.

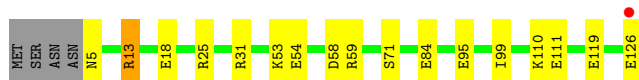
- Molecule 1: PROTEIN (UREASE (CHAIN A))

Chain A:



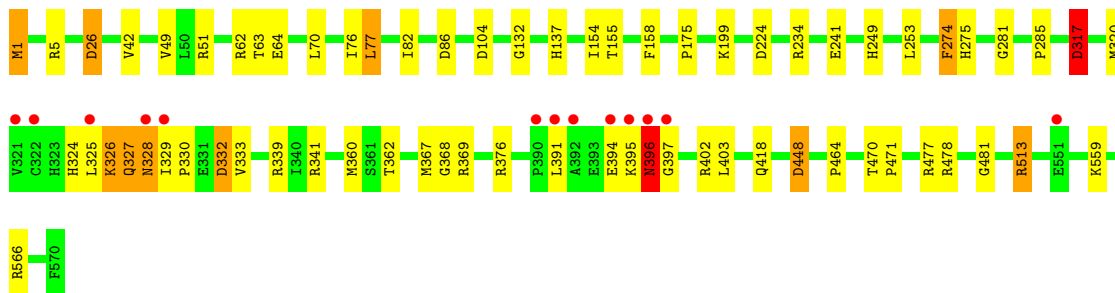
- Molecule 2: PROTEIN (UREASE (CHAIN B))

Chain B:



- Molecule 3: PROTEIN (UREASE (CHAIN C))

Chain C:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	130.88Å 130.88Å 189.00Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	32.73 – 1.55 32.73 – 1.55	Depositor EDS
% Data completeness (in resolution range)	99.5 (32.73-1.55) 99.4 (32.73-1.55)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.12 (at 1.55Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.151 , 0.190 0.169 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	17.7	Xtriage
Anisotropy	0.409	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 51.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	0 of 136971 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	6809	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.35% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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, HAE, ACE, KCX

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	1.06	2/790 (0.3%)	1.30	9/1063 (0.8%)
2	B	2.95	12/963 (1.2%)	1.76	19/1296 (1.5%)
3	C	1.31	15/4392 (0.3%)	1.53	49/5955 (0.8%)
All	All	1.65	29/6145 (0.5%)	1.54	77/8314 (0.9%)

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
2	B	0	1
3	C	0	1
All	All	0	3

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	241	GLU	CD-OE1	-51.68	0.68	1.25
2	B	111	GLU	CD-OE2	44.98	1.75	1.25
2	B	111	GLU	CG-CD	40.13	2.12	1.51
2	B	119	GLU	CG-CD	36.12	2.06	1.51
2	B	13	ARG	NE-CZ	29.66	1.71	1.33

The worst 5 of 77 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	241	GLU	OE1-CD-OE2	34.34	164.51	123.30
3	C	320	MET	CG-SD-CE	27.60	144.37	100.20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	111	GLU	CG-CD-OE2	-21.34	75.62	118.30
3	C	42	VAL	CG1-CB-CG2	20.41	143.56	110.90
3	C	5	ARG	NE-CZ-NH2	-19.93	110.33	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	22	ARG	Sidechain
2	B	13	ARG	Sidechain
3	C	317	ASP	Sidechain

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	782	0	806	12	0
2	B	951	0	937	4	0
3	C	4323	0	4290	23	1
4	C	2	0	0	0	0
5	C	5	0	4	0	0
6	A	99	0	0	1	0
6	B	146	0	0	1	0
6	C	501	0	0	8	2
All	All	6809	0	6037	37	3

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

The worst 5 of 37 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:0:ACE:CH3	1:A:0:ACE:O	1.86	1.20
1:A:0:ACE:H3	1:A:0:ACE:O	1.46	1.14
1:A:0:ACE:C	1:A:1:MET:H2	1.50	1.12
1:A:0:ACE:H1	1:A:0:ACE:C	1.50	1.10

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:0:ACE:H2	1:A:0:ACE:C	1.50	1.08

All (3) 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	Distance(Å)	Clash(Å)
3:C:327:GLN:O	3:C:327:GLN:OE1[7_556]	1.76	0.44
6:C:1191:HOH:O	6:C:1191:HOH:O[11_555]	1.81	0.39
6:C:1198:HOH:O	6:C:1198:HOH:O[10_665]	1.92	0.28

## 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	98/101 (97%)	98 (100%)	0	0	100	100
2	B	120/126 (95%)	115 (96%)	4 (3%)	1 (1%)	27	5
3	C	567/570 (100%)	544 (96%)	21 (4%)	2 (0%)	43	16
All	All	785/797 (98%)	757 (96%)	25 (3%)	3 (0%)	43	16

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	396	ASN
2	B	99	ILE
3	C	367	MET

### 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	85/85 (100%)	84 (99%)	1 (1%)	82	59
2	B	101/105 (96%)	100 (99%)	1 (1%)	85	66
3	C	460/460 (100%)	449 (98%)	11 (2%)	61	27
All	All	646/650 (99%)	633 (98%)	13 (2%)	68	35

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	253	LEU
3	C	285	PRO
3	C	328	ASN
3	C	158	PHE
3	C	326	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	12	GLN
3	C	328	ASN
3	C	418	GLN
3	C	519	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	KCX	C	220	3,4	11,11,12	7.43	5 (45%)	10,12,14	3.64	5 (50%)

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	KCX	C	220	3,4	-	0/8/10/12	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	220	KCX	O-C	20.05	1.25	1.11
3	C	220	KCX	CX-NZ	13.11	1.45	1.32
3	C	220	KCX	OQ1-CX	4.33	1.29	1.21
3	C	220	KCX	CA-C	2.76	1.53	1.48
3	C	220	KCX	CB-CA	2.03	1.55	1.53

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	220	KCX	C-CA-N	-6.95	106.88	113.83
3	C	220	KCX	OQ1-CX-NZ	-6.48	114.17	124.94
3	C	220	KCX	OQ2-CX-NZ	4.93	122.47	116.33
3	C	220	KCX	CE-NZ-CX	-3.49	115.64	121.99
3	C	220	KCX	CD-CE-NZ	-2.05	106.03	112.21

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	$\# Z  > 2$
5	HAE	C	800	4	4,4,4	0.43	0	4,4,4	2.19	2 (50%)

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
5	HAE	C	800	4	-	0/1/2/2	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	800	HAE	C1-C2-N	3.08	117.91	114.98
5	C	800	HAE	O-N-C2	-2.73	116.17	119.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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 ⓘ

### 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	101/101 (100%)	-0.67	1 (0%) 79 84	18, 22, 30, 42	2 (1%)
2	B	122/126 (96%)	-0.33	1 (0%) 83 87	19, 25, 35, 61	6 (4%)
3	C	570/570 (100%)	-0.46	13 (2%) 57 62	17, 21, 42, 106	12 (2%)
All	All	793/797 (99%)	-0.47	15 (1%) 64 70	17, 22, 38, 106	20 (2%)

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	391	LEU	7.6
3	C	390	PRO	5.6
3	C	328	ASN	5.3
3	C	397	GLY	4.7
3	C	329	ILE	4.2

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

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	KCX	C	220	12/13	0.05	-0.73	16,18,19,21	0

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	HAE	C	800	5/5	0.05	-0.65	23,24,28,30	0
4	NI	C	799	1/1	0.02	-4.21	20,20,20,20	0
4	NI	C	798	1/1	0.01	-6.40	22,22,22,22	0

## 6.5 Other polymers

There are no such residues in this entry.