



# wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 02:41 PM GMT

PDB ID : 1VHD  
Title : Crystal structure of an iron containing alcohol dehydrogenase  
Authors : Structural GenomiX  
Deposited on : 2003-12-01  
Resolution : 1.60 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>

---

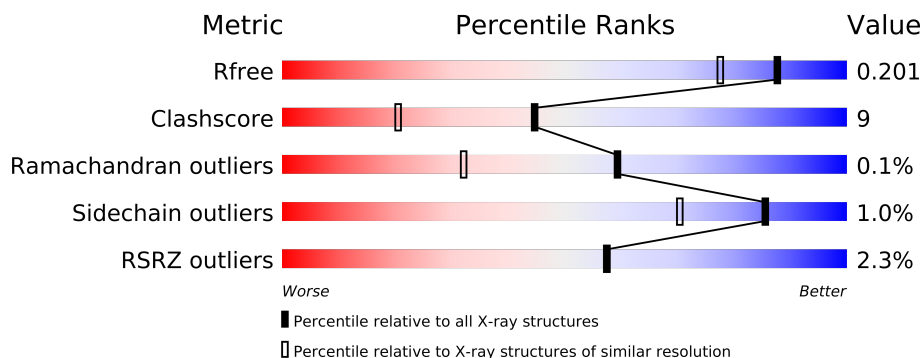
The following versions of software and data (see [references](#)) were used in the production of this report:

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 1.60 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))
$R_{free}$	66092	1872 (1.60-1.60)
Clashscore	79885	2199 (1.60-1.60)
Ramachandran outliers	78287	2126 (1.60-1.60)
Sidechain outliers	78261	2125 (1.60-1.60)
RSRZ outliers	66119	1872 (1.60-1.60)

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	371	
1	B	371	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	CAC	B	370	-	X
3	ZN	A	371	-	X
4	NAP	A	372	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6357 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 alcohol dehydrogenase, iron-containing.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	361	Total	C	N	O	S	Se	0	24	0
			2854	1815	467	559	2	11			
1	B	361	Total	C	N	O	S	Se	0	14	0
			2831	1810	462	545	2	12			

There are 48 discrepancies between the modelled and reference sequences:

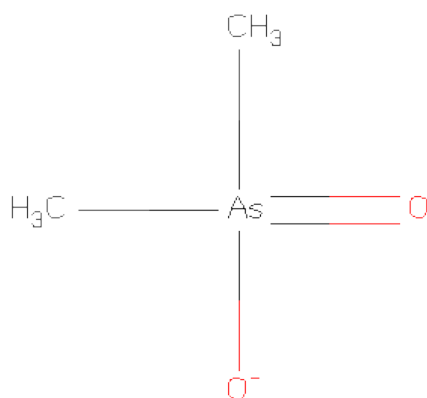
Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MSE	-	cloning artifact	UNP Q9X022
A	0	SER	-	cloning artifact	UNP Q9X022
A	1	LEU	-	cloning artifact	UNP Q9X022
A	6	MSE	MET	modified residue	UNP Q9X022
A	76	MSE	MET	modified residue	UNP Q9X022
A	99	MSE	MET	modified residue	UNP Q9X022
A	163	MSE	MET	modified residue	UNP Q9X022
A	178	MSE	MET	modified residue	UNP Q9X022
A	215	MSE	MET	modified residue	UNP Q9X022
A	236	MSE	MET	modified residue	UNP Q9X022
A	245	MSE	MET	modified residue	UNP Q9X022
A	276	MSE	MET	modified residue	UNP Q9X022
A	282	MSE	MET	modified residue	UNP Q9X022
A	285	MSE	MET	modified residue	UNP Q9X022
A	360	GLU	-	cloning artifact	UNP Q9X022
A	361	GLY	-	cloning artifact	UNP Q9X022
A	362	GLY	-	cloning artifact	UNP Q9X022
A	363	SER	-	cloning artifact	UNP Q9X022
A	364	HIS	-	cloning artifact	UNP Q9X022
A	365	HIS	-	cloning artifact	UNP Q9X022
A	366	HIS	-	cloning artifact	UNP Q9X022
A	367	HIS	-	cloning artifact	UNP Q9X022
A	368	HIS	-	cloning artifact	UNP Q9X022
A	369	HIS	-	cloning artifact	UNP Q9X022
B	-1	MSE	-	cloning artifact	UNP Q9X022

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	SER	-	cloning artifact	UNP Q9X022
B	1	LEU	-	cloning artifact	UNP Q9X022
B	6	MSE	MET	modified residue	UNP Q9X022
B	76	MSE	MET	modified residue	UNP Q9X022
B	99	MSE	MET	modified residue	UNP Q9X022
B	163	MSE	MET	modified residue	UNP Q9X022
B	178	MSE	MET	modified residue	UNP Q9X022
B	215	MSE	MET	modified residue	UNP Q9X022
B	236	MSE	MET	modified residue	UNP Q9X022
B	245	MSE	MET	modified residue	UNP Q9X022
B	276	MSE	MET	modified residue	UNP Q9X022
B	282	MSE	MET	modified residue	UNP Q9X022
B	285	MSE	MET	modified residue	UNP Q9X022
B	360	GLU	-	cloning artifact	UNP Q9X022
B	361	GLY	-	cloning artifact	UNP Q9X022
B	362	GLY	-	cloning artifact	UNP Q9X022
B	363	SER	-	cloning artifact	UNP Q9X022
B	364	HIS	-	cloning artifact	UNP Q9X022
B	365	HIS	-	cloning artifact	UNP Q9X022
B	366	HIS	-	cloning artifact	UNP Q9X022
B	367	HIS	-	cloning artifact	UNP Q9X022
B	368	HIS	-	cloning artifact	UNP Q9X022
B	369	HIS	-	cloning artifact	UNP Q9X022

- Molecule 2 is CACODYLATE ION (three-letter code: CAC) (formula:  $\text{C}_2\text{H}_6\text{AsO}_2$ ).

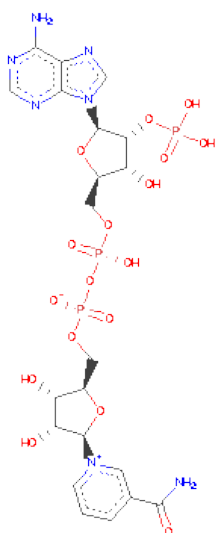


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	As	C	O	0	0
			4	1	2	1		
2	B	1	Total	As	C	O	0	0
			4	1	2	1		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 4 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDEPHOSPHATE (three-letter code: NAP) (formula: C<sub>21</sub>H<sub>28</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
4	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	338	Total	O	0	0
			338	338		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	228	Total 228	O 228	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.90Å 85.07Å 72.35Å 90.00° 96.57° 90.00°	Depositor
Resolution (Å)	11.99 – 1.60 11.99 – 1.60	Depositor EDS
% Data completeness (in resolution range)	(Not available) (11.99-1.60) 99.9 (11.99-1.60)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.99 (at 1.60Å)	Xtriage
Refinement program	REFMAC 4	Depositor
R, $R_{free}$	0.190 , 0.208 0.183 , 0.201	Depositor DCC
$R_{free}$ test set	4585 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	13.5	Xtriage
Anisotropy	0.359	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 45.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 91625 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6357	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.08% of the height of the origin peak. No significant pseudotranslation is detected.*

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



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CAC, ZN, NAP

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	0.29	0/3044	0.57	1/4090 (0.0%)
1	B	0.28	0/2953	0.54	0/3969
All	All	0.29	0/5997	0.55	1/8059 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	25	ASP	CB-CA-C	5.32	121.03	110.40

There are no chirality outliers.

There are no planarity outliers.

### 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	2854	0	2831	47	0
1	B	2831	0	2835	55	0
2	A	4	0	0	4	0
2	B	4	0	0	4	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	48	0	25	3	0
4	B	48	0	25	4	0
5	A	338	0	0	9	0
5	B	228	0	0	6	0
All	All	6357	0	5716	99	0

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 9.

The worst 5 of 99 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:120:TYR:CD1	2:B:370:CAC:C1	2.33	1.11
1:B:120:TYR:CE1	2:B:370:CAC:C1	2.35	1.08
1:A:13:GLY:HA3	1:B:1:LEU:HD12	1.52	0.91
1:A:120:TYR:CD1	2:A:370:CAC:C1	2.60	0.85
1:A:281:VAL:HG12	1:A:285[B]:MSE:HE2	1.63	0.81

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	383/371 (103%)	375 (98%)	8 (2%)	0	100	100
1	B	373/371 (100%)	362 (97%)	10 (3%)	1 (0%)	50	24
All	All	756/742 (102%)	737 (98%)	18 (2%)	1 (0%)	59	32

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	39	SER

### 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	331/306 (108%)	328 (99%)	3 (1%)	87	74
1	B	317/306 (104%)	314 (99%)	3 (1%)	87	74
All	All	648/612 (106%)	642 (99%)	6 (1%)	85	74

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	256	HIS
1	B	256	HIS
1	B	74	ASN
1	A	163	MSE
1	B	163	MSE

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	336	HIS
1	B	22	ASN
1	B	74	ASN
1	B	336	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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	CAC	A	370	1	0,3,4	0.00	-	1,3,6	4.08	1 (100%)
4	NAP	A	372	-	52,52,52	1.47	5 (9%)	80,80,80	1.36	9 (11%)
2	CAC	B	370	1	0,3,4	0.00	-	1,3,6	4.04	1 (100%)
4	NAP	B	372	-	52,52,52	1.52	5 (9%)	80,80,80	1.39	9 (11%)

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	CAC	A	370	1	-	0/0/0/0	0/0/0/0
4	NAP	A	372	-	-	0/35/67/67	0/3/5/5
2	CAC	B	370	1	-	0/0/0/0	0/0/0/0
4	NAP	B	372	-	-	0/35/67/67	0/3/5/5

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	372	NAP	C2N-N1N	4.98	1.41	1.35
4	A	372	NAP	C2N-N1N	4.92	1.41	1.35
4	B	372	NAP	C2N-C3N	4.43	1.45	1.38
4	A	372	NAP	C2N-C3N	4.21	1.44	1.38
4	B	372	NAP	C2D-C1D	3.98	1.59	1.53

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	372	NAP	O4B-C1B-C2B	4.50	111.15	106.95
4	B	372	NAP	O7N-C7N-N7N	-4.34	116.33	122.59

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	372	NAP	O7N-C7N-N7N	-4.28	116.41	122.59
4	A	372	NAP	N3A-C2A-N1A	-4.19	125.20	128.71
4	A	372	NAP	O4B-C1B-C2B	4.12	110.80	106.95

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	361/371 (97%)	-0.23	6 (1%) 67 67	6, 12, 26, 37	0
1	B	361/371 (97%)	0.12	9 (2%) 54 54	7, 18, 33, 39	0
All	All	722/742 (97%)	-0.06	15 (2%) 57 61	6, 16, 32, 39	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	0	SER	4.6
1	B	42	LYS	4.3
1	B	83	ARG	4.0
1	B	84	ASN	3.6
1	B	127	HIS	3.2

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 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( $\text{\AA}^2$ )	Q<0.9
3	ZN	A	371	1/1	0.11	3.48	18,18,18,18	0
4	NAP	A	372	48/48	0.17	2.89	11,20,29,30	0
2	CAC	B	370	4/5	0.20	2.88	29,30,30,31	0
2	CAC	A	370	4/5	0.20	1.98	26,27,28,28	0
4	NAP	B	372	48/48	0.18	1.87	26,33,38,39	0
3	ZN	B	371	1/1	0.11	1.50	21,21,21,21	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.