wwPDB X-ray Structure Validation Summary Report (i)

Sep 11, 2014 – 08:47 PM EDT

PDB ID	:	4B6R
Title	:	Structure of Helicobacter pylori Type II Dehydroquinase inhibited by (2S)-2-(4-methoxy)benzyl-3-dehydroquinicacid
Authors	:	Otero, J.M.; Llamas-Saiz, A.L.; Lence, E.; Tizon, L.; Peon, A.; Prazeres, V.F.V.; Lamb, H.; Hawkins, A.R.; Gonzalez-Bello, C.; van Raaij, M.J.
Deposited on	:	2012-08-14
Resolution	:	2.00 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at http://wwpdb.org/ValidationPDFNotes.html

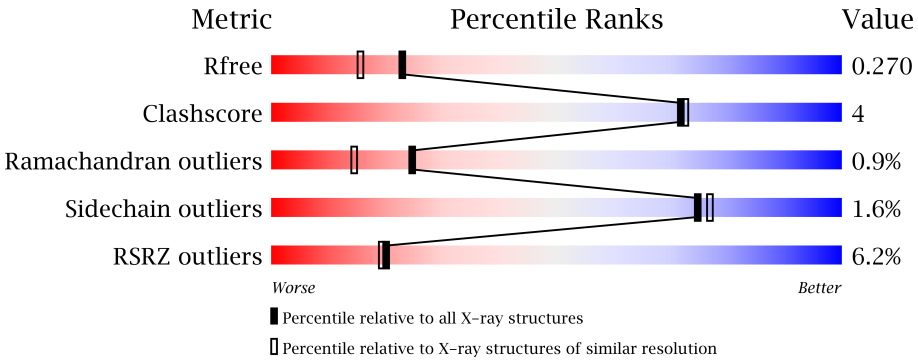
The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable23489
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)	:	stable23489

1 Overall quality at a glance

The reported resolution of this entry is 2.00 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}	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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 >=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	167	
1	B	167	
1	C	167	

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

Mol	Type	Chain	Res	Geometry	Electron density
3	SO4	B	1159	-	X

2 Entry composition

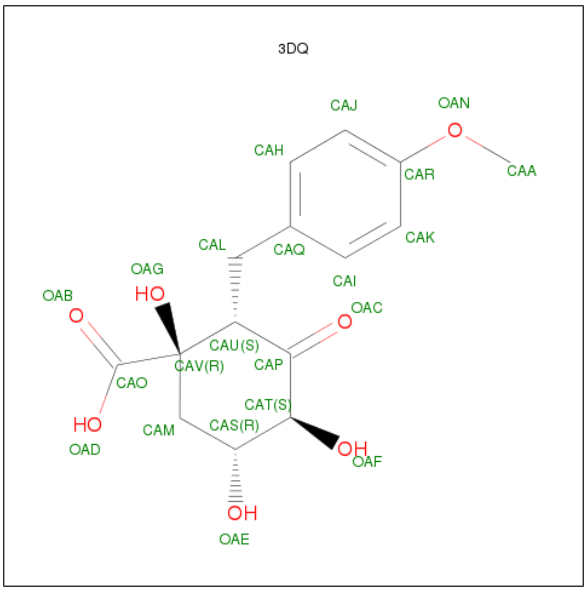
There are 4 unique types of molecules in this entry. The entry contains 3964 atoms, of which 0 are hydrogen and 0 are deuterium.

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 3-DEHYDROQUINATE DEHYDRATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	157	Total	C	N	O	S	0	7	0
			1246	793	206	235	12			
1	B	157	Total	C	N	O	S	0	5	0
			1236	787	204	232	13			
1	C	157	Total	C	N	O	S	0	4	0
			1232	784	205	231	12			

- Molecule 2 is (1R,2S,4S,5R)-2-(4-METHOXYPHENYL)METHYL-1,4,5-TRIHYDROXY-3-OXOCYCLOHEXANE-1-CARBOXYLICACID (three-letter code: 3DQ) (formula: $C_{15}H_{18}O_7$).



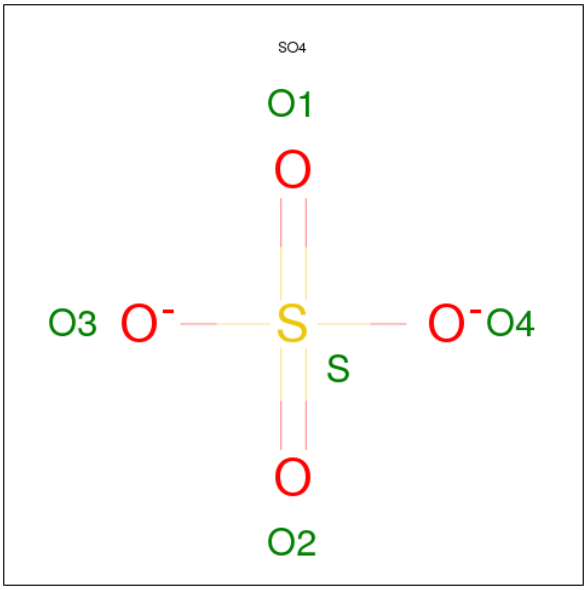
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			22	15	7		
2	B	1	Total	C	O	0	0
			22	15	7		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total C O 22 15 7	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total O S 5 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	86	Total O 86 86	0	0
4	B	49	Total O 49 49	0	0
4	C	44	Total O 44 44	0	0

3 Residue-property plots (i)

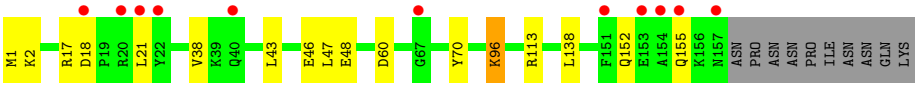
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: 3-DEHYDROQUINATE DEHYDRATASE

Chain A:

• Molecule 1: 3-DEHYDROQUINATE DEHYDRATASE

Chain B:



• Molecule 1: 3-DEHYDROQUINATE DEHYDRATASE

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 42 2 2	Depositor
Cell constants a, b, c, α, β, γ	100.44Å 100.44Å 104.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.70 – 2.00 58.70 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (58.70-2.00) 99.4 (58.70-2.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.43 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.212 , 0.276 0.210 , 0.270	Depositor DCC
R_{free} test set	1876 reflections (5.42%)	DCC
Wilson B-factor (Å^2)	38.1	Xtriage
Anisotropy	0.148	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 39.7	EDS
Estimated twinning fraction	0.155 for -h,l,k 0.038 for -l,-k,-h	Xtriage
L-test for twinning	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	2 of 36460 reflections (0.005%)	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3964	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.12% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] 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: 3DQ, SO4

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	0.75	0/1294	0.68	0/1741
1	B	0.71	0/1276	0.71	1/1717 (0.1%)
1	C	0.70	0/1267	0.67	0/1706
All	All	0.72	0/3837	0.69	1/5164 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	60	ASP	CB-CG-OD1	5.17	122.96	118.30

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	1246	0	1242	13	0
1	B	1236	0	1234	10	0
1	C	1232	0	1235	11	0
2	A	22	0	17	3	0
2	B	22	0	17	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	22	0	17	1	0
3	B	5	0	0	0	0
4	A	86	0	0	0	0
4	B	49	0	0	1	0
4	C	44	0	0	0	0
All	All	3964	0	3762	33	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 4.

The worst 5 of 33 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:16:HIS:O	1:A:17:ARG:HG2	1.58	1.02
1:B:48:GLU:OE1	1:B:70:TYR:OH	1.85	0.92
1:C:48:GLU:OE1	1:C:70:TYR:OH	1.90	0.89
1:C:16:HIS:O	1:C:17:ARG:HG2	1.77	0.82
1:A:36:THR:HG23	1:A:40:GLN:OE1	1.97	0.65

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	162/167 (97%)	154 (95%)	7 (4%)	1 (1%)	33 24
1	B	160/167 (96%)	152 (95%)	7 (4%)	1 (1%)	33 24
1	C	158/167 (95%)	151 (96%)	5 (3%)	2 (1%)	18 8
All	All	480/501 (96%)	457 (95%)	19 (4%)	4 (1%)	25 17

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	18	ASP
1	B	18	ASP
1	A	17	ARG
1	C	20	ARG

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	136/139 (98%)	133 (98%)	3 (2%)	64	65
1	B	134/139 (96%)	131 (98%)	3 (2%)	64	65
1	C	133/139 (96%)	132 (99%)	1 (1%)	89	92
All	All	403/417 (97%)	396 (98%)	7 (2%)	75	75

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

Mol	Chain	Res	Type
1	B	1	MET
1	C	155	GLN
1	B	21	LEU
1	A	143[B]	ASN
1	B	96	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	16	HIS
1	C	155	GLN

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

4 ligands are 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	$\#\|Z\| > 2$
2	3DQ	A	1158	-	23,23,23	1.07	1 (4%)	34,34,34	2.00	6 (17%)
2	3DQ	B	1158	-	23,23,23	0.89	0	34,34,34	1.48	2 (5%)
3	SO4	B	1159	-	4,4,4	0.20	0	6,6,6	0.24	0
2	3DQ	C	1158	-	23,23,23	1.10	1 (4%)	34,34,34	2.40	5 (14%)

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	3DQ	A	1158	-	-	0/12/35/35	0/2/2/2
2	3DQ	B	1158	-	-	0/12/35/35	0/2/2/2
3	SO4	B	1159	-	-	0/0/0/0	0/0/0/0
2	3DQ	C	1158	-	-	0/12/35/35	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1158	3DQ	CAV-CAU	2.75	1.58	1.54
2	A	1158	3DQ	CAM-CAS	2.32	1.56	1.53

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	1158	3DQ	CAL-CAU-CAP	-11.36	96.58	112.36
2	A	1158	3DQ	CAL-CAU-CAP	-8.42	100.66	112.36
2	B	1158	3DQ	CAL-CAU-CAP	-5.98	104.05	112.36
2	C	1158	3DQ	CAQ-CAL-CAU	3.88	121.11	114.08
2	A	1158	3DQ	CAM-CAV-CAU	3.46	114.94	109.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers (i)

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^{th} 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^2)	Q<0.9
1	A	157/167 (94%)	0.24	9 (5%)	23	22	23, 35, 78, 100	0
1	B	157/167 (94%)	0.34	11 (7%)	16	15	27, 40, 83, 106	0
1	C	157/167 (94%)	0.38	9 (5%)	23	22	26, 40, 84, 103	0
All	All	471/501 (94%)	0.32	29 (6%)	20	19	23, 38, 83, 106	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	21	LEU	10.7
1	B	22	TYR	6.5
1	B	20	ARG	6.4
1	C	22	TYR	6.2
1	C	18	ASP	6.0

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^{th} 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^2)	Q<0.9
3	SO4	B	1159	5/5	0.20	4.19	99,99,100,100	0
2	3DQ	C	1158	22/22	0.14	-0.22	32,43,55,57	0
2	3DQ	B	1158	22/22	0.12	-0.27	29,39,56,59	0
2	3DQ	A	1158	22/22	0.12	-0.33	24,39,53,55	0

6.5 Other polymers

There are no such residues in this entry.