wwPDB X-ray Structure Validation Summary Report (i)

Nov 7, 2023 – 03:20 am GMT

PDB ID	:	4B5W
Title	:	Crystal structures of divalent metal dependent pyruvate aldolase R70A mutant, HpaI, in complex with pyruvate
Authors	:	Coincon, M.; Wang, W.; Seah, S.Y.K.; Sygusch, J.
Deposited on	:	2012-08-07
Resolution	:	1.79 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	NOT EXECUTED
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
buster-report	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

PERCENTILES INFOmissingINFO

1 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 24975 atoms, of which 11597 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 4-HYDROXY-2-OXO-HEPTANE-1,7-DIOATE ALDOLASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	256	Total	C	H	N	O	S	0	0	0
			3868	1217	1943	339	363	6			
1	B	253	Total	C	H	N	O	S	0	0	0
			3834	1207	1926	336	359	6			
1	C	252	Total	C	H	N	O	S	0	0	0
			3817	1202	1918	334	357	6			
1	D	253	Total	C	H	N	O	S	0	0	0
			3831	1207	1923	336	359	6			
1	E	255	Total	C	H	N	O	S	0	0	0
			3858	1214	1938	338	362	6			
1	F	254	Total	C	H	N	O	S	0	0	0
			3844	1210	1931	337	360	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	70	ALA	ARG	engineered mutation	UNP B1IS70
B	70	ALA	ARG	engineered mutation	UNP B1IS70
C	70	ALA	ARG	engineered mutation	UNP B1IS70
D	70	ALA	ARG	engineered mutation	UNP B1IS70
E	70	ALA	ARG	engineered mutation	UNP B1IS70
F	70	ALA	ARG	engineered mutation	UNP B1IS70

- Molecule 2 is COBALT (II) ION (three-letter code: CO) (formula: Co).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Co	0	0
			1	1		
2	B	1	Total	Co	0	0
			1	1		
2	C	1	Total	Co	0	0
			1	1		

Continued on next page...

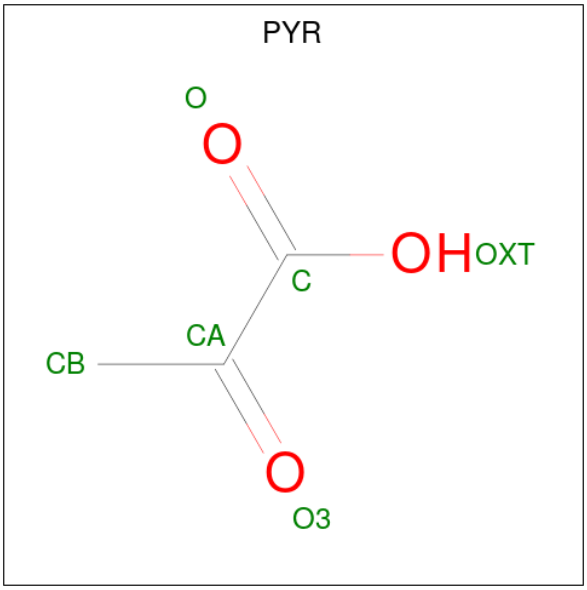
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	1	Total Co 1 1	0	0
2	E	1	Total Co 1 1	0	0
2	F	1	Total Co 1 1	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Ca 1 1	0	0
3	C	1	Total Ca 1 1	0	0
3	D	1	Total Ca 1 1	0	0
3	E	1	Total Ca 1 1	0	0
3	F	1	Total Ca 1 1	0	0

- Molecule 4 is PYRUVIC ACID (three-letter code: PYR) (formula: $C_3H_4O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C H O 9 3 3 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C H O 9 3 3 3	0	0
4	C	1	Total C H O 9 3 3 3	0	0
4	D	1	Total C H O 9 3 3 3	0	0
4	E	1	Total C H O 9 3 3 3	0	0
4	F	1	Total C H O 9 3 3 3	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	365	Total O 365 365	0	0
5	B	322	Total O 322 322	0	0
5	C	290	Total O 290 290	0	0
5	D	306	Total O 306 306	0	0
5	E	306	Total O 306 306	0	0
5	F	269	Total O 269 269	0	0

SEQUENCE-PLOTS INFOmissingINFO

2 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	81.35Å 119.14Å 140.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.49 – 1.79	Depositor
% Data completeness (in resolution range)	99.3 (38.49-1.79)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.71 (at 1.79Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.144 , 0.184	Depositor
Wilson B-factor (Å^2)	17.4	Xtriage
Anisotropy	0.059	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	24975	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.05% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

3 Model quality

3.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, PYR, CO

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	0.47	0/1961	0.60	0/2674
1	B	0.47	0/1944	0.59	0/2650
1	C	0.47	0/1935	0.60	0/2638
1	D	0.45	0/1944	0.63	2/2650 (0.1%)
1	E	0.48	0/1956	0.62	0/2667
1	F	0.45	0/1949	0.60	0/2657
All	All	0.46	0/11689	0.61	2/15936 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	40	LEU	CA-CB-CG	-7.19	98.76	115.30
1	D	90	LEU	CA-CB-CG	5.24	127.35	115.30

There are no chirality outliers.

There are no planarity outliers.

3.2 Too-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 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	1925	1943	1941	15	1
1	B	1908	1926	1924	20	1
1	C	1899	1918	1916	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1908	1923	1924	15	0
1	E	1920	1938	1936	19	1
1	F	1913	1931	1929	12	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	6	3	0	0	0
4	B	6	3	0	0	0
4	C	6	3	0	0	0
4	D	6	3	0	0	0
4	E	6	3	0	0	0
4	F	6	3	0	0	0
5	A	365	0	0	8	2
5	B	322	0	0	7	3
5	C	290	0	0	7	3
5	D	306	0	0	4	1
5	E	306	0	0	8	1
5	F	269	0	0	4	1
All	All	13378	11597	11570	85	7

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 4.

The worst 5 of 85 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:C:164:GLU:OE2	5:C:2167:HOH:O	1.88	0.91
1:E:204:GLU:OE2	5:E:2266:HOH:O	1.96	0.82
1:A:1:MET:N	5:A:2001:HOH:O	2.13	0.82
1:A:106:ARG:NH2	5:A:2209:HOH:O	2.22	0.73
1:C:1:MET:N	5:C:2001:HOH:O	2.23	0.71

The worst 5 of 7 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	Interatomic distance (Å)	Clash overlap (Å)
5:C:2015:HOH:O	5:D:2234:HOH:O[3_555]	1.85	0.35
5:A:2291:HOH:O	5:F:2214:HOH:O[3_545]	2.01	0.19
5:B:2244:HOH:O	5:C:2234:HOH:O[3_645]	2.08	0.12
1:B:190:GLU:OE1	5:E:2192:HOH:O[2_564]	2.10	0.10
5:A:2001:HOH:O	5:B:2299:HOH:O[2_565]	2.12	0.08

3.3 Torsion angles (i)

3.3.1 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	254/256 (99%)	252 (99%)	2 (1%)	0	100	100
1	B	251/256 (98%)	246 (98%)	5 (2%)	0	100	100
1	C	250/256 (98%)	248 (99%)	2 (1%)	0	100	100
1	D	251/256 (98%)	247 (98%)	4 (2%)	0	100	100
1	E	253/256 (99%)	251 (99%)	2 (1%)	0	100	100
1	F	252/256 (98%)	246 (98%)	6 (2%)	0	100	100
All	All	1511/1536 (98%)	1490 (99%)	21 (1%)	0	100	100

There are no Ramachandran outliers to report.

3.3.2 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	196/196 (100%)	195 (100%)	1 (0%)	88 87
1	B	195/196 (100%)	194 (100%)	1 (0%)	88 87
1	C	194/196 (99%)	194 (100%)	0	100 100
1	D	195/196 (100%)	194 (100%)	1 (0%)	88 87
1	E	196/196 (100%)	196 (100%)	0	100 100
1	F	195/196 (100%)	191 (98%)	4 (2%)	53 42
All	All	1171/1176 (100%)	1164 (99%)	7 (1%)	86 84

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

Mol	Chain	Res	Type
1	F	51	GLN
1	F	122	LEU
1	F	253	GLN
1	F	159	GLN
1	D	40	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	188	HIS

3.3.3 RNA (i)

There are no RNA molecules in this entry.

3.4 Non-standard residues in protein, DNA, RNA chains (i)

Mogul was not executed - this section is therefore empty.

3.5 Carbohydrates (i)

Mogul was not executed - this section is therefore empty.

3.6 Ligand geometry (i)

Mogul was not executed - this section is therefore empty.

3.7 Other polymers (i)

Mogul was not executed - this section is therefore empty.

3.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

4 Fit of model and data

4.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

4.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

4.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

4.4 Ligands

EDS failed to run properly - this section is therefore empty.

4.5 Other polymers

EDS failed to run properly - this section is therefore empty.