wwPDB X-ray Structure Validation Summary Report (i)

Oct 24, 2023 – 05:17 PM EDT

PDB ID	:	3HLI
Title	:	diisopropyl fluorophosphatase (DFPase), active site mutants
Authors	:	Chen, J.C.-H.; Blum, M.-M.
Deposited on	:	2009-05-27
Resolution	:	1.40 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:

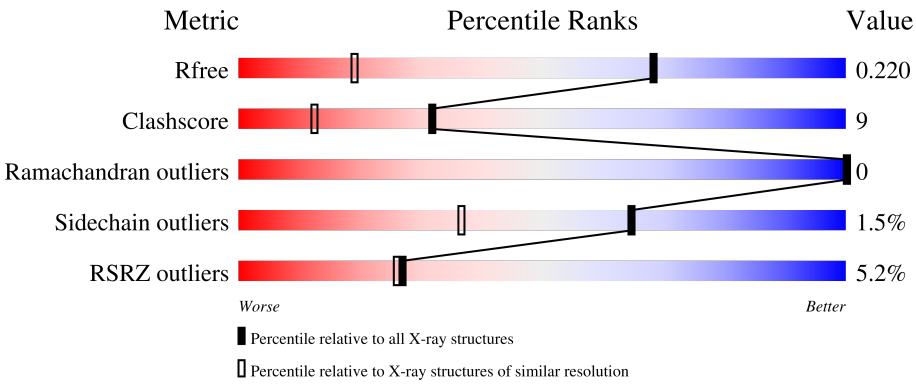
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.40 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}	130704	1714 (1.40-1.40)
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=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	314	4% 84% 15%
1	B	314	6% 80% 20% •
1	C	314	4% 83% 16% •
1	D	314	6% 80% 19% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10803 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 Diisopropyl-fluorophosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	313	2443	1551	414	462	16	0	0	0
1	B	312	2438	1548	413	461	16	0	0	0
1	C	312	2438	1548	413	461	16	0	0	0
1	D	312	2438	1548	413	461	16	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	ASP	GLU	engineered mutation	UNP Q7SIG4
A	144	ALA	TYR	engineered mutation	UNP Q7SIG4
A	146	ALA	ARG	engineered mutation	UNP Q7SIG4
A	195	MET	THR	engineered mutation	UNP Q7SIG4
B	37	ASP	GLU	engineered mutation	UNP Q7SIG4
B	144	ALA	TYR	engineered mutation	UNP Q7SIG4
B	146	ALA	ARG	engineered mutation	UNP Q7SIG4
B	195	MET	THR	engineered mutation	UNP Q7SIG4
C	37	ASP	GLU	engineered mutation	UNP Q7SIG4
C	144	ALA	TYR	engineered mutation	UNP Q7SIG4
C	146	ALA	ARG	engineered mutation	UNP Q7SIG4
C	195	MET	THR	engineered mutation	UNP Q7SIG4
D	37	ASP	GLU	engineered mutation	UNP Q7SIG4
D	144	ALA	TYR	engineered mutation	UNP Q7SIG4
D	146	ALA	ARG	engineered mutation	UNP Q7SIG4
D	195	MET	THR	engineered mutation	UNP Q7SIG4

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
2	A	2	2	2	0	0
			Total	Ca		
2	B	2	2	2	0	0
			Total	Ca		
2	C	2	2	2	0	0
			Total	Ca		
2	D	2	2	2	0	0

- Molecule 3 is water.

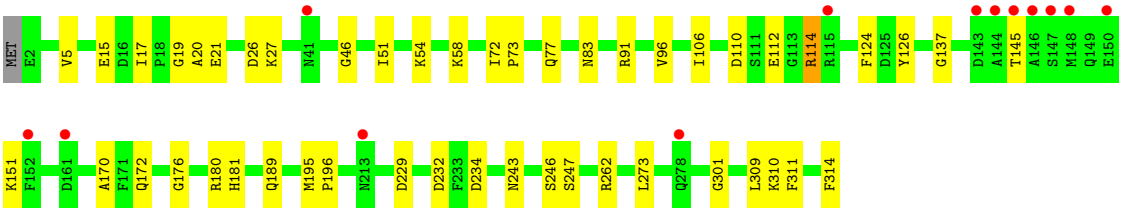
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	250	250	250	0	0
			Total	O		
3	B	273	273	273	0	0
			Total	O		
3	C	270	270	270	0	0
			Total	O		
3	D	245	245	245	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

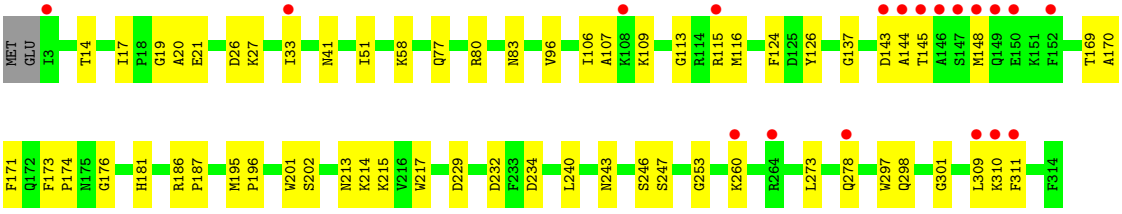
• Molecule 1: Diisopropyl-fluorophosphatase

Chain A:



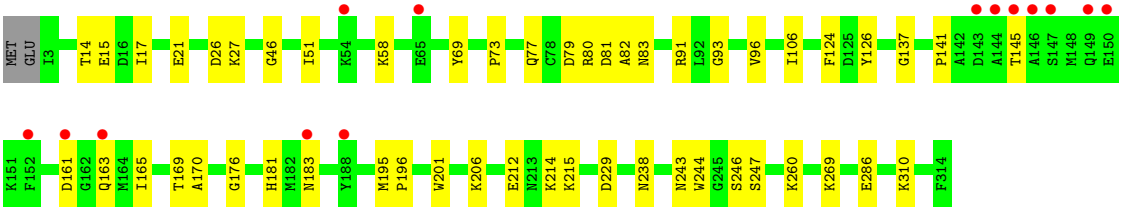
• Molecule 1: Diisopropyl-fluorophosphatase

Chain B:



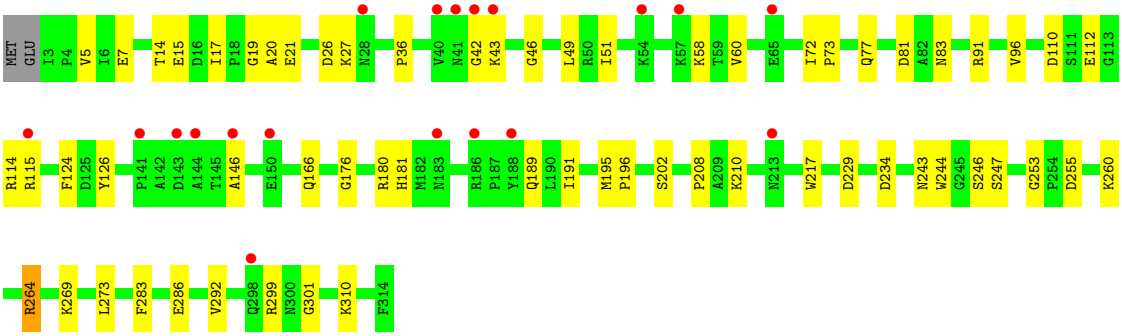
• Molecule 1: Diisopropyl-fluorophosphatase

Chain C:



• Molecule 1: Diisopropyl-fluorophosphatase

Chain D:


MET
GLU
I3
P4
V5
I6
E7
T14
E15
D16
I17
P18
G19
A20
E21
D26
K27
N28
P36
V40
N41
G42
K43
G46
L49
R50
I51
K54
K57
K58
T59
V60
E65
I72
P73
Q77
D81
A82
N83
R91
V96
D110
S111
E112
G113
R114
R115
F124
D125
Y126
P141
A142
D143
A144
T145
A146
E150
Q166
G176
R180
H181
M182
N183
R186
P187
Y188
Q189
L190
I191
M195
P196
S202
P208
A209
K210
N213
W217
D229
D234
N243
W244
G245
S246
S247
G253
P254
D255
K260
R264
K269
L273
F283
E286
V292
Q298
R299
N300
G301
K310
F314

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	66.95Å 74.49Å 118.95Å 90.00° 100.00° 90.00°	Depositor
Resolution (Å)	24.82 – 1.40 24.81 – 1.40	Depositor EDS
% Data completeness (in resolution range)	94.5 (24.82-1.40) 94.4 (24.81-1.40)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	3.85 (at 1.40Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.205 , 0.221 0.203 , 0.220	Depositor DCC
R_{free} test set	10622 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	8.9	Xtriage
Anisotropy	0.178	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.43 , 49.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10803	wwPDB-VP
Average B, all atoms (Å^2)	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.49 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.6378e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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.

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: CA

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	0/2505	0.66	0/3390
1	B	0.30	0/2500	0.66	0/3383
1	C	0.30	0/2500	0.65	0/3383
1	D	0.30	0/2500	0.66	0/3383
All	All	0.30	0/10005	0.66	0/13539

There are no bond length outliers.

There are no bond angle outliers.

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	2443	0	2362	40	0
1	B	2438	0	2360	50	0
1	C	2438	0	2360	37	0
1	D	2438	0	2360	52	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	250	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	273	0	0	3	0
3	C	270	0	0	5	0
3	D	245	0	0	3	0
All	All	10803	0	9442	170	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 9.

The worst 5 of 170 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:D:264:ARG:HB2	1:D:264:ARG:HH11	1.33	0.93
1:A:17:ILE:HD11	1:A:58:LYS:HD2	1.56	0.86
1:A:26:ASP:HB2	1:A:83:ASN:HD21	1.42	0.84
1:B:17:ILE:HD13	1:B:33:ILE:HD12	1.60	0.82
1:D:126:TYR:H	1:D:181:HIS:CE1	1.99	0.80

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	311/314 (99%)	301 (97%)	10 (3%)	0	100	100
1	B	310/314 (99%)	297 (96%)	13 (4%)	0	100	100
1	C	310/314 (99%)	299 (96%)	11 (4%)	0	100	100
1	D	310/314 (99%)	298 (96%)	12 (4%)	0	100	100
All	All	1241/1256 (99%)	1195 (96%)	46 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 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	259/261 (99%)	256 (99%)	3 (1%)	71	47
1	B	259/261 (99%)	255 (98%)	4 (2%)	65	37
1	C	259/261 (99%)	255 (98%)	4 (2%)	65	37
1	D	259/261 (99%)	254 (98%)	5 (2%)	57	25
All	All	1036/1044 (99%)	1020 (98%)	16 (2%)	65	37

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	264	ARG
1	D	243	ASN
1	C	183	ASN
1	D	115	ARG
1	C	21	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	289	ASN
1	D	189	GLN
1	C	304	GLN
1	D	83	ASN
1	D	243	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 95th 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	313/314 (99%)	0.39	13 (4%)	36	37	4, 10, 21, 32	0
1	B	312/314 (99%)	0.51	19 (6%)	21	19	4, 10, 22, 33	0
1	C	312/314 (99%)	0.36	14 (4%)	33	33	4, 9, 20, 30	0
1	D	312/314 (99%)	0.44	19 (6%)	21	19	4, 9, 21, 30	0
All	All	1249/1256 (99%)	0.43	65 (5%)	27	26	4, 10, 21, 33	0

The worst 5 of 65 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	146	ALA	10.4
1	B	144	ALA	10.0
1	C	144	ALA	9.8
1	B	152	PHE	6.1
1	B	145	THR	5.4

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 monosaccharides 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. The B-factors column lists the minimum, median, 95th 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	B-factors(Å^2)	Q<0.9
2	CA	A	315	1/1	1.00	0.07	4,4,4,4	0
2	CA	A	316	1/1	1.00	0.06	4,4,4,4	0
2	CA	B	315	1/1	1.00	0.07	3,3,3,3	0
2	CA	B	316	1/1	1.00	0.04	3,3,3,3	0
2	CA	C	315	1/1	1.00	0.06	4,4,4,4	0
2	CA	C	316	1/1	1.00	0.05	4,4,4,4	0
2	CA	D	315	1/1	1.00	0.06	3,3,3,3	0
2	CA	D	316	1/1	1.00	0.06	4,4,4,4	0

6.5 Other polymers (i)

There are no such residues in this entry.