



Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 01:19 pm BST

PDB ID : 6MT2
Title : Crystal structure of Inorganic Pyrophosphatase from *Medicago truncatula* (I23 crystal form)
Authors : Ruszkowski, M.; Grzechowiak, M.; Dauter, Z.
Deposited on : 2018-10-18
Resolution : 2.89 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

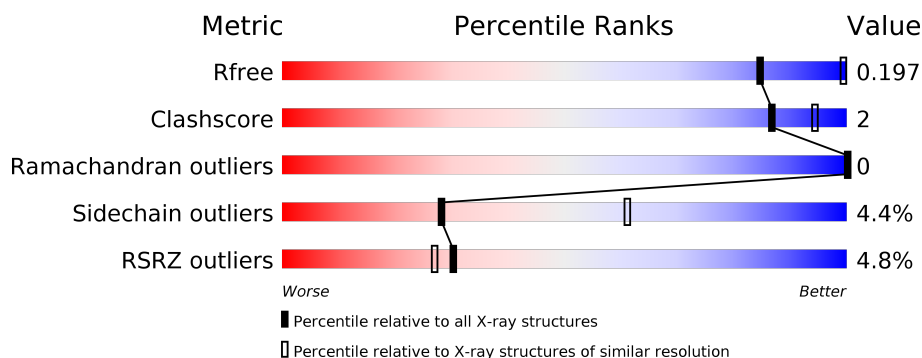
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.89 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 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 $\leq 5\%$. 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	214	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 99%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> % 78% 6% 16% </div> </div>
1	B	214	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 99%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> % 80% 5% 15% </div> </div>
1	C	214	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 5%, green 95%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 5% 71% 8% 19% </div> </div>
1	D	214	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 8%, green 92%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 8% 74% 7% 18% </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5814 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 Soluble inorganic pyrophosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	0	0	0
			1454	934	241	269	10			
1	B	181	Total	C	N	O	S	0	0	0
			1458	938	239	271	10			
1	C	173	Total	C	N	O	S	0	0	0
			1404	903	231	260	10			
1	D	175	Total	C	N	O	S	0	0	0
			1421	914	233	264	10			

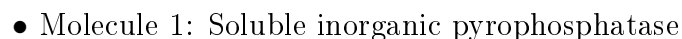
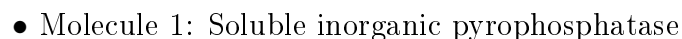
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP A0A072TMB0
A	-1	ASN	-	expression tag	UNP A0A072TMB0
B	-2	SER	-	expression tag	UNP A0A072TMB0
B	-1	ASN	-	expression tag	UNP A0A072TMB0
C	-2	SER	-	expression tag	UNP A0A072TMB0
C	-1	ASN	-	expression tag	UNP A0A072TMB0
D	-2	SER	-	expression tag	UNP A0A072TMB0
D	-1	ASN	-	expression tag	UNP A0A072TMB0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	31	Total	O	0	0
			31	31		
2	B	20	Total	O	0	0
			20	20		
2	C	17	Total	O	0	0
			17	17		
2	D	9	Total	O	0	0
			9	9		

- Molecule 1: Soluble inorganic pyrophosphatase



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α , β , γ	235.20 Å 235.20 Å 235.20 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.01 – 2.89 48.01 – 2.89	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.01-2.89) 99.7 (48.01-2.89)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 2.91 Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.172 , 0.195 0.179 , 0.197	Depositor DCC
R_{free} test set	1060 reflections (2.20%)	wwPDB-VP
Wilson B-factor (Å ²)	83.2	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 45.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.018 for -l,-k,-h	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5814	wwPDB-VP
Average B, all atoms (Å ²)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.41% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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](#)

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	0.56	0/1491	0.79	0/2021
1	B	0.51	0/1495	0.73	0/2028
1	C	0.56	0/1440	0.78	0/1950
1	D	0.50	0/1457	0.72	0/1973
All	All	0.53	0/5883	0.76	0/7972

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	1454	0	1438	8	0
1	B	1458	0	1443	2	0
1	C	1404	0	1383	9	0
1	D	1421	0	1400	6	0
2	A	31	0	0	0	0
2	B	20	0	0	0	0
2	C	17	0	0	0	0
2	D	9	0	0	0	0
All	All	5814	0	5664	24	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 2.

All (24) 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:A:207:LEU:CD1	1:A:208:HIS:CD2	2.68	0.76
1:A:207:LEU:HD12	1:A:208:HIS:CD2	2.43	0.53
1:C:126:MET:HG2	1:C:185:PHE:CE1	2.43	0.53
1:C:176:GLU:HB2	1:C:178:LYS:HE3	1.91	0.53
1:A:207:LEU:HD13	1:A:208:HIS:CD2	2.47	0.49
1:C:143:ALA:O	1:C:149:LYS:HE2	2.13	0.48
1:D:91:ILE:HD12	1:D:102:LEU:HD23	1.95	0.48
1:B:209:THR:HG22	1:B:209:THR:O	2.15	0.47
1:C:207:LEU:HD12	1:C:207:LEU:HA	1.49	0.46
1:C:82:VAL:HG13	1:C:171:ASP:HB3	1.97	0.46
1:A:196:GLN:HE22	1:C:150:HIS:HA	1.81	0.46
1:C:146:PRO:HA	1:C:149:LYS:HG2	1.96	0.46
1:D:57:LYS:HG3	1:D:86:HIS:HA	1.97	0.46
1:A:121:ARG:NH1	1:A:149:LYS:O	2.38	0.45
1:A:131:GLN:NE2	1:A:178:LYS:HD3	2.31	0.44
1:C:112:VAL:HG13	1:C:116:CYS:SG	2.57	0.44
1:A:202:TYR:O	1:A:206:ILE:HB	2.18	0.44
1:D:129:ILE:N	1:D:129:ILE:HD12	2.33	0.44
1:D:171:ASP:O	1:D:175:ASN:ND2	2.52	0.43
1:D:61:VAL:HG11	1:D:63:TYR:CZ	2.54	0.42
1:B:46:PRO:HG2	1:B:192:VAL:HG13	2.02	0.41
1:A:206:ILE:HD12	1:A:206:ILE:HA	1.88	0.41
1:D:51:CYS:O	1:D:117:PHE:HA	2.21	0.41
1:C:57:LYS:HG3	1:C:86:HIS:HA	2.02	0.40

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	178/214 (83%)	173 (97%)	5 (3%)	0	100	100
1	B	179/214 (84%)	176 (98%)	3 (2%)	0	100	100
1	C	169/214 (79%)	156 (92%)	13 (8%)	0	100	100
1	D	171/214 (80%)	163 (95%)	8 (5%)	0	100	100
All	All	697/856 (81%)	668 (96%)	29 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	161/194 (83%)	155 (96%)	6 (4%)	34	68
1	B	162/194 (84%)	155 (96%)	7 (4%)	29	62
1	C	156/194 (80%)	146 (94%)	10 (6%)	17	45
1	D	158/194 (81%)	153 (97%)	5 (3%)	39	73
All	All	637/776 (82%)	609 (96%)	28 (4%)	28	61

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	ARG
1	A	134	LYS
1	A	161	ARG
1	A	172	TYR
1	A	193	GLU
1	A	207	LEU
1	B	31	VAL
1	B	65	LEU
1	B	125	LEU
1	B	134	LYS
1	B	167	ARG
1	B	172	TYR
1	B	207	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	65	LEU
1	C	78	LEU
1	C	116	CYS
1	C	126	MET
1	C	161	ARG
1	C	163	MET
1	C	166	LYS
1	C	178	LYS
1	C	207	LEU
1	C	209	THR
1	D	80	SER
1	D	149	LYS
1	D	167	ARG
1	D	183	ASN
1	D	210	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	131	GLN
1	A	196	GLN
1	A	208	HIS
1	C	109	GLN
1	C	160	HIS
1	D	175	ASN

5.3.3 RNA ⓘ

There are no RNA molecules 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

There are no ligands in this entry.

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, 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	180/214 (84%)	0.07	3 (1%) 70 69	51, 64, 94, 147	0
1	B	181/214 (84%)	0.31	3 (1%) 70 69	61, 81, 111, 165	0
1	C	173/214 (80%)	0.40	11 (6%) 19 15	53, 71, 133, 153	0
1	D	175/214 (81%)	0.65	17 (9%) 7 6	74, 96, 148, 167	0
All	All	709/856 (82%)	0.36	34 (4%) 30 27	51, 80, 131, 167	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	179	GLU	4.6
1	D	131	GLN	4.4
1	D	133	GLU	4.2
1	D	132	GLY	4.2
1	D	129	ILE	4.2
1	C	208	HIS	4.1
1	D	177	ASN	4.1
1	B	32	ALA	3.6
1	D	130	ASP	3.4
1	D	178	LYS	3.3
1	D	134	LYS	3.2
1	A	208	HIS	3.2
1	B	30	SER	3.2
1	D	174	LYS	3.1
1	C	129	ILE	3.1
1	C	131	GLN	3.0
1	C	205	TYR	2.9
1	C	133	GLU	2.8
1	C	209	THR	2.8
1	D	175	ASN	2.7
1	D	126	MET	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	59	SER	2.3
1	C	132	GLY	2.3
1	B	207	LEU	2.2
1	A	32	ALA	2.2
1	A	30	SER	2.1
1	D	128	MET	2.1
1	C	183	ASN	2.1
1	C	206	ILE	2.0
1	D	58	GLY	2.0
1	D	176	GLU	2.0
1	D	191	ALA	2.0
1	C	177	ASN	2.0
1	D	183	ASN	2.0

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.