W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 16, 2023 – 04:01 PM EDT

PDB ID	:	2APJ
Title	:	X-Ray Structure of Protein from Arabidopsis Thaliana AT4G34215 at 1.6 Angstrom Resolution
Authors	:	Wesenberg, G.E.; Phillips Jr., G.N.; Mccoy, J.G.; Bitto, E.; Bingman, C.A.; Allard, S.T.; Center for Eukaryotic Structural Genomics (CESG)
Deposited on	:	2005-08-16
Resolution	:	1.60 Å(reported)

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

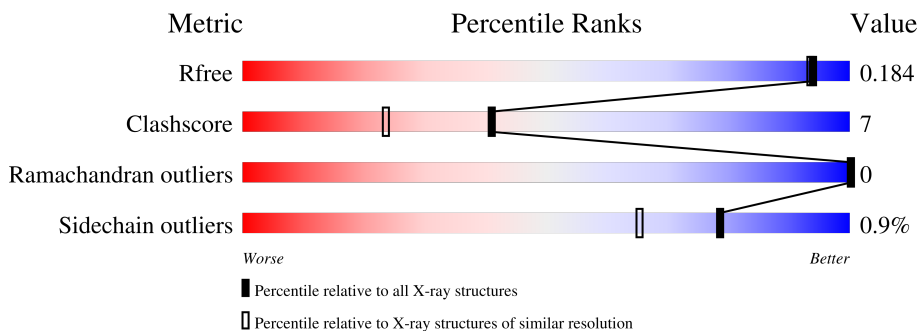|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance ⓘ

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

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	Percentile Ranks	Value
Rfree		0.184
Clashscore		7
Ramachandran outliers		0
Sidechain outliers		0.9%

Worse   *Better*

█ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (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 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%

Mol	Chain	Length	Quality of chain
1	A	260	83% 9% · 6%
1	B	260	87% 5% 7%
1	C	260	85% 7% · 7%
1	D	260	84% 6% · 8%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	SEB	A	31	-	-	X	-
1	SEB	B	31	-	-	X	-

2　Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 9132 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 Putative Esterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	244	Total	C	N	O	S	0	6	0
			1906	1201	336	358	11			
1	B	243	Total	C	N	O	S	0	8	0
			1901	1197	335	358	11			
1	C	241	Total	C	N	O	S	0	2	0
			1869	1172	333	353	11			
1	D	238	Total	C	N	O	S	0	2	0
			1845	1159	330	345	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	31	SEB	SER	modified residue	UNP O49483
B	31	SEB	SER	modified residue	UNP O49483
C	31	SEB	SER	modified residue	UNP O49483
D	31	SEB	SER	modified residue	UNP O49483

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	464	Total	O	0	0
			464	464		
2	B	443	Total	O	0	0
			443	443		
2	C	348	Total	O	0	0
			348	348		
2	D	356	Total	O	0	0
			356	356		

3 Residue-property plots ⓘ

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. 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. 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.
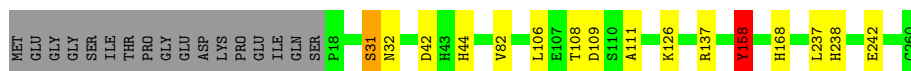
- Molecule 1: Putative Esterase

Chain A:
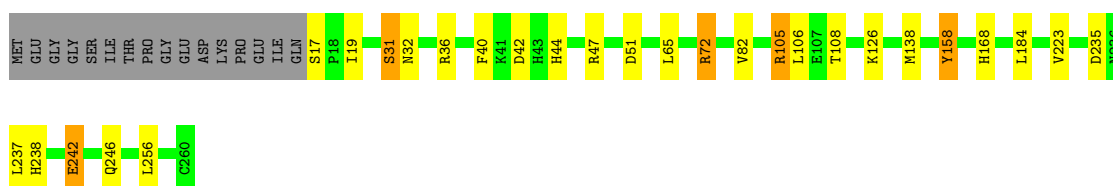
- Molecule 1: Putative Esterase

Chain B:

- Molecule 1: Putative Esterase

Chain C:

- Molecule 1: Putative Esterase

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	40.69Å 71.92Å 93.21Å 108.16° 93.25° 90.40°	Depositor
Resolution (Å)	35.85 – 1.60 35.85 – 1.60	Depositor EDS
% Data completeness (in resolution range)	94.4 (35.85-1.60) 94.4 (35.85-1.60)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.74 (at 1.60Å)	Xtriage
Refinement program	REFMAC refmac_5.2.0005	Depositor
R, R_{free}	0.146 , 0.183 0.148 , 0.184	Depositor DCC
R_{free} test set	6226 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	15.1	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 31.5	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	9132	wwPDB-VP
Average B, all atoms (Å^2)	8.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.61% 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.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SEB

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	0.96	4/1952 (0.2%)	0.89	5/2645 (0.2%)
1	B	0.90	4/1954 (0.2%)	0.87	2/2647 (0.1%)
1	C	0.72	0/1898	0.73	1/2568 (0.0%)
1	D	0.73	0/1872	0.74	1/2531 (0.0%)
All	All	0.84	8/7676 (0.1%)	0.81	9/10391 (0.1%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	242	GLU	CG-CD	7.70	1.63	1.51
1	A	242	GLU	CG-CD	6.50	1.61	1.51
1	A	105	ARG	CB-CG	6.29	1.69	1.52
1	B	242	GLU	CB-CG	6.26	1.64	1.52
1	B	126	LYS	CD-CE	6.02	1.66	1.51

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	126	LYS	CD-CE-NZ	6.18	125.91	111.70
1	B	158	TYR	CB-CG-CD2	-6.06	117.36	121.00
1	A	126	LYS	CD-CE-NZ	5.98	125.45	111.70
1	A	47	ARG	NE-CZ-NH1	-5.87	117.37	120.30
1	C	166	ASP	CB-CG-OD1	5.59	123.33	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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	1906	0	1892	37	0
1	B	1901	0	1892	34	0
1	C	1869	0	1854	20	0
1	D	1845	0	1834	15	0
2	A	464	0	0	17	1
2	B	443	0	0	24	2
2	C	348	0	0	13	0
2	D	356	0	0	7	0
All	All	9132	0	7472	106	2

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 7.

The worst 5 of 106 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:168:HIS:CE1	2:D:289:HOH:O	1.89	1.22
1:A:31:SEB:HE3	2:A:451:HOH:O	1.49	1.12
1:A:168:HIS:CE1	2:A:361:HOH:O	2.04	1.10
1:B:82[B]:VAL:HG22	2:B:540:HOH:O	1.48	1.09
1:B:31:SEB:HE3	2:B:345:HOH:O	1.51	1.09

All (2) 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)
2:A:526:HOH:O	2:B:405:HOH:O[1_455]	2.18	0.02
2:B:475:HOH:O	2:B:629:HOH:O[1_655]	2.18	0.02

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	247/260 (95%)	243 (98%)	4 (2%)	0	100	100
1	B	248/260 (95%)	242 (98%)	6 (2%)	0	100	100
1	C	240/260 (92%)	236 (98%)	4 (2%)	0	100	100
1	D	235/260 (90%)	231 (98%)	4 (2%)	0	100	100
All	All	970/1040 (93%)	952 (98%)	18 (2%)	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	208/215 (97%)	207 (100%)	1 (0%)	88	80
1	B	209/215 (97%)	208 (100%)	1 (0%)	88	80
1	C	201/215 (94%)	199 (99%)	2 (1%)	76	61
1	D	198/215 (92%)	195 (98%)	3 (2%)	65	44
All	All	816/860 (95%)	809 (99%)	7 (1%)	78	65

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

Mol	Chain	Res	Type
1	C	242	GLU
1	D	110	SER
1	D	192	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	158	TYR
1	C	158	TYR

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

Mol	Chain	Res	Type
1	D	133	HIS

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	#$|Z| > 2$
1	SEB	C	31	1	15,16,17	2.27	4 (26%)	15,21,23	1.62	4 (26%)
1	SEB	A	31	1	15,16,17	2.34	4 (26%)	15,21,23	1.91	5 (33%)
1	SEB	B	31	1	15,16,17	2.37	5 (33%)	15,21,23	1.03	1 (6%)
1	SEB	D	31	1	15,16,17	1.93	4 (26%)	15,21,23	1.32	1 (6%)

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
1	SEB	C	31	1	-	5/9/13/15	0/1/1/1
1	SEB	A	31	1	-	5/9/13/15	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	SEB	B	31	1	-	5/9/13/15	0/1/1/1
1	SEB	D	31	1	-	5/9/13/15	0/1/1/1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	31	SEB	CE-CZ	5.83	1.60	1.50
1	C	31	SEB	CE-CZ	5.53	1.60	1.50
1	B	31	SEB	OD1-SD	5.03	1.56	1.44
1	B	31	SEB	CE-SD	-4.88	1.73	1.78
1	A	31	SEB	OD1-SD	4.31	1.55	1.44

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	31	SEB	CE-CZ-CH1	4.48	126.13	120.54
1	C	31	SEB	CE-CZ-CH1	3.13	124.44	120.54
1	D	31	SEB	CB-OG-SD	-3.01	112.76	119.23
1	A	31	SEB	CB-OG-SD	-2.81	113.20	119.23
1	C	31	SEB	OG-SD-CE	2.45	110.92	104.18

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	31	SEB	SD-CE-CZ-CH2
1	A	31	SEB	SD-CE-CZ-CH1
1	A	31	SEB	CB-OG-SD-CE
1	A	31	SEB	CB-OG-SD-OD2
1	B	31	SEB	SD-CE-CZ-CH2

There are no ring outliers.

4 monomers are involved in 27 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	31	SEB	3	0
1	A	31	SEB	11	0
1	B	31	SEB	10	0
1	D	31	SEB	3	0

5.5 Carbohydrates ⓘ

There are no monosaccharides 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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.