wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2023 – 06:57 PM EDT

PDB ID	:	3VX8
Title	:	Crystal structure of Arabidopsis thaliana Atg7NTD-Atg3 complex
Authors	:	Matoba, K.; Fujioka, Y.; Noda, N.N.
Deposited on	:	2012-09-11
Resolution	:	3.11 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 ⓘ 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 ⓘ) were used in the production of this report:

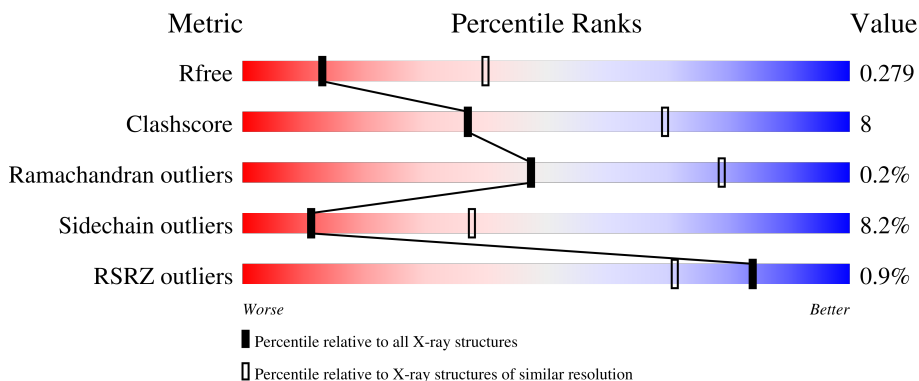|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.11 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		0.279
Clashscore		8
Ramachandran outliers		0.2%
Sidechain outliers		8.2%
RSRZ 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	1292 (3.14-3.10)
Clashscore	141614	1389 (3.14-3.10)
Ramachandran outliers	138981	1337 (3.14-3.10)
Sidechain outliers	138945	1337 (3.14-3.10)
RSRZ outliers	127900	1260 (3.14-3.10)

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	323	73% 18% • 6%
1	D	323	72% 19% • 7%
2	B	292	48% 11% • 40%
2	C	292	47% 10% • 41%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 7608 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 Ubiquitin-like modifier-activating enzyme atg7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	300	Total	C	N	O	S	0	0	0
			2380	1533	396	444	7			
1	A	305	Total	C	N	O	S	0	0	0
			2418	1557	403	451	7			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	3	GLY	-	expression tag	UNP Q94CD5
D	4	PRO	-	expression tag	UNP Q94CD5
D	5	HIS	-	expression tag	UNP Q94CD5
D	6	MET	-	expression tag	UNP Q94CD5
A	3	GLY	-	expression tag	UNP Q94CD5
A	4	PRO	-	expression tag	UNP Q94CD5
A	5	HIS	-	expression tag	UNP Q94CD5
A	6	MET	-	expression tag	UNP Q94CD5

- Molecule 2 is a protein called Autophagy-related protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	175	Total	C	N	O	S	0	0	0
			1416	909	238	259	10			
2	C	172	Total	C	N	O	S	0	0	0
			1391	892	233	256	10			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	22	GLY	-	expression tag	UNP Q0WWQ1
B	23	PRO	-	expression tag	UNP Q0WWQ1
B	24	HIS	-	expression tag	UNP Q0WWQ1
B	25	MET	-	expression tag	UNP Q0WWQ1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	22	GLY	-	expression tag	UNP Q0WWQ1
C	23	PRO	-	expression tag	UNP Q0WWQ1
C	24	HIS	-	expression tag	UNP Q0WWQ1
C	25	MET	-	expression tag	UNP Q0WWQ1

- Molecule 3 is water.

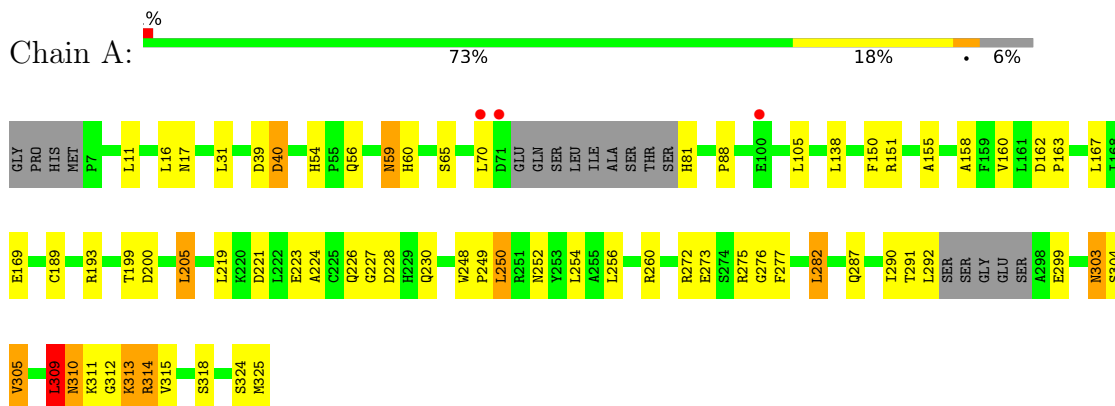| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	O 1	0	0
3	B	1	Total 1	O 1	0	0
3	C	1	Total 1	O 1	0	0

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 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.
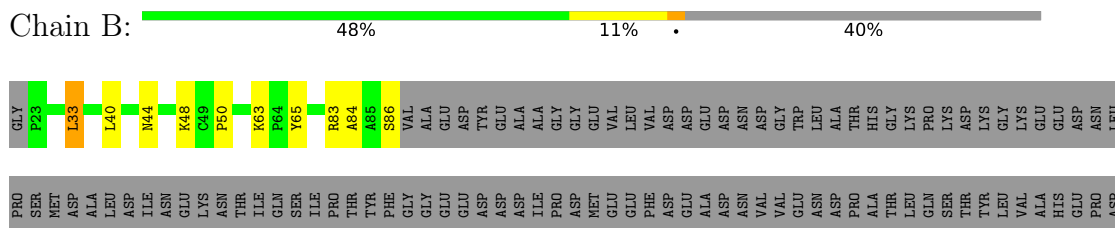
• Molecule 1: Ubiquitin-like modifier-activating enzyme atg7
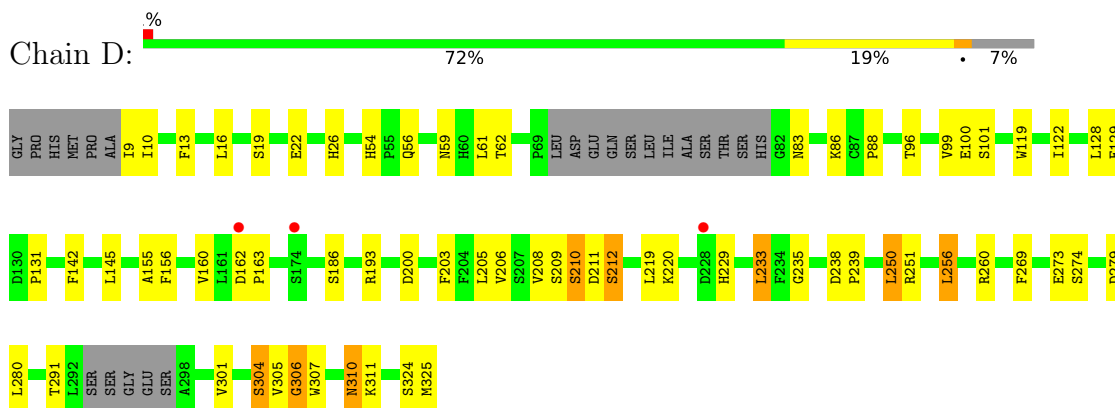
• Molecule 1: Ubiquitin-like modifier-activating enzyme atg7

• Molecule 2: Autophagy-related protein 3

ASP ASP ASN ILE L190 R193 R209 V210 W211 L212 D216 E217 S218 R219 L222 E225 L226 V227 D234 HIS ALA ARG LYS T239 E243 D244 H245 G250 K251 H252 P257 H260 M264 K265 K266 I267 I268 D269 V270 V280 Y283 D301 M304 ASP PHE

ASP LEU GLY SER SER THR

• Molecule 2: Autophagy-related protein 3

Chain C:

%

47%   10% • 41%

GLY PRO HIS M25 L33 T51 L81 R82 R83 S86 VAL ALA GLU ASP TYR GLU ALA ALA GLY GLY VAL LEU VAL ASP ASP GLU ASP ASN ASP GLY TRP LEU ALA THR HIS GLY LYS PRO LYS ASP LYS GLY LYS GLU GLU ASP ASN LEU PRO SER MET ASP ALA LEU ASP

ILE ASN GLU LYS ASN THR ILE GLN SER ILE PRO THR TYR PHE GLY GLY GLU GLU GLU ASP ASP ILE PRO ASP MET GLU GLU PHE ASP GLU ALA ASP ASN VAL VAL GLU ASN ASP PRO ALA THR LEU GLN SER THR TYR LEU VAL ALA HIS GLU PRO ASP ASP ASP ASN ILE LEU R191

T194 L197 S198 I199 K203 Y204 P208 R209 L212 T213 G214 Y215 D216 E217 S218 R219 E225 L226 V227 V231 D234 HIS ALA ARG LYS T239 V240 T241 I242 E243 D244 H245 K251 H252 A253 S254 P257 H260 M264 I268 E277 F287 L288 D301

M304 ASP PHE ASP LEU GLY SER SER THR

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	101.52Å 132.68Å 102.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.85 – 3.11 34.85 – 3.11	Depositor EDS
% Data completeness (in resolution range)	99.1 (34.85-3.11) 99.2 (34.85-3.11)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.69 (at 3.12Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.226 , 0.278 0.226 , 0.279	Depositor DCC
R_{free} test set	1295 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	44.6	Xtriage
Anisotropy	0.076	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 20.7	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.018 for l,-k,h	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	7608	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.51% 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 ⓘ

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	0.50	0/2487	0.60	2/3383 (0.1%)
1	D	0.39	0/2447	0.54	0/3328
2	B	0.28	0/1452	0.51	0/1966
2	C	0.28	0/1425	0.49	0/1929
All	All	0.40	0/7811	0.55	2/10606 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

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	A	227	GLY	N-CA-C	6.10	128.35	113.10
1	A	309	LEU	CA-CB-CG	5.55	128.07	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	305	VAL	Peptide

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	2418	0	2358	47	0
1	D	2380	0	2323	45	0
2	B	1416	0	1412	15	0
2	C	1391	0	1386	17	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
All	All	7608	0	7479	118	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 8.

The worst 5 of 118 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:210:SER:HB2	1:D:229:HIS:O	1.55	1.06
1:A:228:ASP:OD1	1:A:230:GLN:HB3	1.59	1.03
1:D:304:SER:O	1:D:305:VAL:HG23	1.65	0.97
2:B:193:ARG:HD2	2:B:222:LEU:HD11	1.51	0.92
1:A:230:GLN:O	1:A:230:GLN:HG3	1.72	0.89

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	299/323 (93%)	284 (95%)	14 (5%)	1 (0%)	41	73
1	D	294/323 (91%)	275 (94%)	18 (6%)	1 (0%)	41	73
2	B	169/292 (58%)	164 (97%)	5 (3%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	166/292 (57%)	161 (97%)	5 (3%)	0	100	100
All	All	928/1230 (75%)	884 (95%)	42 (4%)	2 (0%)	47	79

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	314	ARG
1	D	306	GLY

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	271/286 (95%)	247 (91%)	24 (9%)	9	34
1	D	267/286 (93%)	250 (94%)	17 (6%)	17	47
2	B	160/259 (62%)	144 (90%)	16 (10%)	7	28
2	C	157/259 (61%)	144 (92%)	13 (8%)	11	37
All	All	855/1090 (78%)	785 (92%)	70 (8%)	11	38

5 of 70 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	81	LEU
2	C	194	THR
2	C	219	ARG
1	A	105	LEU
1	A	65	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	116	ASN
1	A	287	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	247	HIS
1	A	310	ASN
1	A	230	GLN

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 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 ⓘ

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(Å²)	Q<0.9
1	A	305/323 (94%)	-0.43	3 (0%)	82	69	18, 32, 56, 87	0
1	D	300/323 (92%)	0.01	3 (1%)	82	69	33, 56, 85, 108	0
2	B	175/292 (59%)	-0.40	1 (0%)	89	79	18, 33, 58, 89	0
2	C	172/292 (58%)	-0.20	2 (1%)	79	63	33, 49, 74, 102	0
All	All	952/1230 (77%)	-0.25	9 (0%)	84	71	18, 43, 79, 108	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	228	ASP	4.4
1	A	70	LEU	3.5
1	D	174	SER	3.3
1	D	162	ASP	3.2
2	C	277	GLU	2.9

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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

6.5 Other polymers ⓘ

There are no such residues in this entry.