



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 18, 2023 – 08:29 AM EDT

PDB ID : 1UM2
Title : Crystal Structure of the Vma1-Derived Endonuclease with the Ligated Extein Segment
Authors : Mizutani, R.; Anraku, Y.; Satow, Y.
Deposited on : 2003-09-22
Resolution : 2.90 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
Xtrriage (Phenix) : **NOT EXECUTED**
EDS : **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

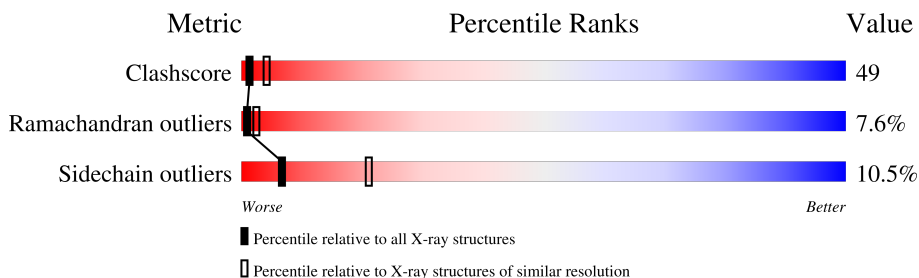
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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))
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (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 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 $\leq 5\%$

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	454	 35% 47% 9% • 8%
1	B	454	 34% 48% 10% • 8%
2	C	21	 10% 5% 14% 5% 67%
2	D	21	 5% 14% 14% 67%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6865 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 ENDONUCLEASE PI-SCEI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	417	Total 3293	C 2089	N 561	O 631	S 12	0	0	0
1	B	419	Total 3313	C 2101	N 566	O 634	S 12	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	284	SER	CYS	engineered mutation	UNP P17255
A	362	ASN	HIS	engineered mutation	UNP P17255
B	284	SER	CYS	engineered mutation	UNP P17255
B	362	ASN	HIS	engineered mutation	UNP P17255

- Molecule 2 is a protein called 21-mer from Vacuolar ATP synthase catalytic subunit A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	7	Total 53	C 32	N 10	O 11	0	0	0
2	D	7	Total 53	C 32	N 10	O 11	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	273	MET	-	initiating methionine	UNP P17255
C	738	SER	CYS	engineered mutation	UNP P17255
D	273	MET	-	initiating methionine	UNP P17255
D	738	SER	CYS	engineered mutation	UNP P17255

- Molecule 3 is water.

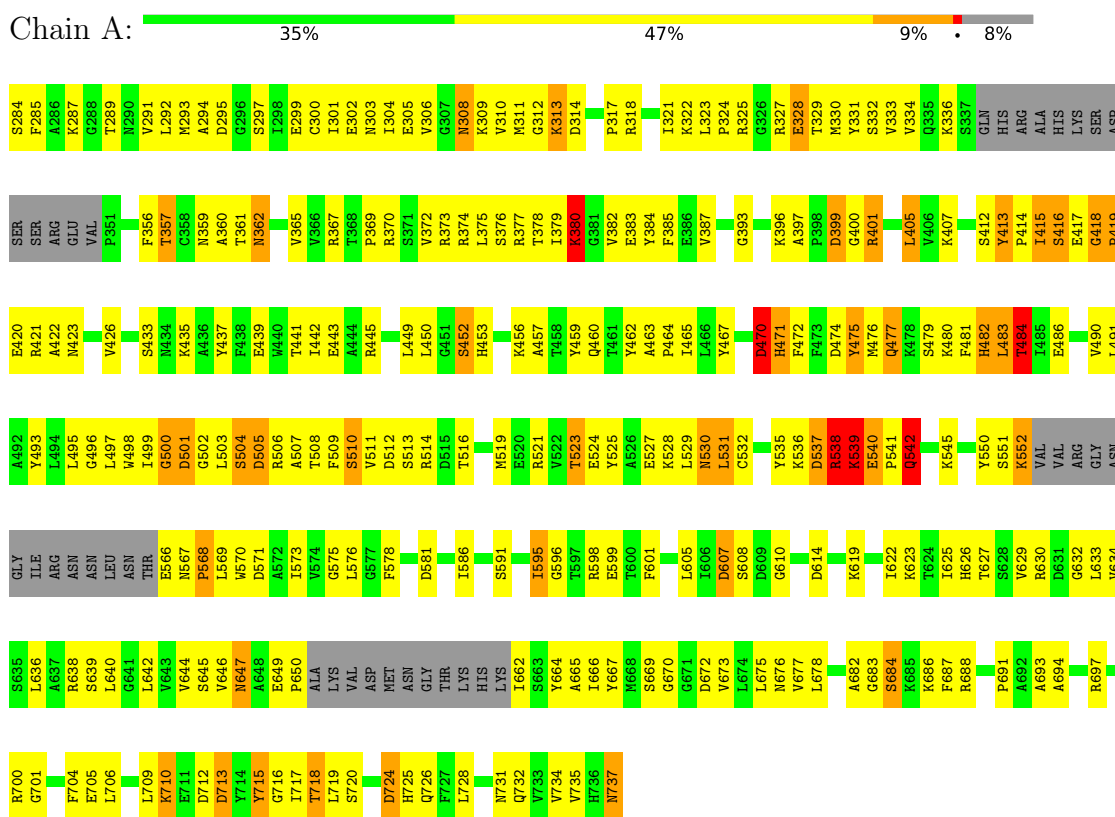
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	74	Total O 74 74	0	0
3	B	79	Total O 79 79	0	0

3 Residue-property plots

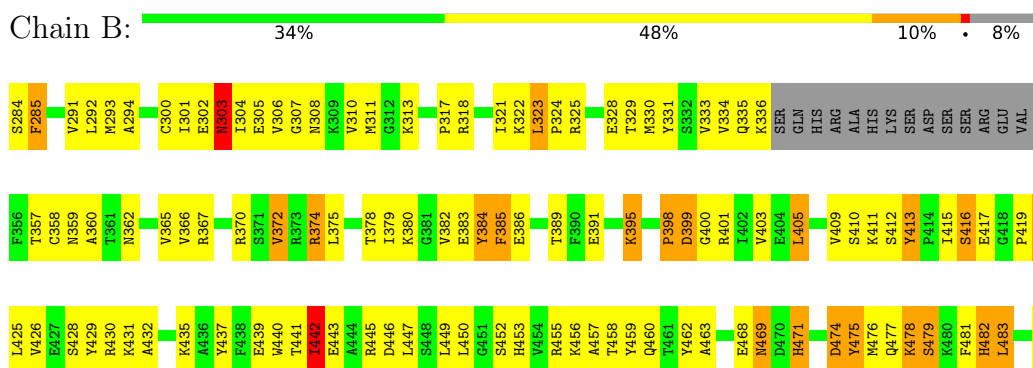
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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.

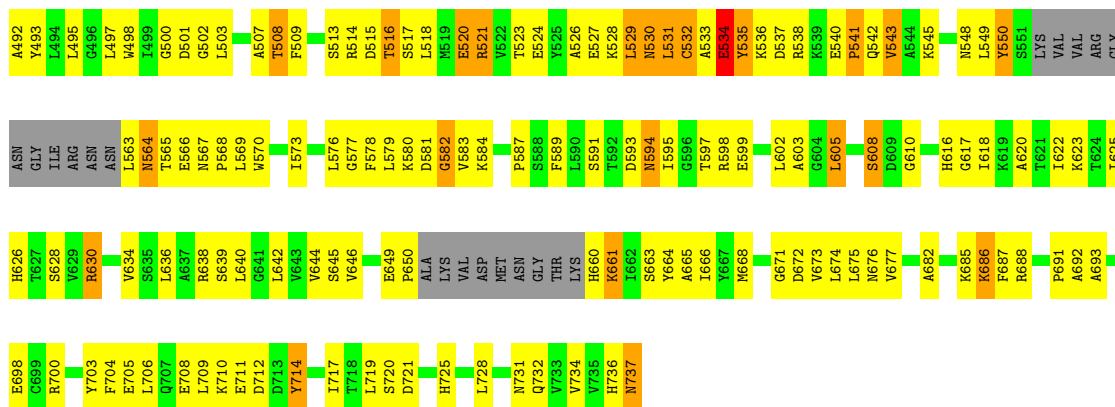
Note EDS was not executed.

- Molecule 1: ENDONUCLEASE PI-SCEI

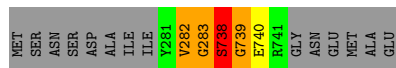


- Molecule 1: ENDONUCLEASE PI-SCEI

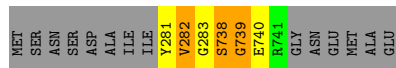




• Molecule 2: 21-mer from Vacuolar ATP synthase catalytic subunit A



• Molecule 2: 21-mer from Vacuolar ATP synthase catalytic subunit A



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	66.40Å 70.40Å 58.00Å 100.40° 98.70° 78.90°	Depositor
Resolution (Å)	30.00 – 2.90	Depositor
% Data completeness (in resolution range)	(Not available) (30.00-2.90)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.214 , 0.283	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6865	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

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	0/3353	0.80	0/4523
1	B	0.62	1/3373 (0.0%)	0.81	0/4551
2	C	1.11	0/53	1.15	1/69 (1.4%)
2	D	0.92	0/53	1.09	1/69 (1.4%)
All	All	0.60	1/6832 (0.0%)	0.81	2/9212 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	534	GLU	CB-CG	5.27	1.62	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	739	GLY	N-CA-C	-5.71	98.84	113.10
2	D	739	GLY	N-CA-C	-5.19	100.12	113.10

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	3293	0	3291	308	0
1	B	3313	0	3309	337	0
2	C	53	0	47	23	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	53	0	47	11	0
3	A	74	0	0	11	0
3	B	79	0	0	13	0
All	All	6865	0	6694	653	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 49.

The worst 5 of 653 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:B:493:TYR:HE2	1:B:497:LEU:HD12	1.13	1.13
1:A:595:ILE:HG13	1:A:598:ARG:HH21	1.13	1.11
1:A:507:ALA:HB1	1:A:569:LEU:HD23	1.40	1.04
1:B:493:TYR:CE2	1:B:497:LEU:HD12	1.92	1.04
1:B:284:SER:HB2	1:B:359:ASN:ND2	1.75	1.01

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	409/454 (90%)	311 (76%)	65 (16%)	33 (8%)	1	2
1	B	411/454 (90%)	326 (79%)	59 (14%)	26 (6%)	1	4
2	C	5/21 (24%)	2 (40%)	0	3 (60%)	0	0
2	D	5/21 (24%)	3 (60%)	1 (20%)	1 (20%)	0	0
All	All	830/950 (87%)	642 (77%)	125 (15%)	63 (8%)	1	2

5 of 63 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	416	SER
1	A	452	SER
1	A	470	ASP
1	A	505	ASP
1	A	510	SER

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	360/392 (92%)	321 (89%)	39 (11%)	6	20
1	B	362/392 (92%)	326 (90%)	36 (10%)	8	24
2	C	5/16 (31%)	4 (80%)	1 (20%)	1	4
2	D	5/16 (31%)	4 (80%)	1 (20%)	1	4
All	All	732/816 (90%)	655 (90%)	77 (10%)	7	21

5 of 77 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	520	GLU
1	B	712	ASP
1	B	529	LEU
1	B	594	ASN
2	C	738	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	308	ASN
1	B	548	ASN
1	B	737	ASN
1	A	471	HIS
1	A	647	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](#)

There are no ligands in this entry.

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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.