Full wwPDB X-ray Structure Validation Report (i)

Nov 2, 2024 – 09:39 AM EDT

PDB ID	:	1OUO
Title	:	Crystal structure of the periplasmic endonuclease Vvn
Authors	:	Yuan, H.S.; Li, C.L.
Deposited on	:	2003-03-25
Resolution	:	2.30 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 (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:

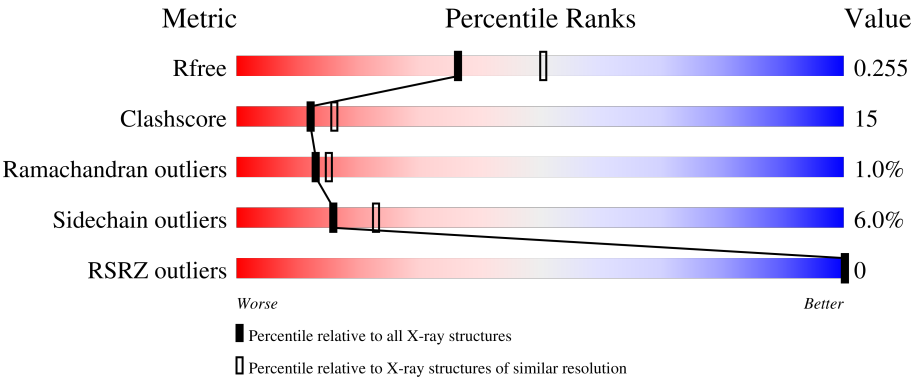
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.30 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}	164625	5963 (2.30-2.30)
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (2.30-2.30)
RSRZ outliers	164620	5963 (2.30-2.30)

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	210	70% 26% .

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1913 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 Nuclease.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	210	1718	1069	321	315	8	5	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	112	MSE	MET	modified residue	UNP Q8DCA6
A	151	MSE	MET	modified residue	UNP Q8DCA6
A	161	MSE	MET	modified residue	UNP Q8DCA6
A	163	PRO	GLN	SEE REMARK 999	UNP Q8DCA6
A	164	ASP	THR	SEE REMARK 999	UNP Q8DCA6
A	165	ARG	GLU	SEE REMARK 999	UNP Q8DCA6
A	166	ALA	LEU	SEE REMARK 999	UNP Q8DCA6
A	177	MSE	MET	modified residue	UNP Q8DCA6
A	193	MSE	MET	modified residue	UNP Q8DCA6

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
2	A	1	1	1	0	0

- Molecule 3 is water.

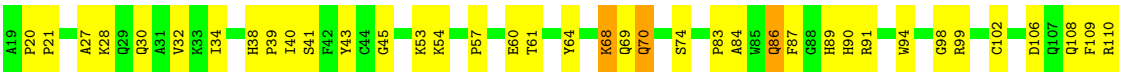
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	194	194	194	0	0

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: Nuclease

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	49.12Å 40.10Å 50.63Å 90.00° 90.53° 90.00°	Depositor
Resolution (Å)	26.55 – 2.30 26.55 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.4 (26.55-2.30) 98.3 (26.55-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	6.24 (at 2.26Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.186 , 0.269 0.188 , 0.255	Depositor DCC
R_{free} test set	731 reflections (8.30%)	wwPDB-VP
Wilson B-factor (Å^2)	19.5	Xtriage
Anisotropy	0.163	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 56.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.000 for l,k,-h 0.044 for h,-k,-l 0.027 for l,-k,h	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	1913	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.37% 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: MG

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	0.60	2/1759 (0.1%)	0.63	0/2363

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	161	MSE	SE-CE	-6.65	1.56	1.95
1	A	151	MSE	SE-CE	-5.12	1.65	1.95

There are no bond angle outliers.

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	1718	0	1619	51	0
2	A	1	0	0	0	0
3	A	194	0	0	9	0
All	All	1913	0	1619	51	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 15.

All (51) 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:161:MSE:HE2	1:A:167:ARG:HH22	1.28	0.97
1:A:137:GLN:HE21	1:A:196:TRP:HE1	1.23	0.84
1:A:153:VAL:HG22	1:A:160:VAL:HG12	1.62	0.81
1:A:161:MSE:HE2	1:A:167:ARG:NH2	1.96	0.78
1:A:137:GLN:NE2	1:A:196:TRP:HE1	1.84	0.76
1:A:106:ASP:O	1:A:110:ARG:HG3	1.87	0.75
1:A:221:ASN:HD22	1:A:221:ASN:C	1.94	0.71
1:A:151:MSE:HG2	1:A:163:PRO:HD3	1.75	0.68
1:A:54:LYS:HE2	3:A:421:HOH:O	1.98	0.64
1:A:84:ALA:HA	1:A:87:PHE:CZ	2.34	0.62
1:A:89:HIS:HA	1:A:94:TRP:CG	2.35	0.62
1:A:186:SER:OG	1:A:188:GLN:HG3	2.00	0.61
1:A:153:VAL:HG22	1:A:160:VAL:CG1	2.31	0.60
1:A:30:GLN:HE21	1:A:216:ILE:HD13	1.67	0.59
1:A:20:PRO:HD3	3:A:376:HOH:O	2.02	0.58
1:A:221:ASN:C	1:A:221:ASN:ND2	2.60	0.56
1:A:43:TYR:CD2	1:A:151:MSE:HE2	2.42	0.55
1:A:161:MSE:CE	1:A:167:ARG:HH22	2.11	0.55
1:A:91:ARG:HH21	1:A:112:MSE:CE	2.23	0.52
1:A:57:PRO:HD2	1:A:74:SER:O	2.11	0.50
1:A:30:GLN:HE21	1:A:216:ILE:HG21	1.75	0.50
1:A:90:HIS:HB2	3:A:355:HOH:O	2.12	0.50
1:A:221:ASN:HD22	1:A:222:PRO:N	2.09	0.50
1:A:30:GLN:O	1:A:34:ILE:HG13	2.11	0.49
1:A:83:PRO:O	1:A:86:GLN:HB2	2.12	0.49
1:A:188:GLN:HG3	1:A:189:GLN:H	1.78	0.49
1:A:70:GLN:H	1:A:70:GLN:NE2	2.11	0.49
1:A:86:GLN:HG2	3:A:469:HOH:O	2.11	0.49
1:A:132:ASN:C	1:A:132:ASN:HD22	2.16	0.48
1:A:30:GLN:NE2	1:A:216:ILE:HD13	2.28	0.47
1:A:106:ASP:HB3	1:A:109:PHE:HB3	1.96	0.47
1:A:94:TRP:CH2	1:A:99:ARG:HB2	2.50	0.47
1:A:21:PRO:HG2	1:A:27:ALA:HA	1.97	0.47
1:A:69:GLN:HB3	3:A:381:HOH:O	2.14	0.47
1:A:139:ASN:ND2	3:A:313:HOH:O	2.48	0.46
1:A:64:TYR:CD1	1:A:123:ILE:HD13	2.51	0.45
1:A:91:ARG:HD3	1:A:106:ASP:OD1	2.16	0.45
1:A:41:SER:O	1:A:45:GLY:HA2	2.17	0.45
1:A:28:LYS:O	1:A:32:VAL:HG23	2.17	0.45
1:A:202:VAL:HA	1:A:206:GLU:OE1	2.17	0.45
1:A:21:PRO:HD2	1:A:114:ALA:O	2.16	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:102:CYS:HB3	1:A:109:PHE:CD2	2.52	0.44
1:A:38:HIS:HE1	3:A:410:HOH:O	2.00	0.44
1:A:86:GLN:HE21	1:A:86:GLN:HB3	1.60	0.42
1:A:196:TRP:HD1	3:A:475:HOH:O	2.03	0.42
1:A:39:PRO:C	1:A:40:ILE:HD12	2.40	0.41
1:A:91:ARG:HE	1:A:112:MSE:HE1	1.85	0.41
1:A:162:PRO:HA	1:A:163:PRO:HD3	1.96	0.41
1:A:91:ARG:HH21	1:A:112:MSE:HE1	1.86	0.41
1:A:130:ARG:C	1:A:130:ARG:HD3	2.42	0.41
1:A:61:THR:HG21	3:A:302:HOH:O	2.21	0.40

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	208/210 (99%)	197 (95%)	9 (4%)	2 (1%)	13 15

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	68	LYS
1	A	98	GLY

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	183/178 (103%)	172 (94%)	11 (6%)	16 23

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	53	LYS
1	A	60	GLU
1	A	68	LYS
1	A	70	GLN
1	A	86	GLN
1	A	108	GLN
1	A	130	ARG
1	A	132	ASN
1	A	151	MSE
1	A	175	LEU
1	A	221	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	29	GLN
1	A	30	GLN
1	A	36	GLN
1	A	38	HIS
1	A	51	GLN
1	A	70	GLN
1	A	80	HIS
1	A	132	ASN
1	A	134	ASN
1	A	137	GLN
1	A	139	ASN
1	A	179	GLN
1	A	184	GLN
1	A	190	GLN
1	A	221	ASN
1	A	225	GLN

5.3.3 RNA

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is 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

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, 95^{th} 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	205/210 (97%)	-0.34	0 100 100	11, 22, 37, 43	0

There are no RSRZ outliers to report.

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

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, 95^{th} 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	MG	A	494	1/1	0.99	0.05	15,15,15,15	0

6.5 Other polymers

There are no such residues in this entry.