wwPDB X-ray Structure Validation Summary Report (i)

May 7, 2025 – 01:09 pm BST

PDB ID	:	9FYV / pdb_00009fyv
Title	:	Crystal structure of the engineered photoenzyme VEnT1.3
Authors	:	Hardy, F.J.; Roberts, G.W.
Deposited on	:	2024-07-04
Resolution	:	1.82 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 (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:

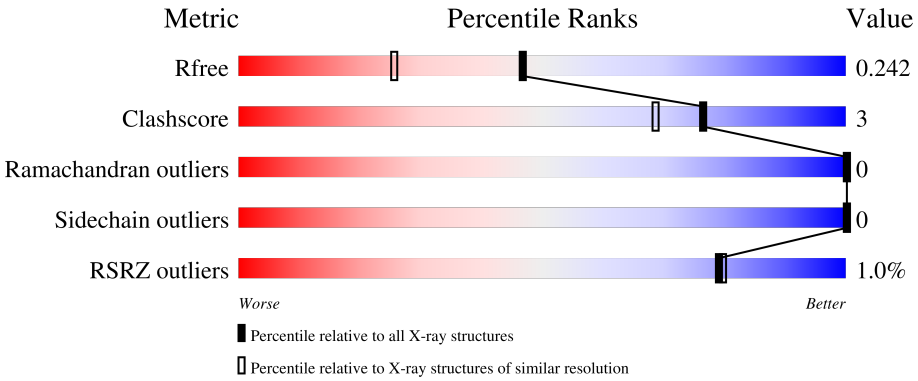
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

1 Overall quality at a glance (i)

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

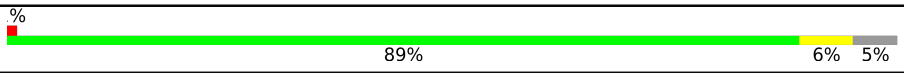
The reported resolution of this entry is 1.82 Å.

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	9242 (1.84-1.80)
Clashscore	180529	1080 (1.82-1.82)
Ramachandran outliers	177936	1073 (1.82-1.82)
Sidechain outliers	177891	1073 (1.82-1.82)
RSRZ outliers	164620	9241 (1.84-1.80)

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	326	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5269 atoms, of which 2524 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.

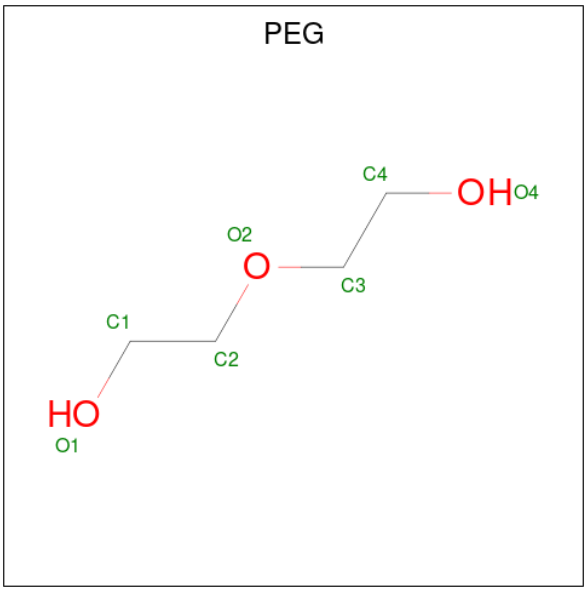
- Molecule 1 is a protein called VEnT1.3.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	310	4975	1593	2482	424	455	21	0	11	0

There are 25 discrepancies between the modelled and reference sequences:

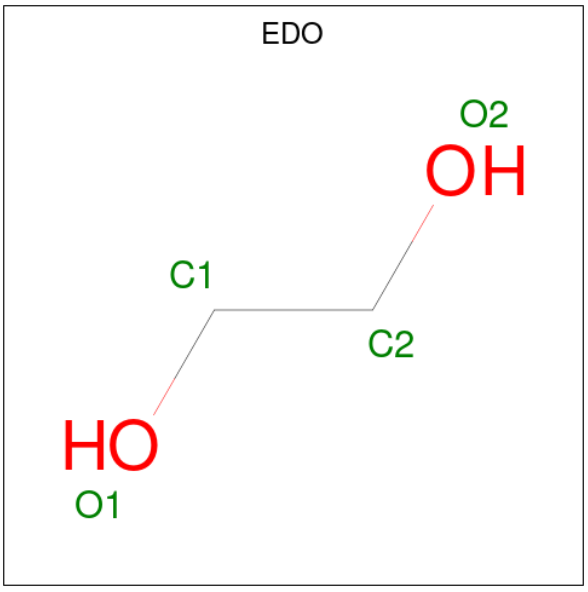
Chain	Residue	Modelled	Actual	Comment	Reference
A	21	MET	GLU	engineered mutation	UNP Q7SIG4
A	37	SER	GLU	engineered mutation	UNP Q7SIG4
A	120	ALA	ASN	engineered mutation	UNP Q7SIG4
A	121	TYR	ASP	engineered mutation	UNP Q7SIG4
A	144	PHE	TYR	engineered mutation	UNP Q7SIG4
A	146	ILE	ARG	engineered mutation	UNP Q7SIG4
A	148	LEU	MET	engineered mutation	UNP Q7SIG4
A	175	ALA	ASN	engineered mutation	UNP Q7SIG4
A	195	GLN	THR	engineered mutation	UNP Q7SIG4
A	229	ALA	ASP	engineered mutation	UNP Q7SIG4
A	244	ARG	TRP	engineered mutation	UNP Q7SIG4
A	255	LYS	ASP	engineered mutation	UNP Q7SIG4
A	272	ALA	ASN	engineered mutation	UNP Q7SIG4
A	315	GLY	-	expression tag	UNP Q7SIG4
A	316	SER	-	expression tag	UNP Q7SIG4
A	317	LEU	-	expression tag	UNP Q7SIG4
A	318	GLU	-	expression tag	UNP Q7SIG4
A	319	TRP	-	expression tag	UNP Q7SIG4
A	320	SER	-	expression tag	UNP Q7SIG4
A	321	HIS	-	expression tag	UNP Q7SIG4
A	322	PRO	-	expression tag	UNP Q7SIG4
A	323	GLN	-	expression tag	UNP Q7SIG4
A	324	PHE	-	expression tag	UNP Q7SIG4
A	325	GLU	-	expression tag	UNP Q7SIG4
A	326	LYS	-	expression tag	UNP Q7SIG4

- Molecule 2 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			17	4	10	3		
2	A	1	Total	C	H	O	0	0
			17	4	10	3		
2	A	1	Total	C	H	O	0	0
			17	4	10	3		

- Molecule 3 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			10	2	6	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 4 is water.

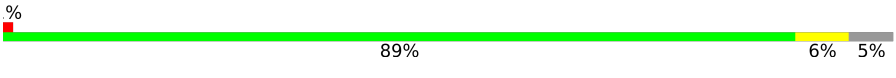
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	223	Total	O	0	0
			223	223		

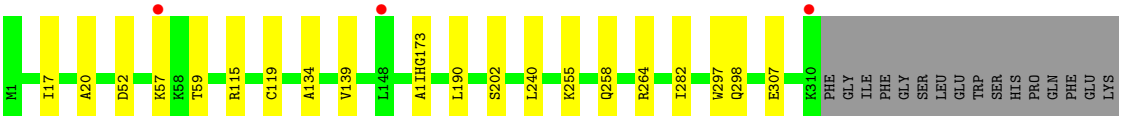
3 Residue-property plots (i)

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: VEnT1.3

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	42.06Å 81.65Å 85.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.95 – 1.82 42.95 – 1.82	Depositor EDS
% Data completeness (in resolution range)	99.7 (42.95-1.82) 99.9 (42.95-1.82)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.08 (at 1.82Å)	Xtriage
Refinement program	PHENIX 1.21_5207	Depositor
R, R_{free}	0.188 , 0.237 0.193 , 0.242	Depositor DCC
R_{free} test set	1335 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å^2)	31.5	Xtriage
Anisotropy	0.303	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.43 , 52.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.002 for -h,l,k	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	5269	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.58% 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: PEG, EDO, A1IHG

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	0.15	0/2564	0.37	0/3458

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

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	2493	2482	2475	14	0
2	A	21	30	30	0	0
3	A	8	12	12	0	0
4	A	223	0	0	4	1
All	All	2745	2524	2517	14	1

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 3.

The worst 5 of 14 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:264[A]:ARG:NH1	4:A:601:HOH:O	2.02	0.91
1:A:115[A]:ARG:NH2	1:A:139:VAL:O	2.13	0.82
1:A:258:GLN:OE1	4:A:602:HOH:O	2.15	0.64
1:A:52:ASP:HB2	1:A:57:LYS:HE3	1.90	0.54
1:A:255:LYS:NZ	4:A:609:HOH:O	2.46	0.49

All (1) 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)
4:A:752:HOH:O	4:A:811:HOH:O[2_554]	2.14	0.06

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	318/326 (98%)	306 (96%)	12 (4%)	0	100 100

There are no Ramachandran outliers to report.

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	266/269 (99%)	266 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

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

Mol	Chain	Res	Type
1	A	67	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)

1 non-standard protein/DNA/RNA residue is 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	A1IHG	A	173	1	21,22,23	2.43	9 (42%)	28,31,33	3.28	10 (35%)

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	A1IHG	A	173	1	-	2/5/6/8	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	173	A1IHG	C10-C09	5.06	1.59	1.48
1	A	173	A1IHG	CZ-S16	4.89	1.84	1.76
1	A	173	A1IHG	C15-S16	4.11	1.83	1.76
1	A	173	A1IHG	CE2-C09	2.94	1.54	1.48
1	A	173	A1IHG	CD2-CE2	2.71	1.44	1.39

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	173	A1IHG	C10-C09-CE2	7.96	130.96	117.60
1	A	173	A1IHG	CZ-CE2-C09	-7.59	118.67	123.72
1	A	173	A1IHG	C15-C10-C09	-7.38	118.81	123.72
1	A	173	A1IHG	C15-S16-CZ	5.44	113.47	100.44
1	A	173	A1IHG	O17-C09-C10	-4.92	113.70	120.91

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	173	A1IHG	N-CA-CB-CG
1	A	173	A1IHG	C-CA-CB-CG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

5 ligands 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	$\#\|Z\| > 2$
2	PEG	A	501	-	6,6,6	0.24	0	5,5,5	0.30	0
2	PEG	A	504	-	6,6,6	0.23	0	5,5,5	0.25	0
3	EDO	A	503	-	3,3,3	0.28	0	2,2,2	0.27	0
2	PEG	A	505	-	6,6,6	0.23	0	5,5,5	0.28	0
3	EDO	A	502	-	3,3,3	0.27	0	2,2,2	0.28	0

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
2	PEG	A	501	-	-	0/4/4/4	-
2	PEG	A	504	-	-	0/4/4/4	-
3	EDO	A	503	-	-	0/1/1/1	-
2	PEG	A	505	-	-	2/4/4/4	-
3	EDO	A	502	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	505	PEG	C1-C2-O2-C3
2	A	505	PEG	O1-C1-C2-O2
3	A	502	EDO	O1-C1-C2-O2

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	309/326 (94%)	-0.27	3 (0%)	79	80	15, 38, 59, 93	11 (3%)

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	310	LYS	2.9
1	A	57	LYS	2.4
1	A	148	LEU	2.2

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

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
1	A1IHG	A	173	20/21	0.91	0.11	28,43,57,64	0

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
3	EDO	A	503	4/4	0.72	0.20	73,88,94,94	0
2	PEG	A	505	7/7	0.73	0.15	48,62,74,75	0
2	PEG	A	504	7/7	0.77	0.16	48,63,83,83	0
3	EDO	A	502	4/4	0.84	0.12	50,62,74,84	0
2	PEG	A	501	7/7	0.88	0.12	51,61,73,76	0

6.5 Other polymers (i)

There are no such residues in this entry.