



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 12, 2026 – 12:40 PM UTC

PDB ID : 4E06 / pdb_00004e06
Title : Anophelin from the malaria vector inhibits thrombin through a novel reverse-binding mechanism
Authors : Figueiredo, A.C.; de Sanctis, D.; Gutierrez-Gallego, R.; Cereija, T.B.; Macedo-Ribeiro, S.; Fuentes-Prior, P.; Pereira, P.J.B.
Deposited on : 2012-03-02
Resolution : 3.20 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

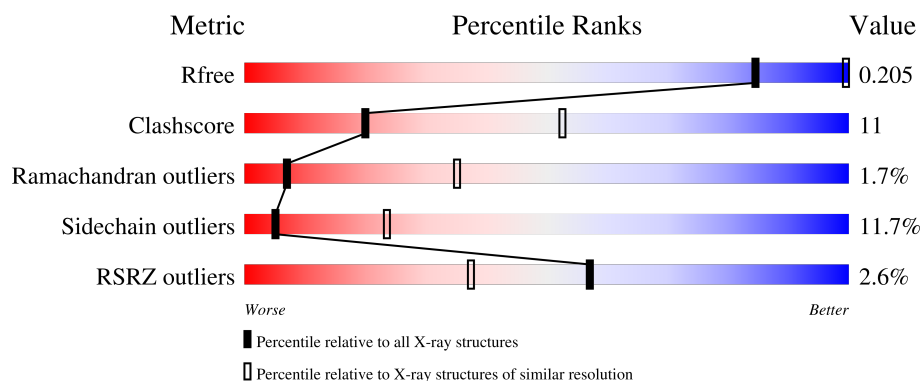
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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}	180053	1466 (3.20-3.20)
Clashscore	190562	1573 (3.20-3.20)
Ramachandran outliers	187476	1548 (3.20-3.20)
Sidechain outliers	187428	1547 (3.20-3.20)
RSRZ outliers	180081	1466 (3.20-3.20)

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 $\leq 5\%$. 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	L	36	<div> <div>3%</div> <div>58%</div> <div>17%</div> <div>•</div> <div>22%</div> </div>
2	H	259	<div> <div>%</div> <div>64%</div> <div>29%</div> <div>• •</div> </div>
3	I	61	<div> <div>8%</div> <div>26%</div> <div>13%</div> <div>7%</div> <div>54%</div> </div>

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 2512 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 Thrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	28	Total	C	N	O	S	0	0	0
			230	144	37	48	1			

- Molecule 2 is a protein called Thrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	252	Total	C	N	O	S	0	0	0
			2039	1299	360	366	14			

- Molecule 3 is a protein called Salivary anti-thrombin peptide anophelin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	I	28	Total	C	N	O	0	0	0
			205	125	36	44			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	H	1	Total	C	N	O	0	0
			14	8	1	5		

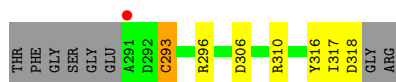
- Molecule 5 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	H	1	Total	Na	0	0
			1	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	H	20	Total	O	0	0
			20	20		
6	I	3	Total	O	0	0
			3	3		

- Molecule 1: Thrombin



S525	S526	F529	M542	G543	I544	V545	G550	C551	Y560	T561	L566	K567	K568	W569	K572	K573	I574	D575	Q576	F577	G578	E579	T445	A446	A447	S448	L449	L450	Q451	Y454	K455	G456	T459	G460	W461	L464	K465	E466	T467	TRP	THR	ALA	ASN	VAL	GLY	LYS	G475	Q476	P477	S478	V479	L480	Q481	V482	L485	P486	I487	R490	T502	D503	N504	M505	K511	F512	D513	E514	G515	K516	R517	G518	D519	A520	C521	D524	T321	V322	E323	D326	A327	E328	L346	K347	C348	R356	A361	A362	L366	L367	P368	P369	W370	N377	T383	S387	R388	T389	R390	E396	S399	W400	L401	E402	K403	L406	H407	P408	R409	Y410	N411	W412	R413	R418	D419	T420	A421	K424	D427	T435
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

ALA PRO GLN TYR ALA PRO GLY ASP GLU PRO SER TYR ASP GLU THR ASP ASP SER ASP LYS LEU VAL GLU ASN ASP THR SER ILE THR ASP GLU ASP ASP Y34 Y35 Y36 Y37 Y38 Y39 Y40 Y41 Y44 Y45 N46 T47 T48 A48 P51 G52 R53 R54 T54

4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	209.66Å 209.66Å 127.41Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.83 – 3.20 46.83 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.83-3.20) 93.1 (46.83-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.39 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.7.2_869	Depositor
R, R_{free}	0.177 , 0.208 0.178 , 0.205	Depositor DCC
R_{free} test set	911 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å ²)	69.8	Xtriage
Anisotropy	0.556	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 94.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.033 for -2/3*h-1/3*k-4/3*l,-1/3*h-2/3*k+4/3*l,-1/3*h+1/3*k+1/3*l 0.018 for -h,1/3*h-1/3*k-4/3*l,-1/3*h-2/3*k+1/3*l 0.019 for -1/3*h+1/3*k+4/3*l,-k,2/3*h+1/3*k+1/3*l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	2512	wwPDB-VP
Average B, all atoms (Å ²)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: NAG, NA

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	L	0.51	0/232	1.04	2/309 (0.6%)
2	H	0.54	0/2091	0.93	5/2823 (0.2%)
3	I	0.60	0/208	1.07	0/279
All	All	0.54	0/2531	0.95	7/3411 (0.2%)

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	485	LEU	CA-C-N	6.07	125.83	119.76
2	H	485	LEU	C-N-CA	6.07	125.83	119.76
2	H	407	HIS	CA-C-N	5.68	125.36	119.05
2	H	407	HIS	C-N-CA	5.68	125.36	119.05
2	H	445	THR	N-CA-C	-5.51	105.35	111.36

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	L	230	0	226	3	0
2	H	2039	0	2009	50	0
3	I	205	0	194	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	H	14	0	13	0	0
5	H	1	0	0	0	0
6	H	20	0	0	0	0
6	I	3	0	0	0	0
All	All	2512	0	2442	56	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 11.

The worst 5 of 56 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)
2:H:450:LEU:HD12	2:H:487:ILE:HD13	1.58	0.86
2:H:466:GLU:HB2	2:H:551:CYS:HB2	1.75	0.69
2:H:409:ARG:HB3	2:H:418:ARG:HD3	1.75	0.68
2:H:413:ARG:HB2	2:H:413:ARG:HH21	1.58	0.68
2:H:513:ASP:OD1	2:H:513:ASP:N	2.28	0.63

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	L	26/36 (72%)	22 (85%)	3 (12%)	1 (4%)	2	18
2	H	248/259 (96%)	235 (95%)	11 (4%)	2 (1%)	16	50
3	I	26/61 (43%)	21 (81%)	3 (12%)	2 (8%)	1	5
All	All	300/356 (84%)	278 (93%)	17 (6%)	5 (2%)	7	35

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	I	37	ILE
2	H	578	GLY
2	H	576	GLN
3	I	48	ALA
1	L	293	CYS

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	L	26/31 (84%)	24 (92%)	2 (8%)	12	41
2	H	220/225 (98%)	197 (90%)	23 (10%)	6	28
3	I	20/50 (40%)	14 (70%)	6 (30%)	0	1
All	All	266/306 (87%)	235 (88%)	31 (12%)	5	23

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	448	SER
3	I	38	GLU
2	H	478	SER
3	I	44	THR
2	H	574	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	H	386	HIS
2	H	411	ASN
2	H	476	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

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	$\# Z > 2$
4	NAG	H	601	2	14,14,15	0.50	0	17,19,21	1.21	3 (17%)

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
4	NAG	H	601	2	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	601	NAG	C1-O5-C5	2.37	115.36	112.19
4	H	601	NAG	C2-N2-C7	-2.24	119.90	122.90
4	H	601	NAG	C4-C3-C2	-2.01	108.08	111.02

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	H	601	NAG	C8-C7-N2-C2
4	H	601	NAG	O7-C7-N2-C2
4	H	601	NAG	C1-C2-N2-C7

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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(Å ²)	Q<0.9
1	L	28/36 (77%)	-0.07	1 (3%) 46 28	39, 74, 103, 142	1 (3%)
2	H	252/259 (97%)	-0.43	2 (0%) 82 67	54, 77, 119, 158	0
3	I	28/61 (45%)	0.59	5 (17%) 3 3	64, 92, 142, 165	0
All	All	308/356 (86%)	-0.31	8 (2%) 57 37	39, 77, 124, 165	1 (0%)

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	467	THR	4.2
3	I	34	TYR	3.3
3	I	35	ALA	2.9
3	I	39	ALA	2.9
1	L	291	ALA	2.7

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

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, 95th 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(\AA^2)	Q<0.9
4	NAG	H	601	14/15	0.87	0.15	109,125,129,130	0
5	NA	H	602	1/1	0.98	0.11	57,57,57,57	0

6.5 Other polymers [i](#)

There are no such residues in this entry.