



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 10:02 pm GMT

PDB ID : 2PV9
Title : Crystal structure of murine thrombin in complex with the extracellular fragment of murine PAR4
Authors : Bah, A.; Chen, Z.; Bush-Pelc, L.A.; Mathews, F.S.; Di Cera, E.
Deposited on : 2007-05-09
Resolution : 3.50 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

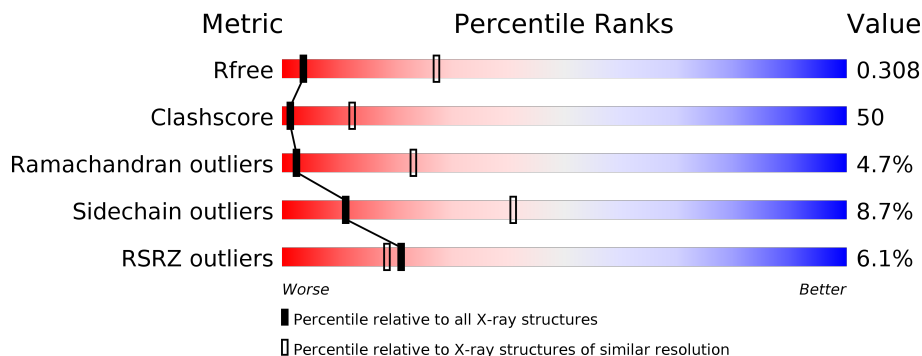
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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}	111664	1391 (3.60-3.40)
Clashscore	122126	1485 (3.60-3.40)
Ramachandran outliers	120053	1446 (3.60-3.40)
Sidechain outliers	120020	1447 (3.60-3.40)
RSRZ outliers	108989	1303 (3.60-3.40)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	A	44	<div> <div>39%</div> <div>50%</div> <div>9%</div> <div>•</div> </div>
2	B	258	<div> <div>35%</div> <div>57%</div> <div>7%</div> </div>
3	C	26	<div> <div>15%</div> <div>35%</div> <div>38%</div> <div>12%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2679 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 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	43	Total	C	N	O	S	0	0	0
			343	221	52	69	1			

- Molecule 2 is a protein called Thrombin heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	258	Total	C	N	O	S	0	0	0
			2109	1352	378	367	12			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	195	ALA	SER	ENGINEERED	UNP P19221

- Molecule 3 is a protein called Proteinase-activated receptor 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	26	Total	C	N	O	S	0	0	0
			199	122	34	42	1			

- Molecule 4 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C₈H₁₅NO₆).



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

4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α , β , γ	111.15Å 111.15Å 179.71Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.52 – 3.50 37.52 – 3.50	Depositor EDS
% Data completeness (in resolution range)	97.2 (37.52-3.50) 97.2 (37.52-3.50)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.90 (at 3.48Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.306 , 0.319 0.272 , 0.308	Depositor DCC
R_{free} test set	472 reflections (5.52%)	wwPDB-VP
Wilson B-factor (Å ²)	77.2	Xtriage
Anisotropy	0.113	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 12.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.33$, $\langle L^2 \rangle = 0.16$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	2679	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.95% 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

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	0.47	0/349	0.72	0/466
2	B	0.52	1/2166 (0.0%)	0.79	4/2933 (0.1%)
3	C	0.47	0/204	1.47	6/274 (2.2%)
All	All	0.51	1/2719 (0.0%)	0.85	10/3673 (0.3%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	186(A)	ASN	CB-CG	6.48	1.66	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	186(A)	ASN	CB-CG-ND2	-8.28	96.84	116.70
3	C	53	SER	N-CA-C	6.90	129.63	111.00
2	B	186(A)	ASN	CB-CG-OD1	5.96	133.53	121.60
3	C	55	LYS	N-CA-C	-5.58	95.95	111.00
3	C	52	SER	N-CA-C	-5.47	96.24	111.00

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	343	0	331	22	0
2	B	2109	0	2100	235	0
3	C	199	0	187	42	0
4	B	28	0	26	0	0
All	All	2679	0	2644	264	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 50.

The worst 5 of 264 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:B:149(A):THR:HG21	3:C:70:SER:HB2	1.28	1.11
3:C:61:TYR:HB2	3:C:62:PRO:HD2	1.39	1.05
2:B:186:VAL:HG11	2:B:186(C):THR:HB	1.45	0.97
2:B:192:GLU:HG2	3:C:62:PRO:HD3	1.55	0.88
2:B:185:LYS:O	2:B:186(A):ASN:N	2.07	0.88

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	41/44 (93%)	31 (76%)	9 (22%)	1 (2%)	6	38
2	B	256/258 (99%)	218 (85%)	32 (12%)	6 (2%)	7	39
3	C	24/26 (92%)	10 (42%)	6 (25%)	8 (33%)	0	0
All	All	321/328 (98%)	259 (81%)	47 (15%)	15 (5%)	2	23

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	149(A)	THR
2	B	186	VAL
2	B	186(C)	THR
3	C	52	SER
3	C	65	PHE

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	37/39 (95%)	33 (89%)	4 (11%)	7	31
2	B	228/228 (100%)	210 (92%)	18 (8%)	13	45
3	C	23/23 (100%)	20 (87%)	3 (13%)	4	23
All	All	288/290 (99%)	263 (91%)	25 (9%)	11	41

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	94	TYR
2	B	148	TRP
3	C	65	PHE
2	B	131	ARG
2	B	149(E)	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	159	ASN
2	B	239	GLN
2	B	204(B)	ASN
2	B	95	ASN
2	B	209	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

2 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
4	NAG	B	301	2	14,14,15	1.59	3 (21%)	17,19,21	2.29	5 (29%)
4	NAG	B	302	2	14,14,15	2.19	7 (50%)	17,19,21	3.14	9 (52%)

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	B	301	2	-	0/6/23/26	0/1/1/1
4	NAG	B	302	2	-	0/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	302	NAG	C2-N2	2.20	1.50	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	302	NAG	C4-C5	2.20	1.57	1.53
4	B	301	NAG	C4-C3	2.38	1.58	1.52
4	B	302	NAG	O4-C4	2.45	1.49	1.43
4	B	302	NAG	C4-C3	2.45	1.58	1.52

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	302	NAG	O5-C1-C2	-8.02	100.45	111.52
4	B	302	NAG	C3-C4-C5	-5.29	100.78	110.24
4	B	301	NAG	C1-O5-C5	-4.77	105.63	112.19
4	B	301	NAG	O5-C1-C2	-4.52	105.28	111.52
4	B	301	NAG	C1-C2-N2	-3.82	103.97	110.49

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 [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	A	43/44 (97%)	-0.57	0 100 100	20, 38, 64, 80	0
2	B	258/258 (100%)	-0.51	1 (0%) 92 90	7, 35, 69, 99	0
3	C	26/26 (100%)	3.67	19 (73%) 0 0	52, 94, 102, 102	0
All	All	327/328 (99%)	-0.18	20 (6%) 21 18	7, 37, 94, 102	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	69	ASP	7.1
3	C	71	ASP	7.1
3	C	74	GLU	7.1
3	C	66	CYS	6.4
3	C	70	SER	6.3

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 carbohydrates 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	B	301	14/15	0.78	0.20	76,85,87,89	0
4	NAG	B	302	14/15	0.84	0.35	87,96,99,101	0

6.5 Other polymers [i](#)

There are no such residues in this entry.