wwPDB X-ray Structure Validation Summary Report (i)

Aug 2, 2023 – 02:45 AM EDT

PDB ID	:	1B7X
Title	:	STRUCTURE OF HUMAN ALPHA-THROMBIN Y225I MUTANT BOUND TO D-PHE-PRO-ARG-CHLOROMETHYLKETONE
Authors	:	Caccia, S.; Futterer, K.; Di Cera, E.; Waksman, G.
Deposited on	:	1999-01-25
Resolution	:	2.10 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 (i)) were used in the production of this report:

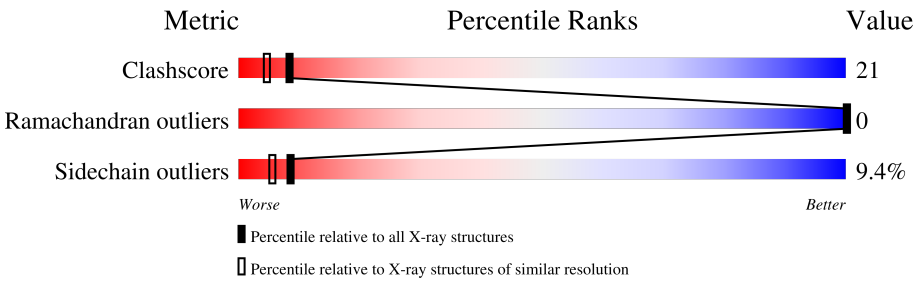
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 2.10 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	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	36	28% 42% • 28%
2	B	259	59% 32% • 5%
3	C	3	67% 33%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2261 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 PROTEIN (THROMBIN LIGHT CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	26	214	134	35	44	1	0	0	0

- Molecule 2 is a protein called PROTEIN (THROMBIN HEAVY CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	246	1872	1200	323	335	14	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	225	ILE	TYR	engineered mutation	UNP P00734

- Molecule 3 is a protein called PROTEIN (INHIBITOR).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	3	29	20	6	3	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	20	20	20	0	0
4	B	125	125	125	0	0
4	C	1	1	1	0	0

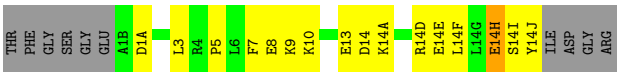
3 Residue-property plots (i)

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. 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.

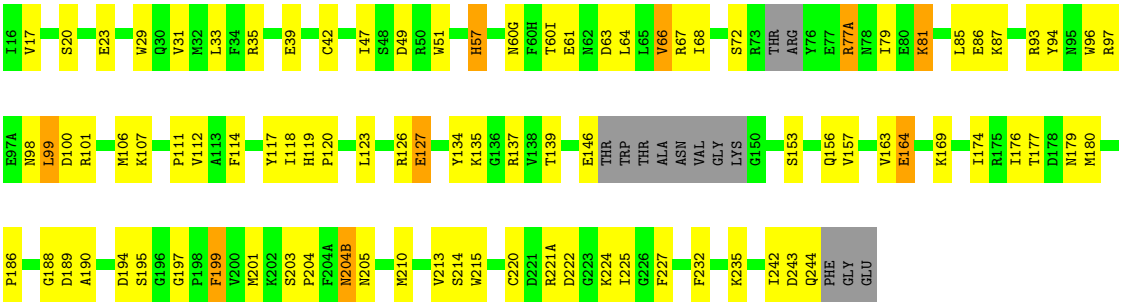
• Molecule 1: PROTEIN (THROMBIN LIGHT CHAIN)

Chain A: 28% 42% • 28%



• Molecule 2: PROTEIN (THROMBIN HEAVY CHAIN)

Chain B: 59% 32% • 5%



• Molecule 3: PROTEIN (INHIBITOR)

Chain C: 67% 33%

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	50.60Å 73.70Å 89.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 2.10	Depositor
% Data completeness (in resolution range)	85.6 (6.00-2.10)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.226 , 0.282	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2261	wwPDB-VP
Average B, all atoms (Å^2)	19.0	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	0.54	0/216	0.69	0/287
2	B	0.53	1/1920 (0.1%)	0.73	1/2610 (0.0%)
3	C	1.36	0/30	1.71	2/39 (5.1%)
All	All	0.55	1/2166 (0.0%)	0.75	3/2936 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	C	1	0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	127	GLU	CB-CG	5.01	1.61	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	303	ARG	NE-CZ-NH1	6.37	123.48	120.30
3	C	303	ARG	NE-CZ-NH2	-5.60	117.50	120.30
2	B	199	PHE	N-CA-C	-5.17	97.03	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	C	301	PHE	CA

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	214	0	214	14	1
2	B	1872	0	1751	78	1
3	C	29	0	28	2	0
4	A	20	0	0	2	0
4	B	125	0	0	11	0
4	C	1	0	0	0	0
All	All	2261	0	1993	86	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 21.

The worst 5 of 86 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:139:THR:HG22	2:B:157:VAL:HG12	1.20	1.12
2:B:156:GLN:HB3	4:B:6015:HOH:O	1.59	1.00
2:B:194:ASP:HB3	4:B:3044:HOH:O	1.68	0.92
2:B:214:SER:HB3	4:B:5006:HOH:O	1.71	0.89
2:B:139:THR:CG2	2:B:157:VAL:HG12	2.03	0.89

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)
1:A:1(A):ASP:OD2	2:B:127:GLU:OE2[4_456]	2.08	0.12

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	24/36 (67%)	23 (96%)	1 (4%)	0	100	100
2	B	240/259 (93%)	229 (95%)	11 (5%)	0	100	100
3	C	1/3 (33%)	1 (100%)	0	0	100	100
All	All	265/298 (89%)	253 (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	24/31 (77%)	22 (92%)	2 (8%)	11	7
2	B	186/225 (83%)	168 (90%)	18 (10%)	8	5
3	C	3/3 (100%)	3 (100%)	0	100	100
All	All	213/259 (82%)	193 (91%)	20 (9%)	8	5

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	123	LEU
2	B	204(B)	ASN
2	B	242	ILE
2	B	235	LYS
2	B	60(G)	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	156	GLN
2	B	204(B)	ASN
2	B	239	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	71	HIS
2	B	60(G)	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 (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.