W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 9, 2021 – 12:22 PM GMT

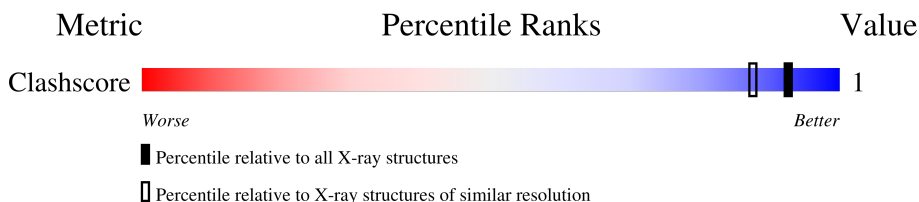|		
PDB ID	:	6Y02
Title	:	Thrombin in complex with 13k
Authors	:	Sandner, A.; Heine, A.; Klebe, G.; Abazi, N.
Deposited on	:	2020-02-06
Resolution	:	1.48 Å(reported)

1 Overall quality at a glance ⓘ

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

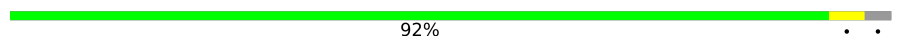The reported resolution of this entry is 1.48 Å.

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	Percentile Ranks	Value
Clashscore		1

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	4955 (1.50-1.46)

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 failed to run properly.

Mol	Chain	Length	Quality of chain
1	L	36	78% 22%
2	H	259	92% · ·
3	C	12	83% 8% 8%

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 4949 atoms, of which 2288 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 Prothrombin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	L	28	Total	C	H	N	O	S	0	7	0
			480	158	231	38	52	1			

- Molecule 2 is a protein called Prothrombin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	H	250	Total	C	H	N	O	S	0	16	0
			4050	1314	1995	361	366	14			

- Molecule 3 is a protein called Hirudin variant-2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	11	Total	C	H	N	O	S	0	0	0
			146	55	62	11	17	1			

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	H	2	Total	Na	0	0
			2	2		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).

PO4

O3

O⁻

O2 O⁻ — P = O O1

P

O⁻

O4

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	H	1	Total 5	O 4	P 1	0	0

- Molecule 6 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).

DMS

C2

S

C1 — S

O

O
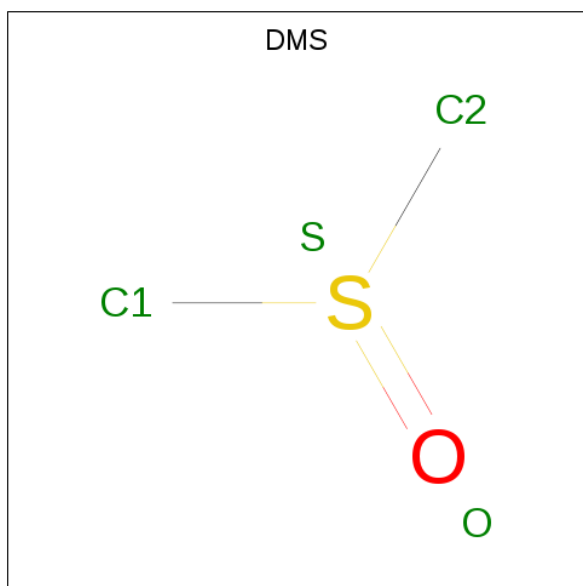
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	H	1	Total 4	C 2	O 1	S 1	0	0
6	H	1	Total 4	C 2	O 1	S 1	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

C3

OH O3

C1 C2

HO

O1

OH

O2

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	H	1	Total	C	O	0	0
			6	3	3		
7	H	1	Total	C	O	0	0
			6	3	3		

- Molecule 8 is (2 {S})-1-[(2 {R})-2-azanyl-3-phenyl-propanoyl]- {N}-[(5-bromanylfuran-2-yl)methyl]pyrrolidine-2-carboxamide (three-letter code: O5Z) (formula: $C_{19}H_{22}BrN_3O_3$) (labeled as "Ligand of Interest" by depositor).

O5Z

C4

C5 C3

C6 C2 C1

C7

H₂N C(R)

N C8 O O

O1

N1 C13 C14

C9 C12(S) C15 O2

N H C18 Br BR

N2

C10 C11 C16

C17
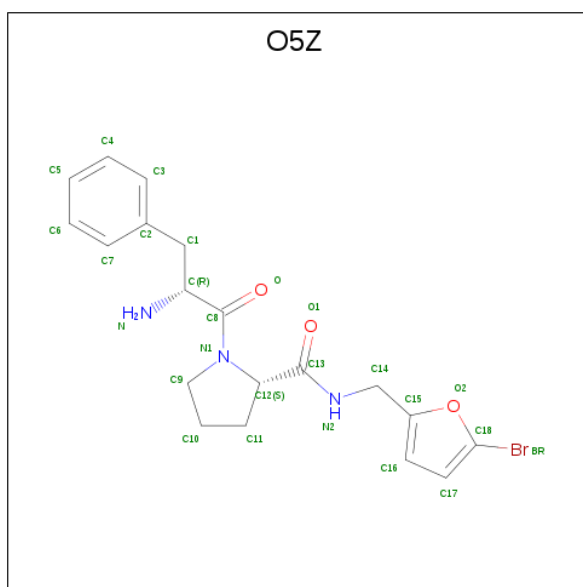
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	H	1	Total	Br	C	N	O	0	0
			26	1	19	3	3		

- Molecule 9 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
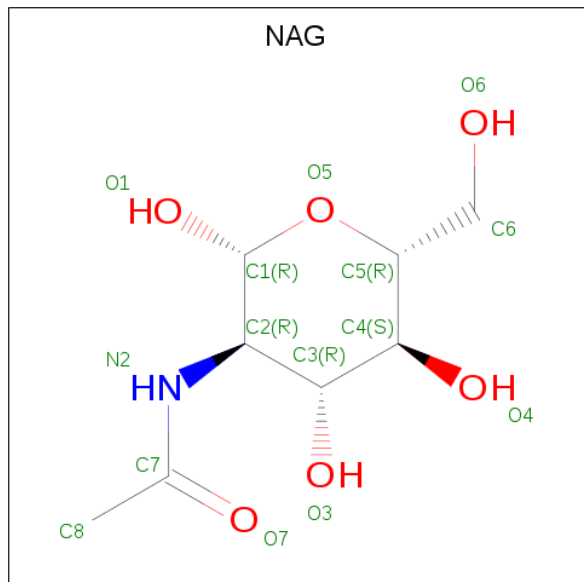
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	H	1	Total 13	C 8	N 1	O 4	0	0

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	L	25	Total 25	O 25	0	0
10	H	177	Total 179	O 179	0	3
10	C	3	Total 3	O 3	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.
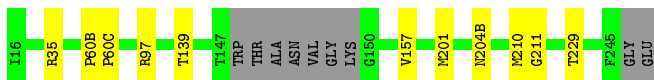
Note EDS failed to run properly.
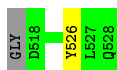
• Molecule 1: Prothrombin

Chain L:

• Molecule 2: Prothrombin

Chain H:

• Molecule 3: Hirudin variant-2

Chain C:

4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	70.13Å 71.40Å 72.89Å 90.00° 100.51° 90.00°	Depositor
Resolution (Å)	43.55 – 1.48	Depositor
% Data completeness (in resolution range)	98.1 (43.55-1.48)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	1.17 (at 1.48Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, R_{free}	0.139 , 0.164	Depositor
Wilson B-factor (Å^2)	18.3	Xtriage
Anisotropy	0.138	Xtriage
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4949	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.83% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, NAG, NA, PO4, O5Z, DMS, TYS

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	0.42	0/273	0.68	0/365
2	H	0.40	0/2181	0.63	0/2952
3	C	0.28	0/68	0.37	0/91
All	All	0.40	0/2522	0.63	0/3408

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 (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	L	249	231	217	0	0
2	H	2055	1995	1926	7	0
3	C	84	62	62	0	0
4	H	2	0	0	0	0
5	H	5	0	0	0	0
6	H	8	0	12	0	0
7	H	12	0	15	1	0
8	H	26	0	0	0	0
9	H	13	0	10	0	0
10	C	3	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	H	179	0	0	2	0
10	L	25	0	0	0	0
All	All	2661	2288	2242	7	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 1.

The worst 5 of 7 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:139:THR:HG22	2:H:157[A]:VAL:HG22	1.83	0.59
2:H:204(B):ASN:ND2	7:H:309:GOL:H12	2.27	0.50
2:H:97[B]:ARG:NH1	10:H:402[B]:HOH:O	2.42	0.47
2:H:211:GLY:HA2	2:H:229:THR:O	2.17	0.44
2:H:35:ARG:NH2	10:H:407:HOH:O	2.52	0.41

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

There are no protein backbone outliers to report in this entry.

5.3.2 Protein sidechains (i)

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

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

wwPDB X-ray Structure Validation Summary Report

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	#$|Z| > 2$
3	TYS	C	526	3	15,16,17	1.27	2 (13%)	18,22,24	1.01	1 (5%)

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
3	TYS	C	526	3	-	0/10/11/13	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	526	TYS	OH-CZ	-3.34	1.37	1.42
3	C	526	TYS	OH-S	-2.51	1.54	1.58

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	526	TYS	CB-CA-C	-2.03	107.66	111.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

LIGAND-GEOMETRY INFOmissingINFO

5.6 Other polymers (i)

There are no such residues in this entry.

5.7 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 failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates (i)

EDS failed to run properly - this section is therefore empty.

6.4 Ligands (i)

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers (i)

EDS failed to run properly - this section is therefore empty.