



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 26, 2017 – 12:57 PM EDT

PDB ID : 1C2F  
Title : RECRUITING ZINC TO MEDIATE POTENT, SPECIFIC INHIBITION OF  
SERINE PROTEASES  
Authors : Katz, B.A.; Luong, C.  
Deposited on : unknown  
Resolution : 1.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

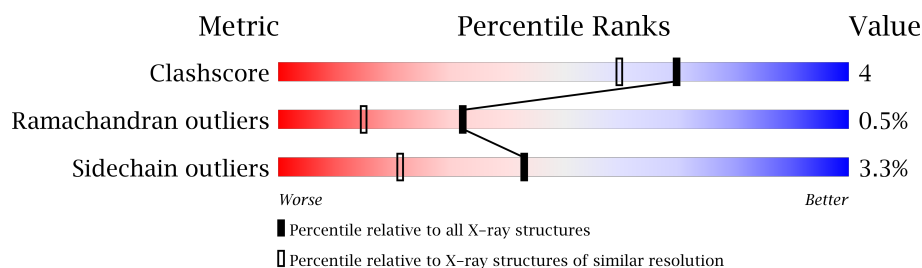
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	112137	3876 (1.70-1.70)
Ramachandran outliers	110173	3815 (1.70-1.70)
Sidechain outliers	110143	3815 (1.70-1.70)

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  $\geq 3$ , 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	223	 84% 15%

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 4044 atoms, of which 2122 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 TRYPSIN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	223	Total	C	H	N	O	S	4	8	0
			3294	1032	1639	281	328	14			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ca	0	0
			1	1		

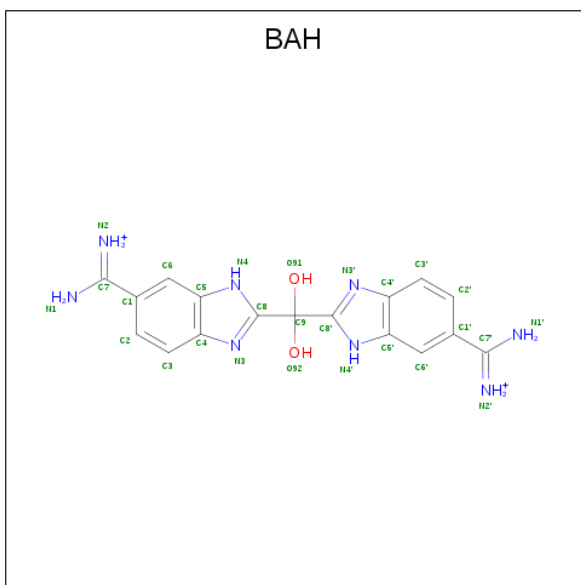
- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Mg	0	0
			1	1		

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

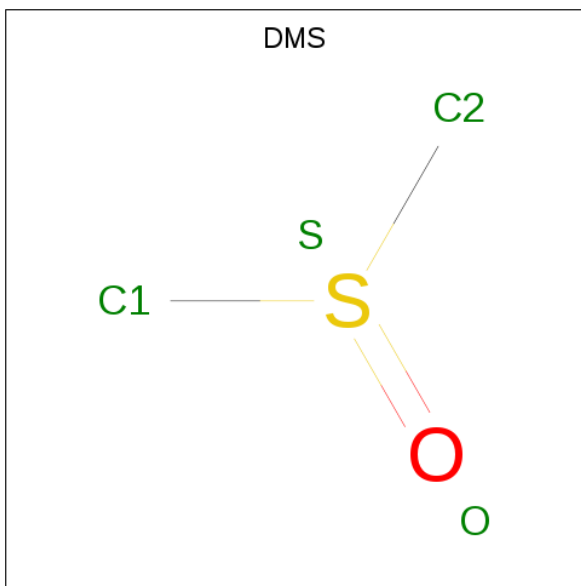
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total	Zn	0	0
			3	3		

- Molecule 5 is BIS(5-AMIDINO-2-BENZIMIDAZOLYL)METHANE KETONE HYDRATE (three-letter code: BAH) (formula: C<sub>17</sub>H<sub>18</sub>N<sub>8</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	H	N	O	0	0
			42	17	15	8	2		

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



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

- Molecule 7 is water.

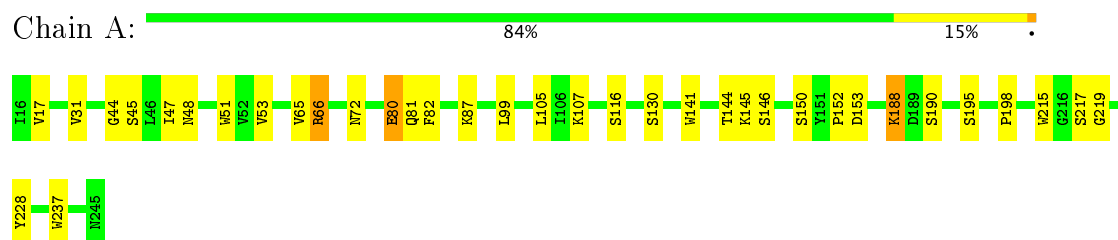
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	231	Total	H	O	0	16
			693	462	231		

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

Note EDS was not executed.

#### • Molecule 1: TRYPSIN



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.99 Å 54.99 Å 109.92 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	7.50 – 1.70	Depositor
% Data completeness (in resolution range)	66.0 (7.50-1.70)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	0.190 , 0.241	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4044	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MG, DMS, CA, BAH

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	1.32	2/1718 (0.1%)	1.48	19/2328 (0.8%)

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
1	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	80	GLU	CD-OE2	-5.11	1.20	1.25
1	A	215	TRP	CG-CD2	-5.01	1.35	1.43

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	51	TRP	CD1-NE1-CE2	9.93	117.93	109.00
1	A	141	TRP	CD1-NE1-CE2	9.17	117.25	109.00
1	A	237	TRP	CD1-NE1-CE2	8.54	116.68	109.00
1	A	215	TRP	CD1-NE1-CE2	8.04	116.24	109.00
1	A	215	TRP	NE1-CE2-CZ2	7.49	138.63	130.40
1	A	141	TRP	NE1-CE2-CZ2	7.22	138.35	130.40
1	A	51	TRP	CG-CD1-NE1	-6.88	103.22	110.10
1	A	141	TRP	CG-CD1-NE1	-6.86	103.24	110.10
1	A	237	TRP	CG-CD1-NE1	-6.59	103.51	110.10
1	A	215	TRP	CG-CD1-NE1	-6.35	103.75	110.10
1	A	51	TRP	NE1-CE2-CZ2	6.32	137.35	130.40

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	190	SER	N-CA-C	-6.06	94.64	111.00
1	A	51	TRP	NE1-CE2-CD2	-5.89	101.41	107.30
1	A	141	TRP	NE1-CE2-CD2	-5.54	101.76	107.30
1	A	215	TRP	NE1-CE2-CD2	-5.41	101.89	107.30
1	A	237	TRP	NE1-CE2-CZ2	5.39	136.33	130.40
1	A	217	SER	N-CA-C	-5.33	96.61	111.00
1	A	228	TYR	CB-CG-CD1	-5.23	117.86	121.00
1	A	81	GLN	N-CA-C	-5.19	96.99	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	66	ARG	Sidechain

## 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	1655	1639	1620	14	0
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	3	0	0	0	0
5	A	27	15	16	1	0
6	A	4	6	6	0	0
7	A	231	462	0	0	0
All	All	1922	2122	1642	14	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 4.

All (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:31[A]:VAL:CG1	1:A:65:VAL:HG13	2.34	0.58

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:31[A]:VAL:HG11	1:A:65:VAL:HG13	1.94	0.50
1:A:72:ASN:HA	1:A:153:ASP:O	2.12	0.48
1:A:66:ARG:HG2	1:A:82:PHE:CD2	2.49	0.47
1:A:47:ILE:HD13	1:A:53[B]:VAL:HG13	1.98	0.45
1:A:145:LYS:HG3	1:A:150:SER:HB2	1.98	0.45
1:A:31[B]:VAL:HG22	1:A:44:GLY:CA	2.46	0.44
1:A:80:GLU:HB3	1:A:82:PHE:CE2	2.54	0.42
1:A:144:THR:HG23	1:A:152:PRO:HD3	2.02	0.42
1:A:53[B]:VAL:HG12	1:A:105:LEU:CD2	2.49	0.42
1:A:87[A]:LYS:HG2	1:A:107:LYS:HB3	2.03	0.41
1:A:219:GLY:O	5:A:246:BAH:N1	2.54	0.41
1:A:17:VAL:O	1:A:188:LYS:HA	2.21	0.41
1:A:45:SER:OG	1:A:198:PRO:HB3	2.21	0.40

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	229/223 (103%)	220 (96%)	8 (4%)	1 (0%)	38	20

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	195	SER

### 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	192/184 (104%)	186 (97%)	6 (3%)	45 24

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	48	ASN
1	A	99	LEU
1	A	116	SER
1	A	130	SER
1	A	146	SER
1	A	188	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	GLN
1	A	48	ASN
1	A	79	ASN
1	A	210	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 5 are monoatomic - leaving 2 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	$\# Z  > 2$
5	BAH	A	246	4	26,30,30	3.28	16 (61%)	21,46,46	1.95	3 (14%)
6	DMS	A	333	-	3,3,3	1.71	1 (33%)	3,3,3	2.06	1 (33%)

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
5	BAH	A	246	4	-	0/8/20/20	0/4/4/4
6	DMS	A	333	-	-	0/0/0/0	0/0/0/0

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	246	BAH	C1-C7	-5.21	1.38	1.48
5	A	246	BAH	C1'-C7'	-4.96	1.39	1.48
5	A	246	BAH	C4'-N3'	-3.48	1.27	1.39
5	A	246	BAH	C3'-C4'	-3.47	1.35	1.41
5	A	246	BAH	C4-N3	-3.27	1.28	1.39
5	A	246	BAH	C5-N4	-2.88	1.29	1.39
6	A	333	DMS	O-S	-2.81	1.31	1.50
5	A	246	BAH	C5'-C4'	-2.58	1.33	1.42
5	A	246	BAH	C5-C4	-2.55	1.33	1.42
5	A	246	BAH	C6-C5	-2.53	1.38	1.41
5	A	246	BAH	C5'-N4'	-2.47	1.30	1.39
5	A	246	BAH	C2-C1	2.40	1.43	1.39
5	A	246	BAH	C6-C1	2.60	1.41	1.37
5	A	246	BAH	C2'-C1'	2.87	1.44	1.39
5	A	246	BAH	C6'-C1'	3.19	1.43	1.37
5	A	246	BAH	C9-C8'	5.61	1.59	1.52
5	A	246	BAH	C9-C8	9.18	1.62	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	246	BAH	C2'-C3'-C4'	-5.05	114.65	120.84
5	A	246	BAH	O92-C9-O91	-4.51	104.51	112.72
5	A	246	BAH	C2-C3-C4	-3.86	116.11	120.84
6	A	333	DMS	O-S-C2	2.80	121.31	106.54

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	246	BAH	1	0

## 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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.