



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 20, 2018 – 02:54 pm GMT

PDB ID : 1D7P  
Title : Crystal structure of the c2 domain of human factor viii at 1.5 a resolution at 1.5 Å  
Authors : Pratt, K.P.; Shen, B.W.; Stoddard, B.L.  
Deposited on : 1999-10-19  
Resolution : 1.50 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

<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)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

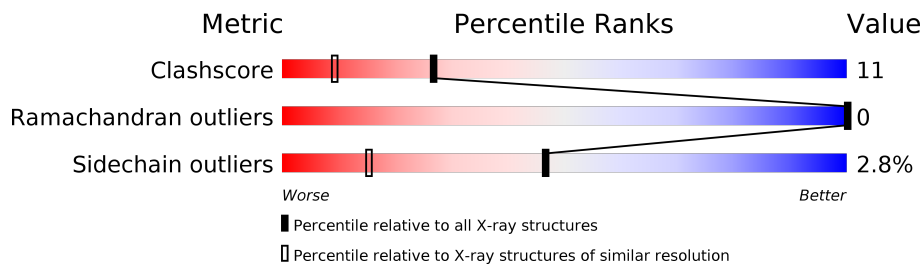
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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))
Clashscore	122078	2726 (1.50-1.50)
Ramachandran outliers	120005	2660 (1.50-1.50)
Sidechain outliers	119972	2658 (1.50-1.50)

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	M	159	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 1538 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.

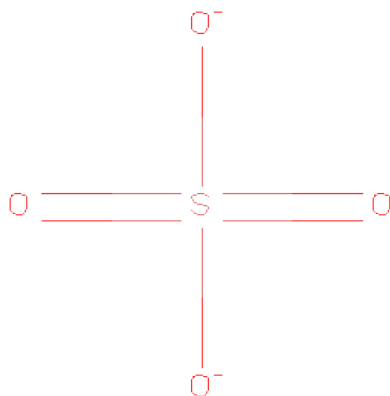
- Molecule 1 is a protein called COAGULATION FACTOR VIII PRECURSOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	M	159	1268	806	221	232	9	0	0	0

There is a discrepancy between the modelled and reference sequences:

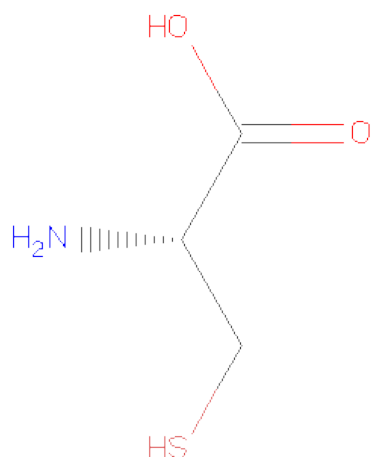
Chain	Residue	Modelled	Actual	Comment	Reference
M	2296	CYS	SER	ENGINEERED	UNP P00451

- Molecule 2 is SULFATE ION (three-letter code: SO<sub>4</sub>) (formula: O<sub>4</sub>S).



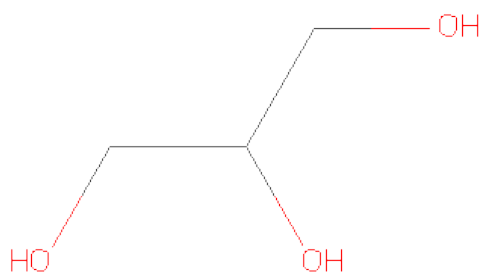
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	M	1	5	4	1	0	0
2	M	1	5	4	1	0	0

- Molecule 3 is CYSTEINE (three-letter code: CYS) (formula: C<sub>3</sub>H<sub>7</sub>NO<sub>2</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	M	1	7	3	1	2	1	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	M	1	6	3	3	0	0

- Molecule 5 is water.

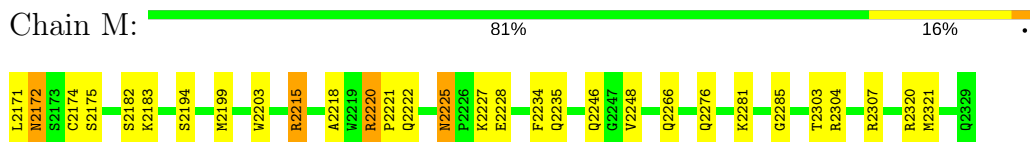
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	M	247	Total 247	O 247	0	0

### 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: COAGULATION FACTOR VIII PRECURSOR



## 4 Data and refinement statistics

Xtrriage (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, $\alpha$ , $\beta$ , $\gamma$	46.44Å 56.41Å 65.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.50	Depositor
% Data completeness (in resolution range)	96.0 (50.00-1.50)	Depositor
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.205 , 0.215	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	1538	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

## 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, SO4

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	M	0.53	0/1300	0.86	6/1762 (0.3%)

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	M	2215	ARG	NE-CZ-NH2	7.04	123.82	120.30
1	M	2220	ARG	NE-CZ-NH2	6.83	123.71	120.30
1	M	2320	ARG	NE-CZ-NH2	6.67	123.64	120.30
1	M	2307	ARG	NE-CZ-NH2	6.41	123.51	120.30
1	M	2304	ARG	NE-CZ-NH2	6.30	123.45	120.30
1	M	2321	MET	CG-SD-CE	5.88	109.61	100.20

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	M	1268	0	1254	27	0
2	M	10	0	0	0	0
3	M	7	0	3	1	0
4	M	6	0	8	2	0
5	M	247	0	0	4	1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	1538	0	1265	27	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 11.

All (27) 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:M:2225:ASN:HD21	1:M:2227:LYS:HB2	1.38	0.89
1:M:2276:GLN:OE1	1:M:2281:LYS:HD3	1.74	0.87
1:M:2246:GLN:HA	5:M:1:HOH:O	1.76	0.83
1:M:2276:GLN:OE1	1:M:2281:LYS:CD	2.35	0.74
1:M:2276:GLN:OE1	1:M:2281:LYS:CE	2.38	0.72
1:M:2266:GLN:NE2	1:M:2303:THR:HA	2.05	0.70
1:M:2218:ALA:HB2	1:M:2248:VAL:HG11	1.80	0.64
1:M:2175:SER:HG	3:M:2330:CYS:N	1.99	0.60
1:M:2266:GLN:HE22	1:M:2303:THR:HA	1.65	0.58
1:M:2220:ARG:HH12	4:M:2331:GOL:H2	1.68	0.58
1:M:2276:GLN:OE1	1:M:2281:LYS:HE2	2.05	0.56
1:M:2225:ASN:HD21	1:M:2227:LYS:CB	2.16	0.56
1:M:2172:ASN:HD21	1:M:2174:CYS:HB3	1.72	0.53
1:M:2225:ASN:ND2	1:M:2227:LYS:H	2.07	0.52
1:M:2182:SER:O	1:M:2183:LYS:HB2	2.10	0.51
1:M:2220:ARG:HH12	4:M:2331:GOL:C2	2.23	0.50
1:M:2194:SER:OG	1:M:2222:GLN:HG2	2.12	0.50
1:M:2225:ASN:ND2	1:M:2227:LYS:HB2	2.17	0.49
1:M:2172:ASN:HD21	1:M:2174:CYS:CB	2.27	0.47
1:M:2225:ASN:HD22	1:M:2228:GLU:HG2	1.80	0.47
1:M:2172:ASN:ND2	1:M:2174:CYS:H	2.14	0.46
1:M:2215:ARG:HB2	5:M:191:HOH:O	2.16	0.46
1:M:2203:TRP:CD2	1:M:2220:ARG:HD3	2.52	0.44
1:M:2171:LEU:HD12	5:M:111:HOH:O	2.18	0.43
1:M:2225:ASN:CG	1:M:2227:LYS:H	2.24	0.41
1:M:2285:GLY:HA3	5:M:424:HOH:O	2.19	0.41
1:M:2234:PHE:O	1:M:2235:GLN:HB2	2.20	0.41

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)
5:M:129:HOH:O	5:M:201:HOH:O[2_664]	0.44	1.76

### 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	M	157/159 (99%)	147 (94%)	10 (6%)	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	M	143/143 (100%)	139 (97%)	4 (3%)	47 15

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	2172	ASN
1	M	2199	MET
1	M	2221	PRO
1	M	2225	ASN

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

Mol	Chain	Res	Type
1	M	2225	ASN
1	M	2266	GLN
1	M	2286	ASN
1	M	2329	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](#)

4 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	$\# Z  > 2$
3	CYS	M	2330	1	2,6,6	0.27	0	1,7,7	0.32	0
4	GOL	M	2331	-	5,5,5	0.99	0	5,5,5	0.30	0
2	SO4	M	2332	-	4,4,4	0.34	0	6,6,6	0.06	0
2	SO4	M	2333	-	4,4,4	0.33	0	6,6,6	0.08	0

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	CYS	M	2330	1	-	0/2/6/6	0/0/0/0
4	GOL	M	2331	-	-	0/4/4/4	0/0/0/0
2	SO4	M	2332	-	-	0/0/0/0	0/0/0/0
2	SO4	M	2333	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	M	2330	CYS	1	0
4	M	2331	GOL	2	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.