



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 07:37 AM GMT

PDB ID : 1GEN  
Title : C-TERMINAL DOMAIN OF GELATINASE A  
Authors : Libson, A.M.; Gittis, A.G.; Collier, I.E.; Marmer, B.L.; Goldberg, G.G.;  
Lattman, E.E.  
Deposited on : 1995-07-19  
Resolution : 2.15 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

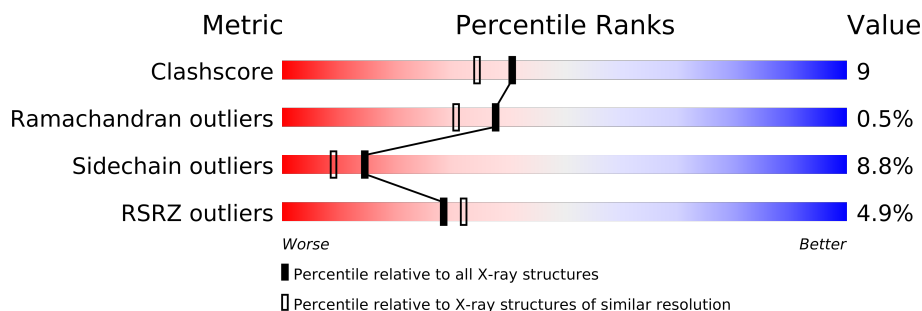
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.15 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	79885	1299 (2.18-2.14)
Ramachandran outliers	78287	1272 (2.18-2.14)
Sidechain outliers	78261	1272 (2.18-2.14)
RSRZ outliers	66119	1094 (2.18-2.14)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	218	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 2114 atoms, of which 437 are hydrogens and 0 are deuterium.

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 GELATINASE A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	200	Total	C	H	N	O	S	0	0	0
			1957	1067	335	262	289	4			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Cl	0	0
			1	1		

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Na	0	0
			1	1		

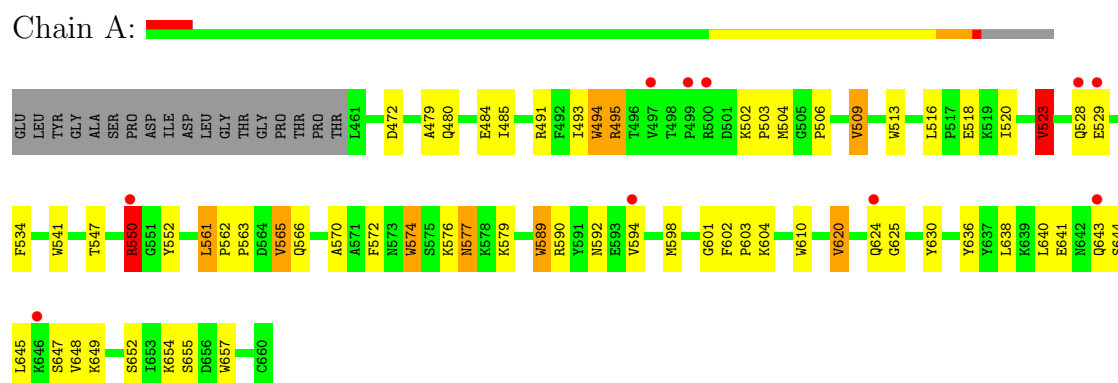
- Molecule 6 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	51	Total	H	O	0	0
			153	102	51		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

#### • Molecule 1: GELATINASE A



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.90Å 77.40Å 53.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.50 – 2.15 32.17 – 2.15	Depositor EDS
% Data completeness (in resolution range)	92.0 (6.50-2.15) 90.2 (32.17-2.15)	Depositor EDS
$R_{merge}$	0.03	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.26 (at 2.16Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.188 , 0.260 0.183 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	30.0	Xtriage
Anisotropy	0.261	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 43.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 12169 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2114	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.56% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, ZN, CA, CL

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	0.93	0/1675	1.63	35/2272 (1.5%)

There are no bond length outliers.

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	610	TRP	CD1-CG-CD2	9.28	113.72	106.30
1	A	494	TRP	CD1-CG-CD2	8.65	113.22	106.30
1	A	574	TRP	CD1-CG-CD2	8.50	113.10	106.30
1	A	513	TRP	CD1-CG-CD2	8.41	113.03	106.30
1	A	610	TRP	CE2-CD2-CG	-8.17	100.76	107.30
1	A	574	TRP	CE2-CD2-CG	-7.99	100.91	107.30
1	A	589	TRP	CD1-CG-CD2	7.74	112.49	106.30
1	A	494	TRP	CE2-CD2-CG	-7.38	101.40	107.30
1	A	589	TRP	CE2-CD2-CG	-7.28	101.47	107.30
1	A	523	VAL	CB-CA-C	-7.23	97.67	111.40
1	A	541	TRP	CE2-CD2-CG	-7.17	101.57	107.30
1	A	513	TRP	CE2-CD2-CG	-7.12	101.60	107.30
1	A	657	TRP	CD1-CG-CD2	7.12	112.00	106.30
1	A	657	TRP	CE2-CD2-CG	-7.10	101.62	107.30
1	A	552	TYR	CB-CG-CD2	-6.87	116.88	121.00
1	A	598	MET	CG-SD-CE	-6.78	89.35	100.20
1	A	541	TRP	CD1-CG-CD2	6.76	111.71	106.30
1	A	589	TRP	CB-CG-CD1	-6.16	119.00	127.00
1	A	494	TRP	CG-CD1-NE1	-5.88	104.22	110.10
1	A	541	TRP	CB-CG-CD1	-5.70	119.59	127.00
1	A	550	ARG	NE-CZ-NH1	5.68	123.14	120.30
1	A	620	VAL	CG1-CB-CG2	-5.62	101.92	110.90
1	A	610	TRP	CG-CD1-NE1	-5.60	104.50	110.10
1	A	541	TRP	CG-CD2-CE3	5.56	138.90	133.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	589	TRP	CG-CD2-CE3	5.52	138.87	133.90
1	A	561	LEU	CA-CB-CG	5.38	127.68	115.30
1	A	641	GLU	N-CA-CB	-5.30	101.05	110.60
1	A	625	GLY	N-CA-C	-5.30	99.86	113.10
1	A	630	TYR	CB-CG-CD1	-5.27	117.84	121.00
1	A	574	TRP	CG-CD1-NE1	-5.24	104.86	110.10
1	A	589	TRP	CG-CD1-NE1	-5.17	104.93	110.10
1	A	636	TYR	CB-CG-CD1	-5.17	117.90	121.00
1	A	649	LYS	CA-CB-CG	-5.11	102.15	113.40
1	A	504	MET	CA-CB-CG	5.01	121.82	113.30
1	A	610	TRP	CG-CD2-CE3	5.01	138.41	133.90

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1622	335	1257	30	0
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	1	0	0	0	0
6	A	51	102	0	3	0
All	All	1677	437	1257	30	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 9.

All (30) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:493:ILE:HG23	1:A:509:VAL:HG12	1.65	0.79
1:A:491:ARG:NH1	1:A:518:GLU:HG2	2.05	0.71
1:A:479:ALA:HB2	1:A:523:VAL:HB	1.80	0.63

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:491:ARG:HH12	1:A:518:GLU:HG2	1.65	0.59
1:A:601:GLY:O	1:A:604:LYS:HE3	2.07	0.55
1:A:574:TRP:CH2	1:A:576:LYS:HD2	2.43	0.54
1:A:480:GLN:HE21	1:A:624:GLN:HG3	1.72	0.53
1:A:520:ILE:CD1	1:A:523:VAL:HG22	2.38	0.53
1:A:472:ASP:HB3	6:A:314:HOH:H1	1.73	0.53
1:A:523:VAL:HG13	1:A:534:PHE:CD2	2.45	0.52
1:A:589:TRP:CE2	1:A:603:PRO:HB3	2.45	0.51
1:A:494:TRP:HB3	1:A:503:PRO:HB2	1.94	0.50
1:A:640:LEU:HD21	1:A:645:LEU:HD22	1.94	0.49
1:A:654:LYS:HG3	6:A:349:HOH:H2	1.78	0.48
1:A:502:LYS:HA	1:A:503:PRO:HD3	1.73	0.47
1:A:577:ASN:ND2	1:A:579:LYS:H	2.12	0.47
1:A:528:GLN:OE1	1:A:594:VAL:HG12	2.16	0.46
1:A:590:ARG:HH21	1:A:602:PHE:HE2	1.63	0.46
1:A:562:PRO:HD2	1:A:565:VAL:CG1	2.48	0.44
1:A:570:ALA:HB1	1:A:620:VAL:HG23	2.00	0.43
1:A:494:TRP:CZ2	1:A:506:PRO:HB3	2.54	0.43
1:A:652:SER:HB3	1:A:655:SER:OG	2.17	0.43
1:A:520:ILE:HD12	1:A:523:VAL:HG22	2.01	0.43
1:A:579:LYS:HG2	1:A:590:ARG:HD2	2.01	0.43
1:A:472:ASP:HB3	6:A:314:HOH:O	2.20	0.42
1:A:563:PRO:O	1:A:566:GLN:NE2	2.53	0.42
1:A:480:GLN:HA	1:A:484:GLU:O	2.20	0.41
1:A:485:ILE:O	1:A:495:ARG:HA	2.20	0.41
1:A:640:LEU:HD12	1:A:648:VAL:HG22	2.02	0.40
1:A:550:ARG:O	1:A:550:ARG:HG3	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	198/218 (91%)	182 (92%)	15 (8%)	1 (0%)	38	31



All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	550	ARG

### 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	170/184 (92%)	155 (91%)	15 (9%)	14 9

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	495	ARG
1	A	509	VAL
1	A	516	LEU
1	A	523	VAL
1	A	529	GLU
1	A	547	THR
1	A	561	LEU
1	A	565	VAL
1	A	572	PHE
1	A	577	ASN
1	A	592	ASN
1	A	638	LEU
1	A	643	GLN
1	A	644	SER
1	A	647	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	573	ASN
1	A	577	ASN
1	A	592	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

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.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	200/218 (91%)	0.28	10 (5%)	28 31	12, 26, 53, 78	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	550	ARG	4.6
1	A	528	GLN	4.5
1	A	643	GLN	3.5
1	A	500	ARG	2.9
1	A	529	GLU	2.9
1	A	497	VAL	2.4
1	A	594	VAL	2.3
1	A	646	LYS	2.2
1	A	499	PRO	2.1
1	A	624	GLN	2.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	CL	A	303	1/1	0.14	-0.38	18,18,18,18	0
2	ZN	A	301	1/1	0.05	-1.53	36,36,36,36	0
5	NA	A	304	1/1	0.09	-1.77	16,16,16,16	0
3	CA	A	302	1/1	0.07	-4.09	15,15,15,15	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.