



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2023 – 04:19 AM EDT

PDB ID : 2D26  
Title : Active site distortion is sufficient for proteinase inhibit second crystal structure of covalent serpin-proteinase complex  
Authors : Dementiev, A.; Dobo, J.; Gettins, P.G.  
Deposited on : 2005-09-03  
Resolution : 3.30 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](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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

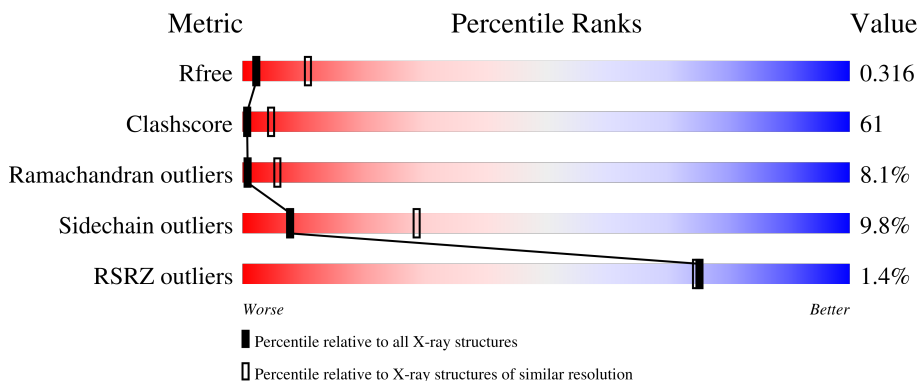
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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))
$R_{free}$	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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  $\geq 3$ , 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  $\leq 5\%$ . 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.

Mol	Chain	Length	Quality of chain
1	A	358	
2	B	36	
3	C	240	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4189 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 Alpha-1-antitrypsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	336	2302	1474	394	429	5	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	GLU	engineered mutation	UNP P01009
A	51	LEU	PHE	engineered mutation	UNP P01009
A	59	ALA	THR	engineered mutation	UNP P01009
A	68	ALA	THR	engineered mutation	UNP P01009
A	70	GLY	ALA	engineered mutation	UNP P01009
A	101	HIS	ARG	engineered mutation	UNP P01009
A	232	SER	CYS	engineered mutation	UNP P01009

- Molecule 2 is a protein called Alpha-1-antitrypsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	30	221	148	34	38	1	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	374	ILE	MET	engineered mutation	UNP P01009
B	381	ALA	SER	engineered mutation	UNP P01009
B	387	ARG	LYS	engineered mutation	UNP P01009

- Molecule 3 is a protein called Elastase-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	223	1591	1001	288	292	10	0	0	0

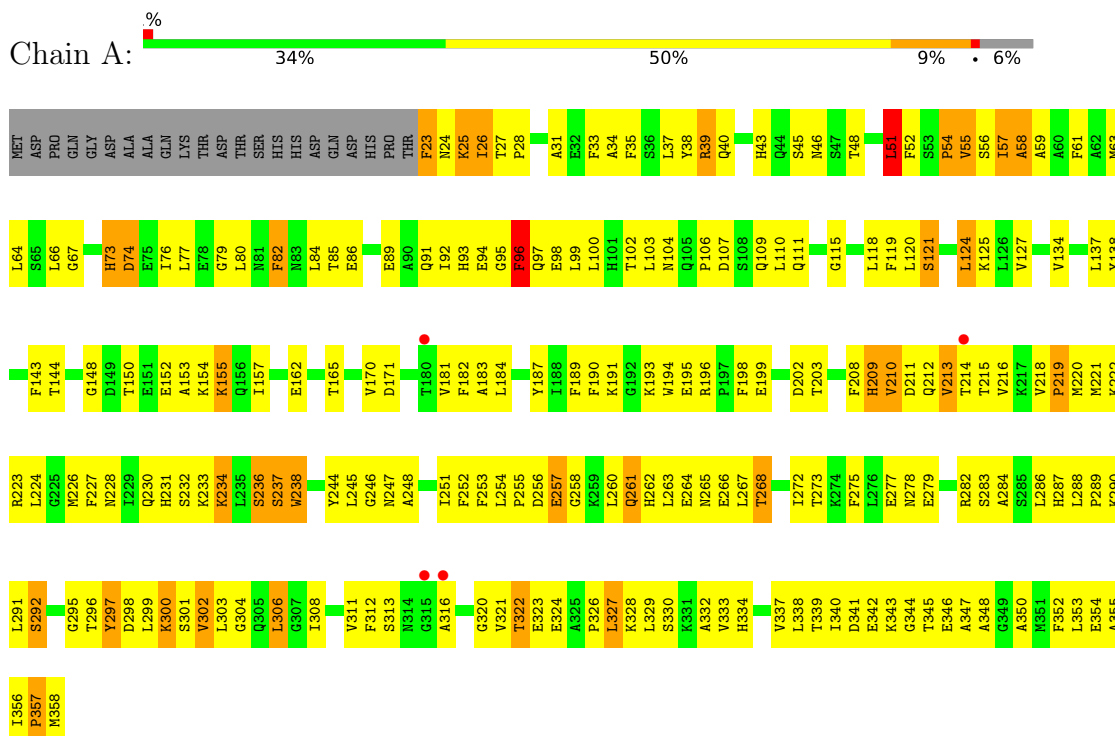
- Molecule 4 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	A	43	Total 43	O 43	0	0
4	B	3	Total 3	O 3	0	0
4	C	29	Total 29	O 29	0	0

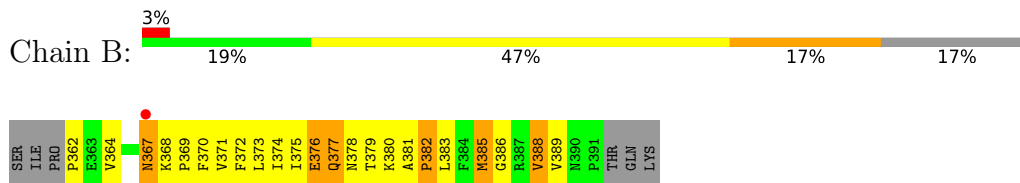
### 3 Residue-property plots

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 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.

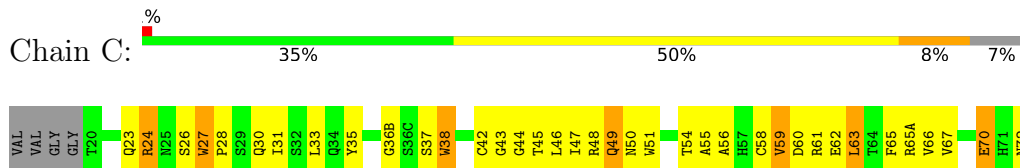
- Molecule 1: Alpha-1-antitrypsin

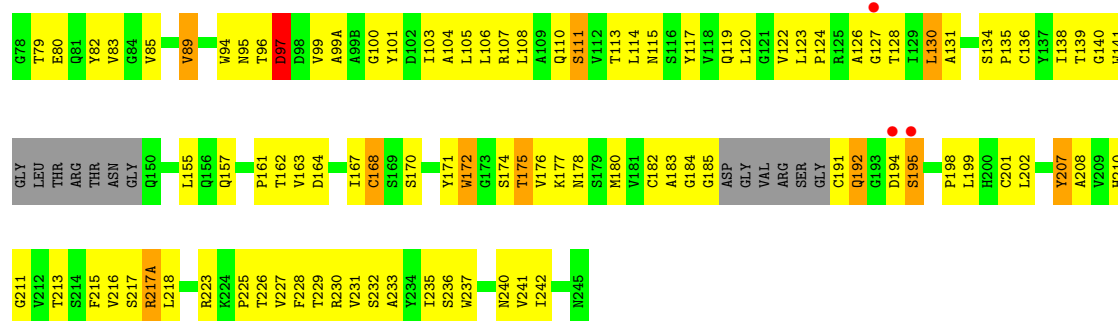


- Molecule 2: Alpha-1-antitrypsin



- Molecule 3: Elastase-1





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	109.83Å 85.18Å 76.26Å 90.00° 121.01° 90.00°	Depositor
Resolution (Å)	50.00 – 3.30 39.16 – 2.93	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-3.30) 91.3 (39.16-2.93)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.50 (at 2.90Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.251 , 0.312 0.254 , 0.316	Depositor DCC
$R_{free}$ test set	821 reflections (6.58%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	80.8	Xtrriage
Anisotropy	0.373	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 184.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	4189	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.26% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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](#)

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	0.49	1/2353 (0.0%)	0.82	5/3222 (0.2%)
2	B	0.56	0/228	0.83	0/311
3	C	0.62	3/1628 (0.2%)	0.79	0/2229
All	All	0.55	4/4209 (0.1%)	0.81	5/5762 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	238	TRP	NE1-CE2	8.70	1.48	1.37
3	C	38	TRP	NE1-CE2	8.69	1.48	1.37
3	C	27	TRP	NE1-CE2	8.67	1.48	1.37
3	C	172	TRP	NE1-CE2	8.66	1.48	1.37

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	51	LEU	CB-CG-CD2	10.07	128.12	111.00
1	A	278	ASN	N-CA-C	-9.67	84.90	111.00
1	A	278	ASN	CB-CA-C	-9.25	91.89	110.40
1	A	279	GLU	N-CA-CB	-9.00	94.40	110.60
1	A	51	LEU	CA-CB-CG	-6.20	101.04	115.30

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	A	2302	0	1920	271	2
2	B	221	0	198	39	0
3	C	1591	0	1401	171	1
4	A	43	0	0	5	0
4	B	3	0	0	0	0
4	C	29	0	0	5	1
All	All	4189	0	3519	468	2

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 61.

The worst 5 of 468 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:299:LEU:HB2	1:A:332:ALA:CB	1.54	1.37
1:A:299:LEU:CD1	1:A:332:ALA:HB1	1.57	1.33
3:C:122:VAL:O	3:C:208:ALA:HB1	1.43	1.16
1:A:45:SER:O	1:A:46:ASN:ND2	1.79	1.14
1:A:299:LEU:HB2	1:A:332:ALA:HB3	1.19	1.14

All (2) 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:23:PHE:N	3:C:127:GLY:CA[3_445]	1.71	0.49
1:A:24:ASN:ND2	4:C:269:HOH:O[3_445]	1.80	0.40

## 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	334/358 (93%)	244 (73%)	57 (17%)	33 (10%)	<b>0</b> <b>3</b>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	28/36 (78%)	18 (64%)	7 (25%)	3 (11%)	0	3
3	C	217/240 (90%)	152 (70%)	54 (25%)	11 (5%)	2	13
All	All	579/634 (91%)	414 (72%)	118 (20%)	47 (8%)	1	6

5 of 47 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	25	LYS
1	A	55	VAL
1	A	210	VAL
1	A	213	VAL
1	A	257	GLU

### 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	182/311 (58%)	166 (91%)	16 (9%)	10	33
2	B	21/34 (62%)	15 (71%)	6 (29%)	0	1
3	C	144/198 (73%)	132 (92%)	12 (8%)	11	36
All	All	347/543 (64%)	313 (90%)	34 (10%)	8	29

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	130	LEU
3	C	168	CYS
3	C	217(A)	ARG
1	A	306	LEU
1	A	297	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	192	GLN
3	C	240	ASN
3	C	245	ASN
3	C	210	HIS
2	B	367	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](#)

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	336/358 (93%)	-0.27	4 (1%) 79 78	31, 71, 96, 98	0
2	B	30/36 (83%)	-0.28	1 (3%) 46 44	29, 52, 86, 92	0
3	C	223/240 (92%)	-0.26	3 (1%) 77 77	28, 54, 91, 99	5 (2%)
All	All	589/634 (92%)	-0.26	8 (1%) 75 75	28, 62, 95, 99	5 (0%)

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	180	THR	3.8
1	A	315	GLY	3.1
2	B	367	ASN	2.7
3	C	194	ASP	2.5
1	A	316	ALA	2.3

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

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.