



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 11:34 AM GMT

PDB ID : 3PEP
Title : REVISED 2.3 ANGSTROMS STRUCTURE OF PORCINE PEPSIN. EVIDENCE FOR A FLEXIBLE SUBDOMAIN
Authors : Abad-Zapatero, C.; Erickson, J.W.
Deposited on : 1989-10-24
Resolution : 2.30 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
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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

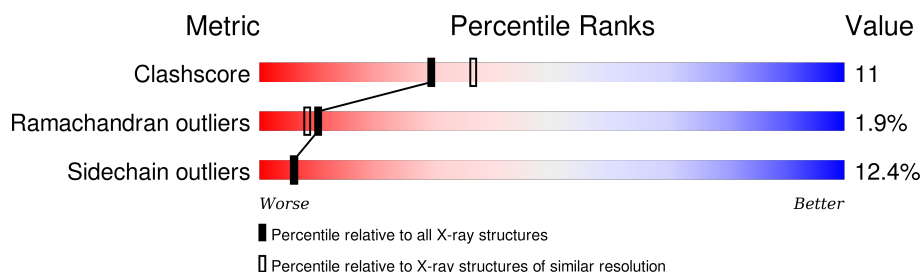
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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))
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)

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 ≥ 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	326	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 2641 atoms, of which 2 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.

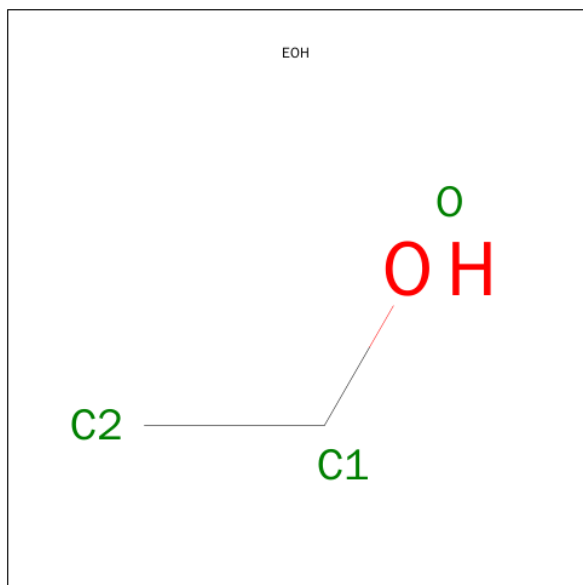
- Molecule 1 is a protein called PEPSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	326	2429	1530	365	524	10	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ILE	DELETION	UNP P00791
A	263	ASP	ASN	CONFLICT	UNP P00791

- Molecule 2 is ETHANOL (three-letter code: EOH) (formula: C₂H₆O).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	H		
2	A	1	3	2	1	0	0
2	A	1	3	2	1	0	0

- Molecule 3 is water.

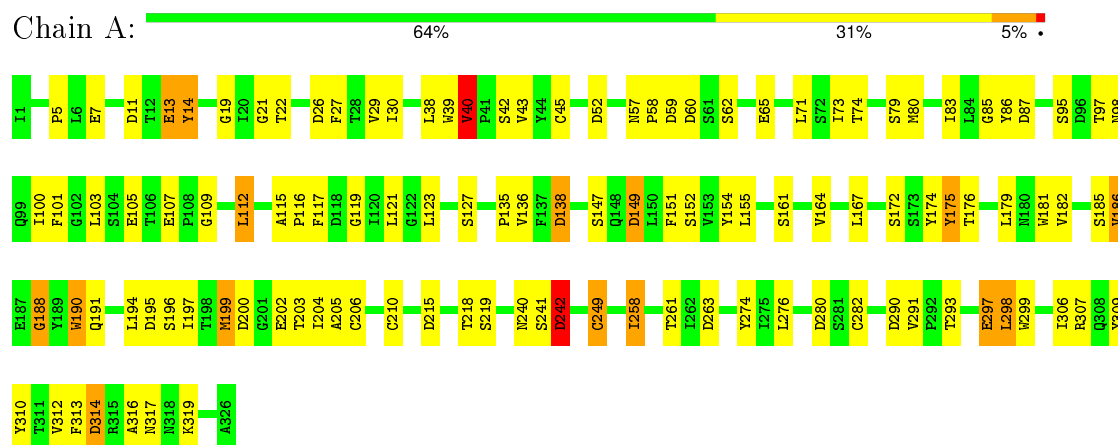
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	206	Total 206	O 206	0	0

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.

Note EDS was not executed.

- Molecule 1: PEPSIN



4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 1 21	Depositor
Cell constants a, b, c, α , β , γ	55.27Å 73.82Å 36.44Å 90.00° 90.00° 103.38°	Depositor
Resolution (Å)	5.00 – 2.30	Depositor
% Data completeness (in resolution range)	(Not available) (5.00-2.30)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.171 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2641	wwPDB-VP
Average B, all atoms (Å ²)	26.0	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: EOH

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	1.01	2/2484 (0.1%)	1.64	32/3400 (0.9%)

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	1	5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	40	VAL	N-CA	-10.16	1.26	1.46
1	A	19	GLY	N-CA	5.13	1.53	1.46

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	40	VAL	CA-C-O	-10.63	97.77	120.10
1	A	39	TRP	C-N-CA	10.44	147.79	121.70
1	A	40	VAL	CA-C-N	8.80	141.73	117.10
1	A	40	VAL	N-CA-CB	7.30	127.55	111.50
1	A	307	ARG	NE-CZ-NH2	7.07	123.83	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	40	VAL	CA

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	154	TYR	Sidechain
1	A	175	TYR	Sidechain
1	A	310	TYR	Sidechain
1	A	40	VAL	Mainchain,Peptide

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	2429	0	2252	53	0
2	A	4	2	0	0	0
3	A	206	0	0	1	0
All	All	2639	2	2252	53	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 11.

The worst 5 of 53 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:27:PHE:CZ	1:A:40:VAL:HG21	2.09	0.86
1:A:27:PHE:CE1	1:A:40:VAL:HG21	2.19	0.77
1:A:149:ASP:HB3	1:A:316:ALA:HB2	1.77	0.67
1:A:27:PHE:CE2	1:A:40:VAL:CG2	2.78	0.66
1:A:199:MET:HB2	1:A:204:ILE:HD13	1.84	0.60

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	324/326 (99%)	298 (92%)	20 (6%)	6 (2%)	10 8

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	293	THR
1	A	297	GLU
1	A	109	GLY
1	A	136	VAL
1	A	62	SER

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	275/275 (100%)	241 (88%)	34 (12%)	6 6

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

Mol	Chain	Res	Type
1	A	161	SER
1	A	185	SER
1	A	297	GLU
1	A	174	TYR
1	A	60	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	230	ASN
1	A	317	ASN

5.3.3 RNA ⓘ

There are no RNA molecules 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 ⓘ

2 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	$\# Z > 2$
2	EOH	A	901	-	0,1,2	0.00	-	0,0,1	0.00	-
2	EOH	A	902	-	0,1,2	0.00	-	0,0,1	0.00	-

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
2	EOH	A	901	-	-	0/0/0/0	0/0/0/0
2	EOH	A	902	-	-	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.

No monomer is involved in short contacts.

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.