wwPDB X-ray Structure Validation Summary Report (i)

May 23, 2020 – 02:26 pm BST

PDB ID	:	1EPM
Title	:	A STRUCTURAL COMPARISON OF 21 INHIBITOR COMPLEXES OF THE ASPARTIC PROTEINASE FROM ENDOTHIA PARASITICA
Authors	:	Crawford, M.; Cooper, J.B.; Strop, P.; Blundell, T.L.
Deposited on	:	1994-07-27
Resolution	:	1.60 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

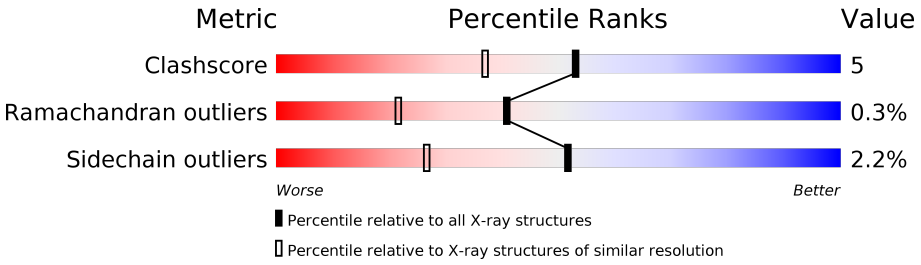
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.60 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	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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 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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	E	330	65% 32%
2	I	8	25% 25% 25% 25%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2759 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.

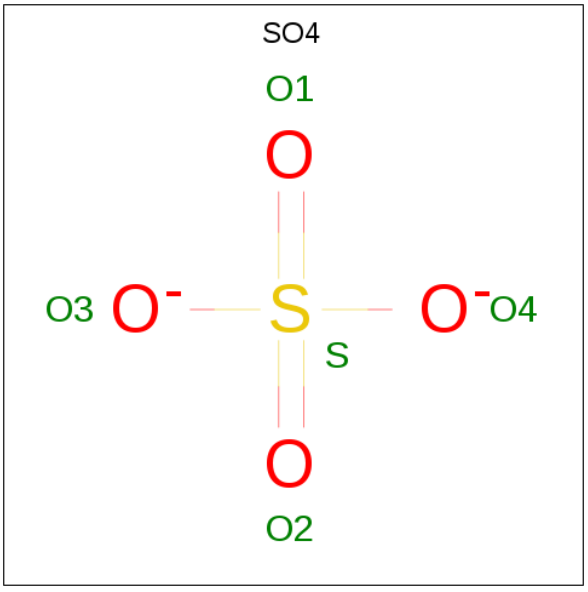
- Molecule 1 is a protein called ENDOTHIAPEPSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	330	Total	C	N	O	S	0	0	0
			2389	1514	366	507	2			

- Molecule 2 is a protein called PS2, THR-PHE-GLN-ALA-PSA-LEU-ARG-GLU.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	I	8	Total	C	N	O	0	0	0
			74	49	12	13			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	E	1	Total O S 5 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	E	274	Total O 274 274	0	0
4	I	7	Total O 7 7	0	0

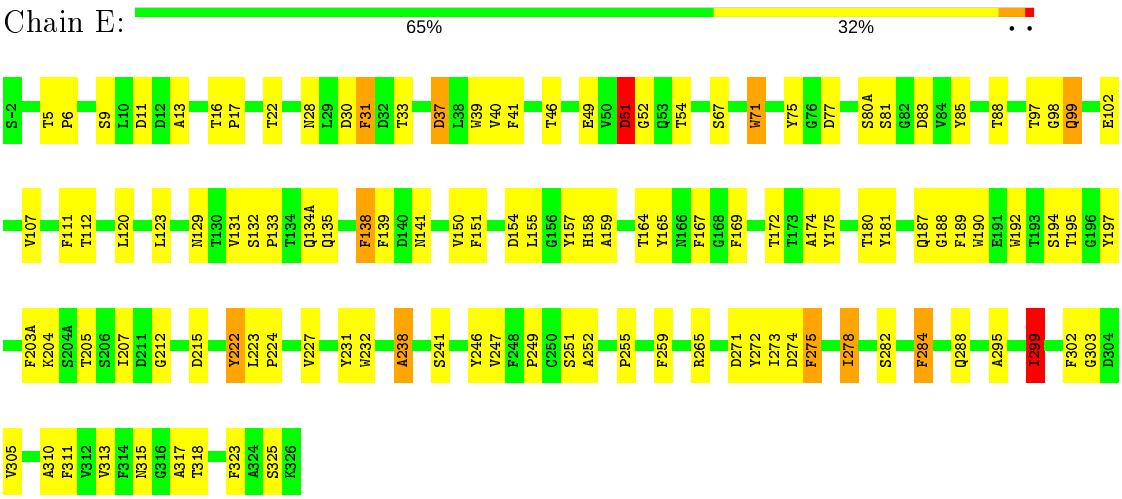
3 Residue-property plots (i)

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. 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. 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.

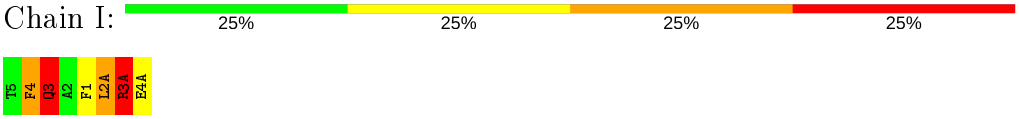
- Molecule 1: ENDOTHIAPEPSIN

Chain E:



- Molecule 2: PS2, THR-PHE-GLN-ALA-PSA-LEU-ARG-GLU

Chain I:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	43.11Å 75.88Å 42.98Å 90.00° 96.76° 90.00°	Depositor
Resolution (Å)	20.00 – 1.60	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-1.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	RESTRAIN	Depositor
R, R_{free}	0.170 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2759	wwPDB-VP
Average B, all atoms (Å^2)	0.0	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: PSA, 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	E	1.64	22/2445 (0.9%)	2.52	156/3345 (4.7%)
2	I	1.19	1/59 (1.7%)	2.90	7/76 (9.2%)
All	All	1.63	23/2504 (0.9%)	2.53	163/3421 (4.8%)

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
2	I	0	3

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	134(A)	GLN	CD-OE1	7.21	1.39	1.24
1	E	246	TYR	CD1-CE1	7.08	1.50	1.39
1	E	188	GLY	N-CA	6.71	1.56	1.46
1	E	282	SER	CB-OG	6.51	1.50	1.42
1	E	49	GLU	CD-OE2	6.06	1.32	1.25

The worst 5 of 163 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	275	PHE	CB-CG-CD1	-17.27	108.71	120.80
1	E	51	ASP	CA-C-O	15.94	153.56	120.10
1	E	265	ARG	NE-CZ-NH1	15.11	127.85	120.30
1	E	157	TYR	CD1-CE1-CZ	-15.00	106.30	119.80
1	E	157	TYR	CG-CD1-CE1	14.64	133.02	121.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	I	1	PSA	Mainchain,Peptide
2	I	3(A)	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	E	2389	0	2280	22	0
2	I	74	0	70	12	0
3	E	15	0	0	0	0
4	E	274	0	0	1	1
4	I	7	0	0	0	0
All	All	2759	0	2350	26	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 5.

The worst 5 of 26 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)
2:I:2(A):LEU:O	2:I:3(A):ARG:CB	1.76	1.25
2:I:2(A):LEU:O	2:I:3(A):ARG:HB2	0.93	0.80
2:I:2(A):LEU:HG	2:I:3(A):ARG:N	1.97	0.79
1:E:275:PHE:CE2	2:I:4:PHE:HE1	2.04	0.76
1:E:299:ILE:HD11	2:I:3(A):ARG:HD3	1.71	0.73

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)
4:E:564:HOH:O	4:E:575:HOH:O[2_555]	0.29	1.91

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	E	328/330 (99%)	326 (99%)	2 (1%)	0	100 100
2	I	5/8 (62%)	4 (80%)	0	1 (20%)	0 0
All	All	333/338 (98%)	330 (99%)	2 (1%)	1 (0%)	41 21

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	I	3(A)	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	E	263/263 (100%)	259 (98%)	4 (2%)	65 44
2	I	6/6 (100%)	4 (67%)	2 (33%)	0 0
All	All	269/269 (100%)	263 (98%)	6 (2%)	52 27

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	241	SER
2	I	3(A)	ARG
1	E	299	ILE
1	E	51	ASP
2	I	3	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	141	ASN
2	I	3	GLN
1	E	187	GLN
1	E	134(A)	GLN
1	E	166	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)

1 non-standard protein/DNA/RNA residue is 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	PSA	I	1	2	14,14,15	1.20	1 (7%)	15,17,19	1.63	4 (26%)

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	PSA	I	1	2	-	3/11/11/12	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	1	PSA	CZ-CE2	2.73	1.45	1.38

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	1	PSA	CB-CA-CH	-2.92	107.07	111.84
2	I	1	PSA	CE1-CD1-CG	2.71	124.80	120.63
2	I	1	PSA	CE2-CD2-CG	2.71	124.79	120.63
2	I	1	PSA	CD2-CG-CD1	-2.09	114.87	118.17

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	I	1	PSA	CA-CB-CG-CD1
2	I	1	PSA	CA-CB-CG-CD2
2	I	1	PSA	O-C-CM-CH

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

3 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	SO4	E	329	-	4,4,4	0.18	0	6,6,6	0.54	0
3	SO4	E	328	-	4,4,4	0.19	0	6,6,6	0.66	0
3	SO4	E	327	-	4,4,4	0.16	0	6,6,6	0.50	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

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.