Full wwPDB X-ray Structure Validation Report (i)

May 18, 2020 – 11:50 pm BST

PDB ID	:	1EPL
Title	:	A STRUCTURAL COMPARISON OF 21 INHIBITOR COMPLEXES OF THE ASPARTIC PROTEINASE FROM ENDOTHIA PARASITICA
Authors	:	Al-Karadaghi, S.; Cooper, J.B.; Strop, P.; Blundell, T.L.
Deposited on	:	1994-07-27
Resolution	:	2.00 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*
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:

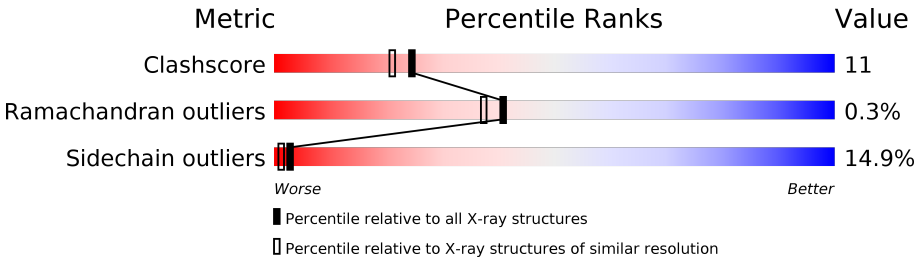
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (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 2.00 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	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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	67% 27% 6%
2	I	6	83% 17%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2764 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 PS1, PRO-LEU-GLU-PSA-ARG-LEU.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	I	6	Total	C	N	O	0	0	0
			58	39	9	10			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	308	Total	O	0	0
			308	308		
3	I	9	Total	O	0	0
			9	9		

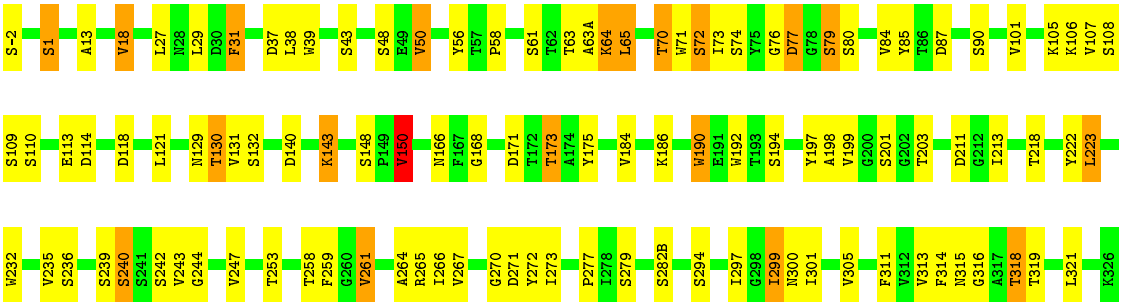
3 Residue-property plots (i)

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.

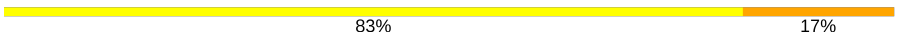
- Molecule 1: ENDOTHIAPEPSIN

Chain E:



- Molecule 2: PS1, PRO-LEU-GLU-PSA-ARG-LEU

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	53.74Å 74.11Å 45.71Å 90.00° 109.76° 90.00°	Depositor
Resolution (Å)	8.00 – 2.00	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-2.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	RESTRAIN	Depositor
R, R_{free}	0.188 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2764	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

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	0.80	4/2445 (0.2%)	1.40	34/3345 (1.0%)
2	I	1.02	0/43	1.56	1/53 (1.9%)
All	All	0.80	4/2488 (0.2%)	1.41	35/3398 (1.0%)

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	E	0	2
2	I	0	2
All	All	0	4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	106	LYS	C-N	-7.77	1.16	1.34
1	E	297	ILE	N-CA	7.12	1.60	1.46
1	E	297	ILE	CA-C	-6.32	1.36	1.52
1	E	244	GLY	C-N	-6.23	1.21	1.33

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	222	TYR	CB-CG-CD2	-12.08	113.75	121.00
1	E	197	TYR	CB-CG-CD1	-9.51	115.29	121.00
1	E	261	VAL	CA-CB-CG2	9.25	124.78	110.90
1	E	85	TYR	CB-CG-CD2	-7.92	116.25	121.00
1	E	87	ASP	CB-CG-OD2	6.95	124.56	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	184	VAL	CA-CB-CG2	6.61	120.81	110.90
1	E	175	TYR	CB-CG-CD1	-6.56	117.07	121.00
1	E	272	TYR	CB-CG-CD2	-6.43	117.14	121.00
1	E	106	LYS	CA-CB-CG	6.41	127.50	113.40
1	E	106	LYS	C-N-CA	6.36	137.61	121.70
1	E	271	ASP	CB-CG-OD2	6.32	123.99	118.30
2	I	2(A)	ARG	NE-CZ-NH2	6.28	123.44	120.30
1	E	313	VAL	CA-CB-CG2	6.08	120.02	110.90
1	E	101	VAL	CA-CB-CG2	6.03	119.94	110.90
1	E	267	VAL	CA-CB-CG2	5.93	119.79	110.90
1	E	305	VAL	CA-CB-CG1	5.91	119.76	110.90
1	E	18	VAL	CA-CB-CG2	5.74	119.51	110.90
1	E	150	VAL	CA-CB-CG2	5.70	119.44	110.90
1	E	56	TYR	CB-CG-CD2	-5.66	117.60	121.00
1	E	85	TYR	CB-CG-CD1	5.63	124.38	121.00
1	E	222	TYR	CD1-CG-CD2	5.61	124.08	117.90
1	E	311	PHE	CB-CG-CD2	-5.57	116.90	120.80
1	E	253	THR	N-CA-CB	5.55	120.85	110.30
1	E	223	LEU	CB-CG-CD1	5.51	120.37	111.00
1	E	106	LYS	O-C-N	-5.46	113.96	122.70
1	E	39	TRP	CA-CB-CG	5.45	124.06	113.70
1	E	222	TYR	CG-CD1-CE1	-5.41	116.97	121.30
1	E	84	VAL	CA-CB-CG2	5.33	118.90	110.90
1	E	50	VAL	CA-CB-CG2	5.31	118.87	110.90
1	E	253	THR	O-C-N	5.28	131.15	122.70
1	E	261	VAL	CA-CB-CG1	5.18	118.67	110.90
1	E	140	ASP	CB-CG-OD1	5.11	122.90	118.30
1	E	199	VAL	CA-CB-CG2	5.10	118.55	110.90
1	E	31	PHE	CB-CG-CD1	-5.02	117.28	120.80
1	E	31	PHE	CD1-CG-CD2	5.02	124.83	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	318	THR	Mainchain,Peptide
2	I	1	PSA	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	E	2389	0	2278	51	0
2	I	58	0	60	5	0
3	E	308	0	0	3	0
3	I	9	0	0	0	0
All	All	2764	0	2338	52	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.

All (52) 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:E:171:ASP:OD1	1:E:173:THR:HB	1.86	0.76
1:E:71:TRP:HB2	1:E:130:THR:HG22	1.76	0.67
1:E:299:ILE:HG12	3:E:492:HOH:O	1.95	0.66
1:E:71:TRP:HZ3	1:E:73:ILE:HG22	1.59	0.66
1:E:77:ASP:HB3	1:E:79:SER:H	1.61	0.65
1:E:71:TRP:HB2	1:E:130:THR:CG2	2.30	0.61
1:E:270:GLY:HA2	1:E:273:ILE:HD12	1.82	0.61
1:E:1:SER:HB3	1:E:166:ASN:ND2	2.16	0.60
1:E:18:VAL:CG1	1:E:29:LEU:HG	2.32	0.59
1:E:18:VAL:HG11	1:E:29:LEU:HG	1.84	0.59
1:E:240:SER:HB3	1:E:243:VAL:HG13	1.87	0.56
1:E:113:GLU:HG3	3:E:447:HOH:O	2.06	0.56
1:E:43:SER:OG	1:E:58:PRO:HD2	2.07	0.54
1:E:213:ILE:HG23	1:E:299:ILE:HD13	1.90	0.54
1:E:71:TRP:CZ3	1:E:73:ILE:HG22	2.43	0.53
1:E:301:ILE:HD11	2:I:3(A):LEU:CD2	2.39	0.52
1:E:129:ASN:ND2	1:E:131:VAL:H	2.08	0.52
1:E:13:ALA:HB2	2:I:3:LEU:HD12	1.92	0.51
1:E:213:ILE:HG23	1:E:299:ILE:CD1	2.41	0.51
1:E:70:THR:HG22	1:E:132:SER:HB3	1.92	0.51
1:E:314:PHE:CD1	1:E:314:PHE:N	2.80	0.50
1:E:76:GLY:HA3	2:I:2:GLU:OE2	2.13	0.49
2:I:4:PRO:C	2:I:3:LEU:HG	2.33	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:218:THR:HA	2:I:3:LEU:O	2.12	0.48
1:E:243:VAL:HG21	1:E:247:VAL:CG1	2.45	0.47
1:E:27:LEU:HD23	1:E:118:ASP:HB3	1.96	0.46
1:E:72:SER:H	1:E:130:THR:HG23	1.81	0.46
1:E:61:SER:HB3	1:E:63(A):ALA:HB2	1.97	0.45
1:E:264:ALA:C	1:E:265:ARG:HG2	2.38	0.44
1:E:273:ILE:HG22	1:E:273:ILE:O	2.18	0.44
1:E:18:VAL:HG11	1:E:29:LEU:CD1	2.47	0.44
1:E:150:VAL:HG23	1:E:315:ASN:HA	1.98	0.44
1:E:64:LYS:HG3	1:E:65:LEU:N	2.32	0.44
1:E:192:TRP:CZ3	1:E:194:SER:HA	2.53	0.44
1:E:190:TRP:CE3	1:E:321:LEU:CD1	3.01	0.43
1:E:232:TRP:CZ3	1:E:235:VAL:HG21	2.52	0.43
1:E:38:LEU:HD23	1:E:38:LEU:C	2.39	0.43
1:E:31:PHE:N	1:E:31:PHE:CD1	2.85	0.43
1:E:294:SER:HB3	1:E:300:ASN:HD22	1.83	0.43
1:E:31:PHE:HD1	1:E:31:PHE:N	2.16	0.42
1:E:18:VAL:HG11	1:E:29:LEU:CG	2.49	0.42
1:E:240:SER:CB	1:E:243:VAL:HG13	2.49	0.42
1:E:129:ASN:HD21	1:E:131:VAL:HB	1.84	0.41
1:E:198:ALA:HB3	1:E:258:THR:HB	2.01	0.41
1:E:130:THR:O	1:E:130:THR:HG23	2.20	0.41
1:E:113:GLU:CB	3:E:447:HOH:O	2.67	0.41
1:E:143:LYS:HE3	1:E:316:GLY:O	2.21	0.41
1:E:294:SER:CB	1:E:300:ASN:ND2	2.84	0.41
1:E:148:SER:O	1:E:168:GLY:HA2	2.21	0.40
1:E:259:PHE:CE1	1:E:266:ILE:HG13	2.56	0.40
1:E:277:PRO:HA	1:E:282(B):SER:O	2.20	0.40
1:E:211:ASP:OD1	1:E:211:ASP:C	2.59	0.40

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	E	328/330 (99%)	320 (98%)	7 (2%)	1 (0%)	41 37
2	I	3/6 (50%)	3 (100%)	0	0	100 100
All	All	331/336 (98%)	323 (98%)	7 (2%)	1 (0%)	41 37

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	190	TRP

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	E	263/263 (100%)	224 (85%)	39 (15%)	3 1
2	I	5/5 (100%)	4 (80%)	1 (20%)	1 0
All	All	268/268 (100%)	228 (85%)	40 (15%)	3 1

All (40) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	-2	SER
1	E	1	SER
1	E	37	ASP
1	E	48	SER
1	E	50	VAL
1	E	63	THR
1	E	64	LYS
1	E	65	LEU
1	E	70	THR
1	E	72	SER
1	E	74	SER
1	E	77	ASP
1	E	79	SER
1	E	80	SER
1	E	90	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	105	LYS
1	E	107	VAL
1	E	108	SER
1	E	109	SER
1	E	110	SER
1	E	114	ASP
1	E	121	LEU
1	E	130	THR
1	E	143	LYS
1	E	150	VAL
1	E	173	THR
1	E	186	LYS
1	E	201	SER
1	E	203	THR
1	E	223	LEU
1	E	236	SER
1	E	239	SER
1	E	240	SER
1	E	242	SER
1	E	261	VAL
1	E	279	SER
1	E	299	ILE
1	E	318	THR
1	E	319	THR
2	I	2(A)	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	28	ASN
1	E	99	GLN
1	E	129	ASN
1	E	135	GLN
1	E	141	ASN
1	E	166	ASN
1	E	300	ASN

5.3.3 RNA

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	0.79	0	15,17,19	1.03	1 (6%)

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	-	4/11/11/12	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	1	PSA	CB-CA-CH	-3.09	106.79	111.84

There are no chirality outliers.

All (4) torsion outliers are listed below:

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

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	E	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	106:LYS	C	107:VAL	N	1.16

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.