Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 07:15 PM GMT

PDB ID	:	1EST
Title	:	THE ATOMIC STRUCTURE OF CRYSTALLINE PORCINE PANCREATIC ELASTASE AT 2.5 ANGSTROMS RESOLUTION. COMPARISONS WITH THE STRUCTURE OF ALPHA-CHYMOTRYPSIN
Authors	:	Sawyer, L.; Shotton, D.M.; Watson, H.C.
Deposited on	:	1976-05-17
Resolution	:	2.50 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
http://wwpdb.org/validation/2016/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:

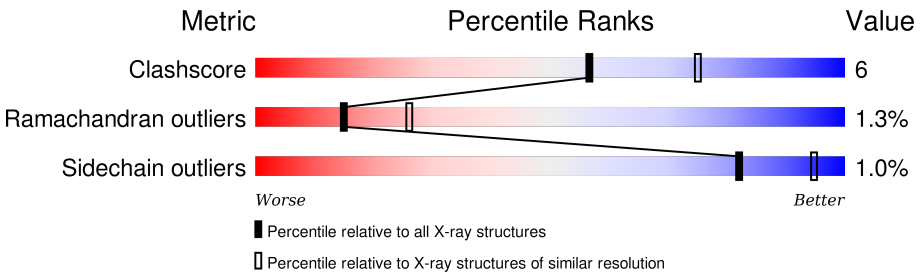
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.50 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	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)

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 <=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	240	55% 37% 8%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1858 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.

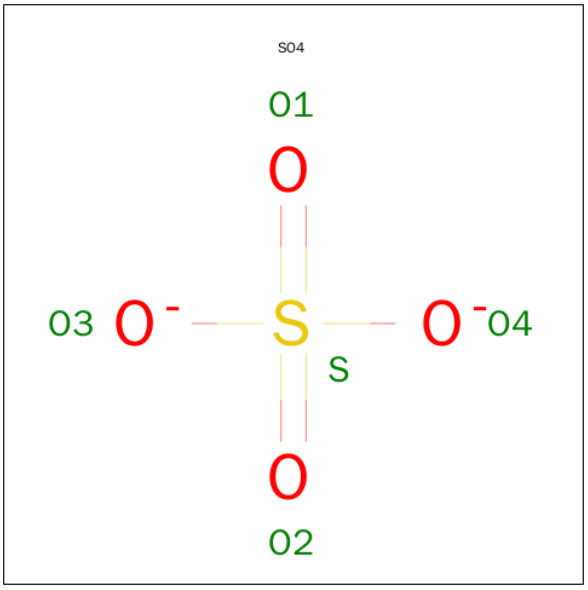
- Molecule 1 is a protein called PORCINE PANCREATIC ELASTASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	240	1822	1135	330	347	10	0	0	0

There is a discrepancy between the modelled and reference sequences:

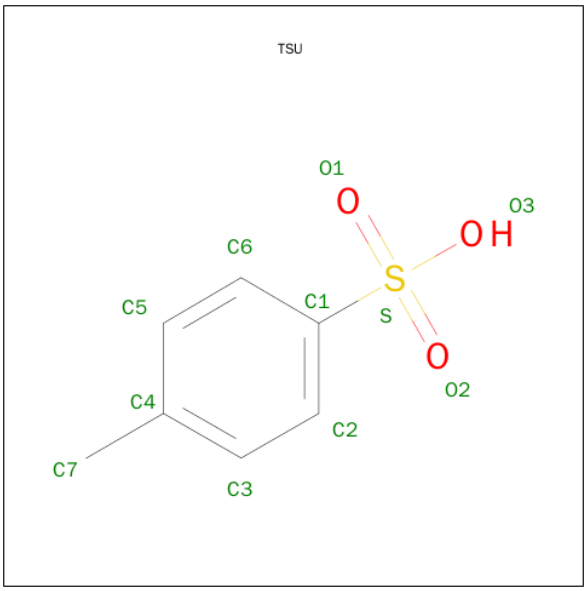
Chain	Residue	Modelled	Actual	Comment	Reference
A	77	ASN	ASP	CONFLICT	UNP P00772

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 3 is PARA-TOLUENE SULFONATE (three-letter code: TSU) (formula: $C_7H_8O_3S$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
3	A	1	10	7	2	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	25	25	25	0	0

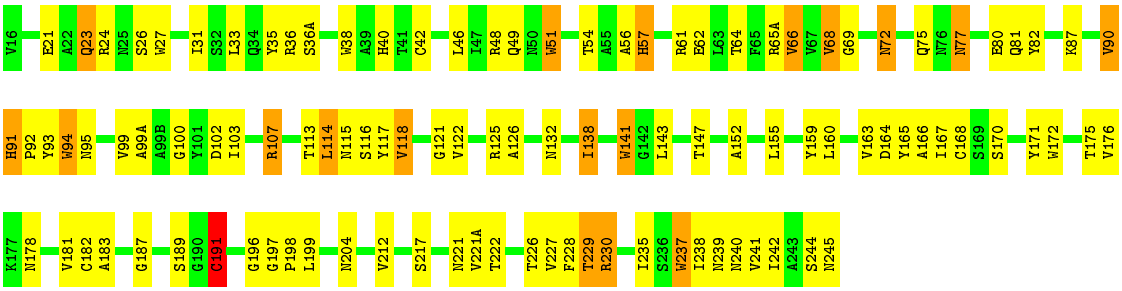
3 Residue-property plots (i)

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

Chain A: 55% 37% 8%



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.50Å 58.00Å 75.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	unknown	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1858	wwPDB-VP
Average B, all atoms ($Å^2$)	0.0	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: TSU, 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	A	0.72	7/1862 (0.4%)	0.93	0/2543

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	0	87

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	172	TRP	NE1-CE2	-7.47	1.27	1.37
1	A	51	TRP	NE1-CE2	-7.42	1.27	1.37
1	A	38	TRP	NE1-CE2	-7.41	1.27	1.37
1	A	237	TRP	NE1-CE2	-7.38	1.27	1.37
1	A	94	TRP	NE1-CE2	-7.37	1.27	1.37
1	A	27	TRP	NE1-CE2	-7.36	1.27	1.37
1	A	141	TRP	NE1-CE2	-7.28	1.28	1.37

There are no bond angle outliers.

There are no chirality outliers.

All (87) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	100	GLY	Mainchain
1	A	103	ILE	Mainchain
1	A	107	ARG	Mainchain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	113	THR	Mainchain
1	A	114	LEU	Mainchain
1	A	115	ASN	Sidechain
1	A	116	SER	Mainchain
1	A	117	TYR	Sidechain
1	A	118	VAL	Mainchain
1	A	121	GLY	Mainchain
1	A	122	VAL	Mainchain
1	A	125	ARG	Mainchain
1	A	126	ALA	Mainchain
1	A	132	ASN	Mainchain
1	A	138	ILE	Mainchain
1	A	141	TRP	Mainchain
1	A	147	THR	Mainchain
1	A	152	ALA	Mainchain
1	A	155	LEU	Mainchain
1	A	159	TYR	Mainchain
1	A	163	VAL	Mainchain
1	A	164	ASP	Mainchain
1	A	165	TYR	Mainchain
1	A	166	ALA	Mainchain
1	A	167	ILE	Mainchain
1	A	168	CYS	Mainchain
1	A	170	SER	Mainchain
1	A	171	TYR	Sidechain,Mainchain
1	A	175	THR	Mainchain
1	A	176	VAL	Mainchain
1	A	178	ASN	Mainchain
1	A	181	VAL	Mainchain
1	A	182	CYS	Mainchain
1	A	189	SER	Mainchain
1	A	191	CYS	Mainchain
1	A	196	GLY	Mainchain
1	A	197	GLY	Mainchain,Peptide
1	A	198	PRO	Mainchain
1	A	199	LEU	Mainchain
1	A	204	ASN	Mainchain
1	A	21	GLU	Mainchain
1	A	217	SER	Mainchain
1	A	221	ASN	Mainchain
1	A	221(A)	VAL	Mainchain
1	A	222	THR	Mainchain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	226	THR	Mainchain
1	A	227	VAL	Mainchain
1	A	229	THR	Mainchain
1	A	23	GLN	Mainchain
1	A	235	ILE	Mainchain
1	A	237	TRP	Mainchain
1	A	238	ILE	Mainchain
1	A	239	ASN	Mainchain
1	A	240	ASN	Mainchain
1	A	241	VAL	Mainchain
1	A	242	ILE	Mainchain
1	A	244	SER	Mainchain
1	A	245	ASN	Sidechain
1	A	33	LEU	Mainchain
1	A	35	TYR	Sidechain
1	A	36	ARG	Mainchain
1	A	36(A)	SER	Mainchain
1	A	40	HIS	Mainchain
1	A	42	CYS	Mainchain
1	A	46	LEU	Mainchain
1	A	49	GLN	Mainchain
1	A	54	THR	Mainchain
1	A	57	HIS	Mainchain
1	A	61	ARG	Mainchain
1	A	64	THR	Mainchain
1	A	65(A)	ARG	Mainchain
1	A	66	VAL	Mainchain
1	A	68	VAL	Mainchain
1	A	69	GLY	Mainchain
1	A	72	ASN	Mainchain
1	A	77	ASN	Sidechain
1	A	80	GLU	Mainchain
1	A	82	TYR	Sidechain
1	A	90	VAL	Mainchain
1	A	91	HIS	Mainchain
1	A	93	TYR	Sidechain
1	A	94	TRP	Mainchain
1	A	95	ASN	Mainchain
1	A	99	VAL	Mainchain
1	A	99(A)	ALA	Mainchain

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	1822	0	1758	22	0
2	A	1	0	0	1	0
3	A	10	0	7	0	0
4	A	25	0	0	6	0
All	All	1858	0	1765	22	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 6.

All (22) 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:75:GLN:O	4:A:272:HOH:O	1.92	0.87
1:A:77:ASN:OD1	4:A:272:HOH:O	2.07	0.72
1:A:230:ARG:NH1	2:A:246:SO4:S	2.63	0.71
1:A:72:ASN:O	4:A:272:HOH:O	2.10	0.70
1:A:31:ILE:HG22	1:A:68:VAL:HG12	1.74	0.69
1:A:48:ARG:HB2	1:A:51:TRP:HB2	1.79	0.64
1:A:183:ALA:HB3	1:A:228:PHE:HE2	1.71	0.55
1:A:212:VAL:HB	1:A:229:THR:HB	1.93	0.51
1:A:68:VAL:CG2	1:A:81:GLN:HB2	2.42	0.50
1:A:87:LYS:HB3	1:A:107:ARG:HB3	1.94	0.49
1:A:102:ASP:HB2	4:A:266:HOH:O	2.12	0.49
1:A:187:GLY:N	4:A:250:HOH:O	2.25	0.48
1:A:31:ILE:CG2	1:A:68:VAL:HG12	2.44	0.47
1:A:56:ALA:HB1	1:A:90:VAL:HG23	1.97	0.47
1:A:31:ILE:HD13	1:A:66:VAL:HG13	1.98	0.45
1:A:143:LEU:HB2	1:A:191:CYS:SG	2.57	0.45
1:A:23:GLN:HB2	1:A:26:SER:OG	2.18	0.44
1:A:114:LEU:HD23	1:A:118:VAL:HG12	2.01	0.42
1:A:91:HIS:HA	1:A:92:PRO:HD2	1.79	0.42
1:A:138:ILE:HG21	1:A:160:LEU:HD22	2.01	0.42
1:A:187:GLY:CA	4:A:250:HOH:O	2.67	0.41
1:A:68:VAL:HG23	1:A:81:GLN:HB2	2.01	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	A	238/240 (99%)	207 (87%)	28 (12%)	3 (1%)	15 26

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	62	GLU
1	A	230	ARG
1	A	24	ARG

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	198/198 (100%)	196 (99%)	2 (1%)	82 95

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	57	HIS
1	A	191	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	95	ASN
1	A	119	GLN
1	A	150	GLN
1	A	153	GLN
1	A	178	ASN
1	A	204	ASN
1	A	206	GLN
1	A	239	ASN
1	A	240	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is modelled with single atom - leaving 1 for Mogul analysis.

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	TSU	A	247	-	9,10,11	1.15	1 (11%)	10,13,16	0.67	0

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
3	TSU	A	247	-	-	0/2/4/6	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	247	TSU	C1-S	-2.39	1.75	1.80

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 (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)

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.