Full wwPDB X-ray Structure Validation Report (i)

Aug 15, 2023 – 12:09 PM EDT

PDB ID	:	1TM3
Title	:	crystal structure of the complex of subtilisin BPN' with chymotrypsin inhibitor 2 M59k mutant
Authors	:	Radisky, E.S.; Kwan, G.; Karen Lu, C.J.; Koshland Jr., D.E.
Deposited on	:	2004-06-10
Resolution	:	1.57 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

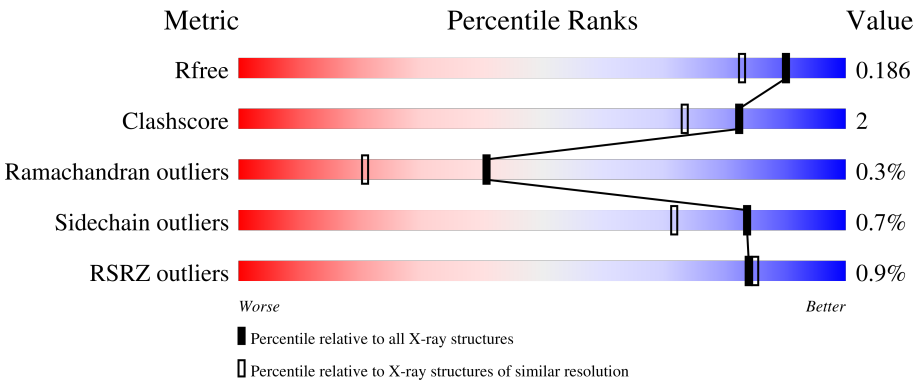
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.57 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	5534 (1.60-1.56)
Clashscore	141614	5861 (1.60-1.56)
Ramachandran outliers	138981	5708 (1.60-1.56)
Sidechain outliers	138945	5703 (1.60-1.56)
RSRZ outliers	127900	5431 (1.60-1.56)

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 >=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% 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	E	281	98% •
2	I	64	3% 89% 8% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	CIT	E	452	-	-	X	-
5	CIT	E	453	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 3071 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 Subtilisin BPN' precursor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	E	281	2030	1252	359	414	5	0	13	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	276	HIS	-	expression tag	UNP P00782
E	277	HIS	-	expression tag	UNP P00782
E	278	HIS	-	expression tag	UNP P00782
E	279	HIS	-	expression tag	UNP P00782
E	280	HIS	-	expression tag	UNP P00782
E	281	HIS	-	expression tag	UNP P00782

- Molecule 2 is a protein called chymotrypsin inhibitor 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	I	64	528	338	94	96	0	6	1

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	20	MET	-	initiating methionine	UNP Q40059
I	59	LYS	MET	engineered mutation	UNP Q40059

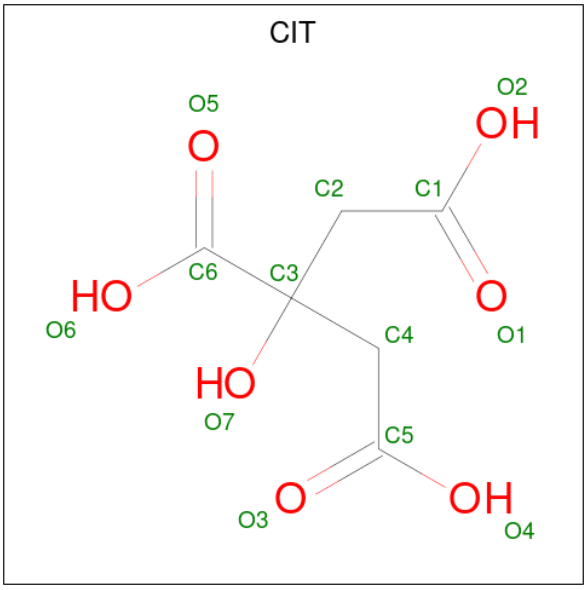
- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
3	E	1	1	1	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

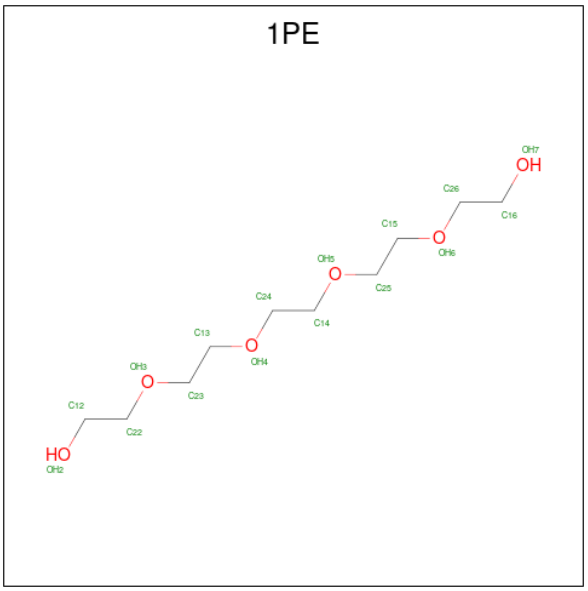
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
4	E	1	1	1	0	0

- Molecule 5 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	E	1	13	6	7	0	0
			Total	C	O		
5	E	1	13	6	7	0	0

- Molecule 6 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
6	E	1	9	5	4	0	1
			Total	C	O		
6	E	1	16	10	6	0	0
			Total	C	O		
6	E	1	13	8	5	0	0

- Molecule 7 is water.

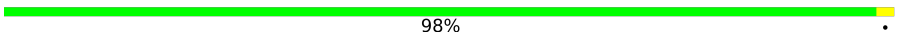
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	E	343	343	343	0	0
			Total	O		
7	I	104	104	104	0	0

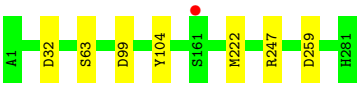
3 Residue-property plots (i)

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: Subtilisin BPN' precursor

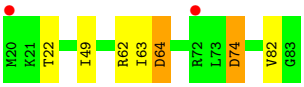
Chain E:





- Molecule 2: chymotrypsin inhibitor 2

Chain I:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	94.20Å 94.20Å 187.62Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	81.65 – 1.57 23.37 – 1.57	Depositor EDS
% Data completeness (in resolution range)	98.9 (81.65-1.57) 98.9 (23.37-1.57)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	3.21 (at 1.57Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.166 , 0.184 0.168 , 0.186	Depositor DCC
R_{free} test set	3452 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	14.6	Xtriage
Anisotropy	0.296	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 44.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	3071	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.31% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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)

Bond lengths and bond angles in the following residue types are not validated in this section: 1PE, CIT, NA, CA

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	0.82	0/2139	0.81	3/2916 (0.1%)
2	I	0.77	0/565	0.84	3/763 (0.4%)
All	All	0.81	0/2704	0.82	6/3679 (0.2%)

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	259	ASP	CB-CG-OD2	6.28	123.95	118.30
2	I	64[A]	ASP	CB-CG-OD2	5.43	123.19	118.30
2	I	64[B]	ASP	CB-CG-OD2	5.43	123.19	118.30
1	E	32	ASP	CB-CG-OD2	5.36	123.12	118.30
2	I	74	ASP	CB-CG-OD2	5.25	123.03	118.30
1	E	247	ARG	NE-CZ-NH2	-5.04	117.78	120.30

There are no chirality outliers.

There are no planarity outliers.

CLOSE-CONTACTS INFOmissingINFO

5.2 Torsion angles (i)

5.2.1 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	293/281 (104%)	283 (97%)	9 (3%)	1 (0%)	41	21
2	I	68/64 (106%)	68 (100%)	0	0	100	100
All	All	361/345 (105%)	351 (97%)	9 (2%)	1 (0%)	41	21

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	63	SER

5.2.2 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	225/211 (107%)	224 (100%)	1 (0%)	91	84
2	I	62/57 (109%)	61 (98%)	1 (2%)	62	39
All	All	287/268 (107%)	285 (99%)	2 (1%)	84	72

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

Mol	Chain	Res	Type
1	E	222	MET
2	I	74	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	245	GLN
1	E	252	ASN
2	I	41	GLN

5.2.3 RNA (i)

There are no RNA molecules in this entry.

5.3 Non-standard residues in protein, DNA, RNA chains (i)

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

5.4 Carbohydrates (i)

There are no monosaccharides in this entry.

5.5 Ligand geometry (i)

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 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	$\#\|Z\| > 2$
6	1PE	E	454[A]	-	7,7,15	0.58	0	6,6,14	0.86	0
6	1PE	E	456	-	12,12,15	0.50	0	11,11,14	0.30	0
6	1PE	E	454[B]	-	7,7,15	0.57	0	6,6,14	0.85	0
5	CIT	E	453	-	12,12,12	3.09	3 (25%)	17,17,17	3.78	11 (64%)
6	1PE	E	455	-	15,15,15	0.49	0	14,14,14	0.27	0
5	CIT	E	452	-	12,12,12	4.86	5 (41%)	17,17,17	4.28	7 (41%)

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
6	1PE	E	454[A]	-	-	1/5/5/13	-
6	1PE	E	456	-	-	4/10/10/13	-
6	1PE	E	454[B]	-	-	1/5/5/13	-
5	CIT	E	453	-	-	3/16/16/16	-
6	1PE	E	455	-	-	6/13/13/13	-
5	CIT	E	452	-	-	0/16/16/16	-

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	452	CIT	C3-C6	-13.74	1.39	1.53
5	E	453	CIT	C3-C6	-9.45	1.43	1.53
5	E	452	CIT	O7-C3	7.95	1.58	1.43
5	E	452	CIT	C4-C3	3.74	1.58	1.53
5	E	453	CIT	O7-C3	2.70	1.48	1.43
5	E	452	CIT	C2-C3	2.32	1.56	1.53
5	E	453	CIT	O5-C6	2.22	1.29	1.22
5	E	452	CIT	O5-C6	2.04	1.28	1.22

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	452	CIT	O7-C3-C6	-12.41	91.44	108.86
5	E	453	CIT	O7-C3-C6	-11.97	92.06	108.86
5	E	452	CIT	O7-C3-C2	6.28	124.08	109.40
5	E	452	CIT	O5-C6-C3	-6.15	113.54	122.25
5	E	452	CIT	O7-C3-C4	5.68	122.68	109.40
5	E	453	CIT	O7-C3-C4	4.70	120.40	109.40
5	E	453	CIT	O5-C6-C3	-4.17	116.34	122.25
5	E	452	CIT	O6-C6-O5	3.70	135.59	123.82
5	E	453	CIT	C4-C3-C6	-3.49	102.62	110.11
5	E	453	CIT	O7-C3-C2	3.02	116.47	109.40
5	E	452	CIT	C4-C3-C2	-2.95	101.45	109.16
5	E	453	CIT	O4-C5-O3	-2.63	116.75	123.30
5	E	453	CIT	O2-C1-C2	2.57	122.60	114.35
5	E	453	CIT	O6-C6-O5	2.37	131.35	123.82
5	E	453	CIT	O2-C1-O1	-2.31	117.55	123.30
5	E	453	CIT	C2-C3-C6	-2.30	105.17	110.11
5	E	452	CIT	O4-C5-C4	2.08	121.04	114.35
5	E	453	CIT	C4-C3-C2	2.04	114.47	109.16

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	453	CIT	C1-C2-C3-C4
6	E	455	1PE	OH6-C15-C25-OH5
6	E	455	1PE	OH5-C14-C24-OH4
5	E	453	CIT	C1-C2-C3-C6
6	E	455	1PE	OH4-C13-C23-OH3
5	E	453	CIT	C1-C2-C3-O7
6	E	456	1PE	OH6-C15-C25-OH5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	E	455	1PE	OH7-C16-C26-OH6
6	E	455	1PE	C16-C26-OH6-C15
6	E	456	1PE	C23-C13-OH4-C24
6	E	456	1PE	C24-C14-OH5-C25
6	E	454[A]	1PE	C15-C25-OH5-C14
6	E	454[B]	1PE	C15-C25-OH5-C14
6	E	456	1PE	C25-C15-OH6-C26
6	E	455	1PE	C24-C14-OH5-C25

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	E	454[A]	1PE	0	1
6	E	454[B]	1PE	0	1
5	E	453	CIT	7	0
5	E	452	CIT	7	0

5.6 Other polymers (i)

There are no such residues in this entry.

5.7 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, 95th 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^2)	Q<0.9
1	E	281/281 (100%)	-0.36	1 (0%)	92	93	8, 13, 24, 33	0
2	I	64/64 (100%)	-0.27	2 (3%)	49	50	12, 16, 28, 38	0
All	All	345/345 (100%)	-0.34	3 (0%)	84	85	8, 14, 25, 38	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	I	20	MET	4.8
1	E	161	SER	3.2
2	I	72	ARG	2.1

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)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	1PE	E	456	13/16	0.66	0.21	55,59,60,61	0
5	CIT	E	453	13/13	0.73	0.24	19,31,34,36	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	1PE	E	455	16/16	0.74	0.28	49,54,61,61	0
6	1PE	E	454[A]	8/16	0.79	0.20	26,27,33,34	1
6	1PE	E	454[B]	8/16	0.79	0.20	26,27,32,34	1
5	CIT	E	452	13/13	0.91	0.10	11,12,17,24	0
4	NA	E	451	1/1	0.98	0.04	15,15,15,15	0
3	CA	E	450	1/1	1.00	0.03	13,13,13,13	0

6.5 Other polymers (i)

There are no such residues in this entry.