Full wwPDB X-ray Structure Validation Report (i)

May 17, 2020 – 03:30 pm BST

PDB ID	:	2SFA
Title	:	SERINE PROTEINASE FROM STREPTOMYCES FRADIAE ATCC 14544
Authors	:	Kitadokoro, K.; Tsuzuki, H.
Deposited on	:	1994-04-25
Resolution	:	1.60 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 (1)) were used in the production of this report:

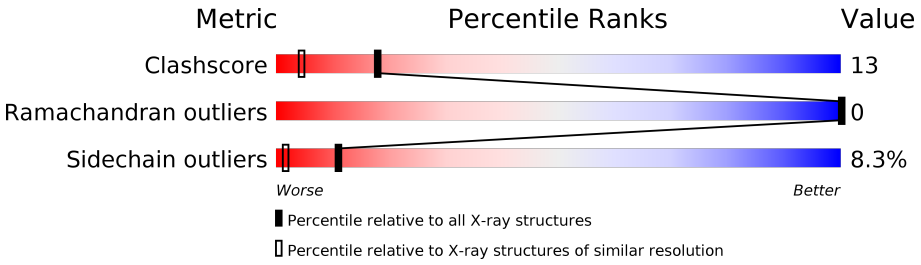
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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	A	191	79% 16%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1423 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 SERINE PROTEINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	191	Total	C	N	O	S	0	0	0
			1337	816	240	277	4			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	86	Total	O	0	0
			86	86		

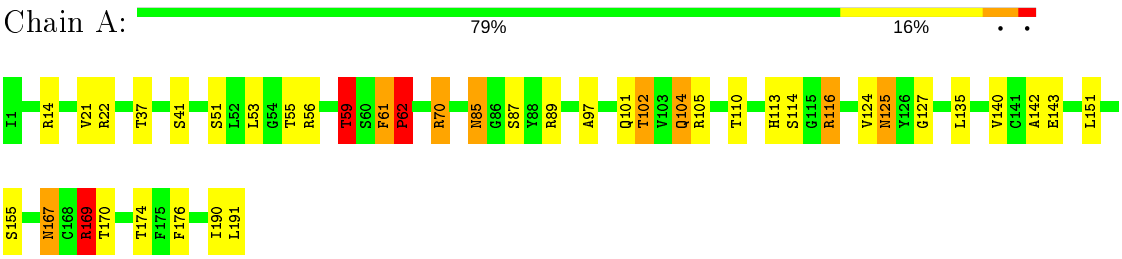
3 Residue-property plots (i)

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: SERINE PROTEINASE

Chain A:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	69.05Å 72.76Å 29.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 1.60	Depositor
% Data completeness (in resolution range)	85.2 (8.00-1.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.182 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1423	wwPDB-VP
Average B, all atoms (Å^2)	12.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	1.48	3/1359 (0.2%)	1.91	25/1844 (1.4%)

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	0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	62	PRO	N-CD	34.42	1.96	1.47
1	A	62	PRO	N-CA	-18.21	1.16	1.47
1	A	62	PRO	CA-C	5.56	1.64	1.52

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	62	PRO	N-CA-CB	30.76	140.21	103.30
1	A	62	PRO	CA-N-CD	-20.31	83.07	111.50
1	A	61	PHE	C-N-CA	13.73	179.66	122.00
1	A	56	ARG	NE-CZ-NH1	12.45	126.53	120.30
1	A	116	ARG	NE-CZ-NH1	11.93	126.26	120.30
1	A	22	ARG	NE-CZ-NH1	11.16	125.88	120.30
1	A	62	PRO	CA-CB-CG	-11.04	83.03	104.00
1	A	61	PHE	C-N-CD	-10.61	97.26	120.60
1	A	169	ARG	NE-CZ-NH1	10.55	125.58	120.30
1	A	105	ARG	NE-CZ-NH1	10.35	125.48	120.30
1	A	70	ARG	NE-CZ-NH1	-8.37	116.11	120.30
1	A	62	PRO	N-CA-C	-8.19	90.80	112.10
1	A	22	ARG	NE-CZ-NH2	-7.37	116.61	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	70	ARG	CD-NE-CZ	-6.87	113.99	123.60
1	A	59	THR	CA-CB-CG2	6.74	121.84	112.40
1	A	116	ARG	NE-CZ-NH2	-6.67	116.97	120.30
1	A	62	PRO	CB-CA-C	6.63	128.58	112.00
1	A	14	ARG	NE-CZ-NH1	-6.22	117.19	120.30
1	A	22	ARG	CD-NE-CZ	6.08	132.11	123.60
1	A	169	ARG	CD-NE-CZ	5.89	131.85	123.60
1	A	70	ARG	NH1-CZ-NH2	5.71	125.68	119.40
1	A	56	ARG	NH1-CZ-NH2	-5.33	113.54	119.40
1	A	104	GLN	O-C-N	5.29	131.16	122.70
1	A	89	ARG	NE-CZ-NH1	5.20	122.90	120.30
1	A	102	THR	N-CA-CB	5.09	119.98	110.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	62	PRO	CA

There are no planarity outliers.

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	1337	0	1266	35	0
2	A	86	0	0	2	5
All	All	1423	0	1266	35	5

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 13.

All (35) 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:62:PRO:CD	1:A:62:PRO:N	1.96	1.21
1:A:167:ASN:HD22	1:A:169:ARG:H	1.01	0.98

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:85:ASN:HD21	1:A:87:SER:HB2	1.35	0.89
1:A:167:ASN:ND2	1:A:169:ARG:H	1.73	0.86
1:A:61:PHE:C	1:A:62:PRO:CD	2.49	0.81
1:A:85:ASN:H	1:A:85:ASN:HD22	1.34	0.74
1:A:167:ASN:HD22	1:A:169:ARG:N	1.83	0.72
1:A:104:GLN:HG2	1:A:114:SER:HB3	1.73	0.71
1:A:70:ARG:HH11	1:A:70:ARG:HG2	1.57	0.68
1:A:70:ARG:HG2	1:A:70:ARG:NH1	2.07	0.67
1:A:125:ASN:HD22	1:A:127:GLY:H	1.46	0.62
1:A:61:PHE:O	1:A:62:PRO:CD	2.49	0.60
1:A:167:ASN:HD21	1:A:169:ARG:HB2	1.70	0.57
1:A:125:ASN:ND2	1:A:127:GLY:H	2.03	0.56
1:A:167:ASN:ND2	1:A:170:THR:H	2.03	0.55
1:A:37:THR:OG1	1:A:59:THR:HB	2.07	0.54
1:A:70:ARG:HH11	1:A:70:ARG:CG	2.15	0.52
1:A:167:ASN:HD21	1:A:170:THR:H	1.58	0.51
1:A:85:ASN:N	1:A:85:ASN:HD22	2.01	0.50
1:A:167:ASN:ND2	1:A:169:ARG:N	2.52	0.49
1:A:124:VAL:HG12	1:A:135:LEU:CD1	2.44	0.48
1:A:167:ASN:ND2	1:A:169:ARG:HB2	2.27	0.48
1:A:116:ARG:NE	2:A:273:HOH:O	2.19	0.48
1:A:97:ALA:HA	1:A:101:GLN:OE1	2.15	0.47
1:A:116:ARG:CD	2:A:273:HOH:O	2.61	0.47
1:A:85:ASN:ND2	1:A:87:SER:H	2.14	0.46
1:A:70:ARG:NH1	1:A:70:ARG:CG	2.68	0.45
1:A:110:THR:O	1:A:113:HIS:HE1	2.00	0.45
1:A:41:SER:O	1:A:55:THR:HA	2.17	0.45
1:A:142:ALA:CB	1:A:174:THR:HG21	2.48	0.43
1:A:21:VAL:HG12	1:A:190:ILE:HD13	2.00	0.43
1:A:110:THR:O	1:A:113:HIS:CE1	2.72	0.42
1:A:143:GLU:OE2	1:A:167:ASN:HB2	2.20	0.42
1:A:151:LEU:HB2	1:A:176:PHE:CD2	2.56	0.41
1:A:140:VAL:O	1:A:174:THR:HG23	2.22	0.40

All (5) 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)
2:A:225:HOH:O	2:A:228:HOH:O[1_554]	0.30	1.90
2:A:248:HOH:O	2:A:276:HOH:O[1_554]	0.31	1.89

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:259:HOH:O	2:A:277:HOH:O[1_554]	0.34	1.86
2:A:211:HOH:O	2:A:263:HOH:O[1_554]	0.88	1.32
2:A:245:HOH:O	2:A:254:HOH:O[2_555]	1.78	0.42

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	189/191 (99%)	185 (98%)	4 (2%)	0	100 100

There are no Ramachandran outliers to report.

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	133/133 (100%)	122 (92%)	11 (8%)	11 2

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51	SER
1	A	53	LEU
1	A	59	THR
1	A	62	PRO
1	A	85	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	102	THR
1	A	125	ASN
1	A	155	SER
1	A	167	ASN
1	A	169	ARG
1	A	191	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	85	ASN
1	A	113	HIS
1	A	125	ASN
1	A	167	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

There are no ligands in this entry.

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.