Full wwPDB X-ray Structure Validation Report (i)

Oct 23, 2021 – 11:58 AM EDT

PDB ID	:	1SUC
Title	:	CALCIUM-INDEPENDENT SUBTILISIN BY DESIGN
Authors	:	Gallagher, T.; Bryan, P.; Gilliland, G.L.
Deposited on	:	1992-06-10
Resolution	:	1.80 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:

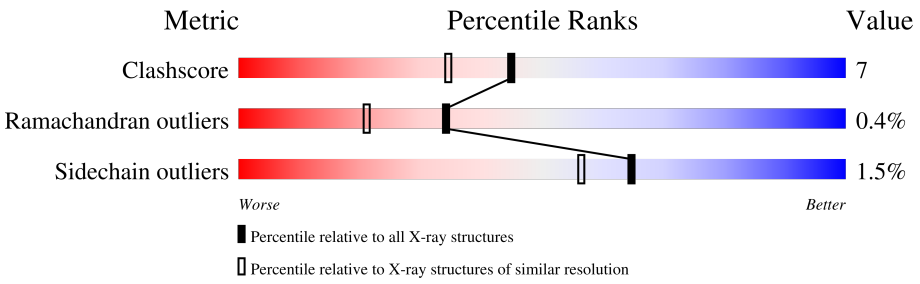
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.80 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	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	275	72% 21% • •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2019 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' CRB-S3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	263	1854	1152	320	377	5	0	0	0

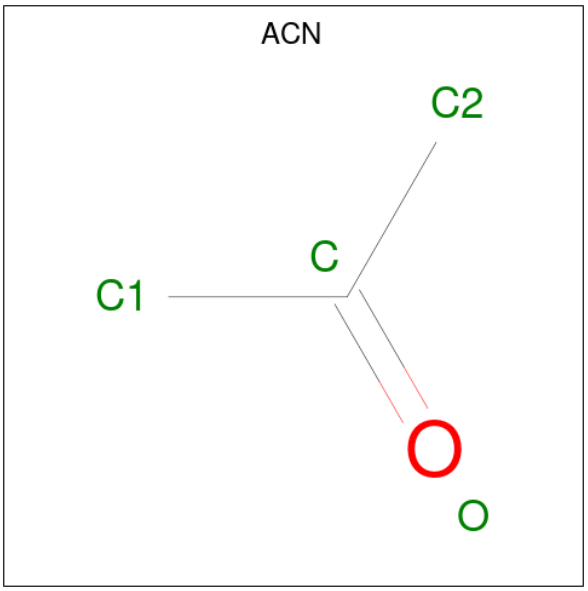
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	50	PHE	MET	engineered mutation	UNP P00782
A	217	LYS	TYR	engineered mutation	UNP P00782
A	218	SER	ASN	engineered mutation	UNP P00782
A	221	CSD	SER	engineered mutation	UNP P00782

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

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

- Molecule 3 is ACETONE (three-letter code: ACN) (formula: C_3H_6O).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	4	3	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	160	160	160	0	0

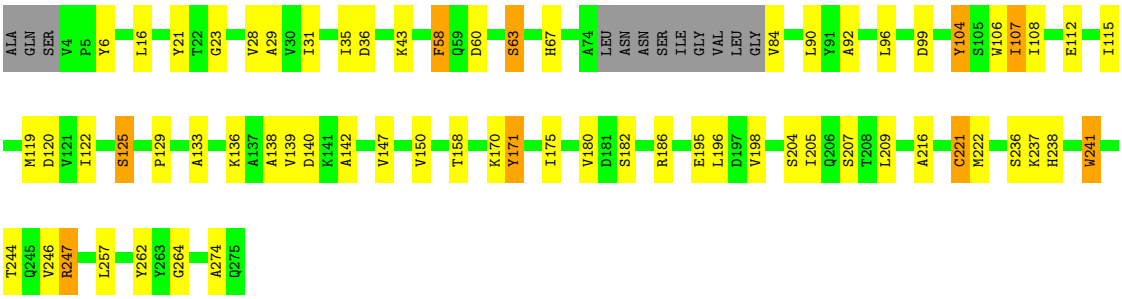
3 Residue-property plots (i)

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. 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: SUBTILISIN BPN' CRB-S3

Chain A: 72% 21%



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	41.37Å 78.67Å 36.78Å 90.00° 114.73° 90.00°	Depositor
Resolution (Å)	8.00 – 1.80	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-1.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROFFT	Depositor
R, R_{free}	0.177 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2019	wwPDB-VP
Average B, all atoms (Å^2)	9.0	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: CSD, K, ACN

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	1.04	1/1884 (0.1%)	1.88	31/2570 (1.2%)

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	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	63	SER	CB-OG	6.22	1.50	1.42

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	247	ARG	NE-CZ-NH1	14.12	127.36	120.30
1	A	247	ARG	NE-CZ-NH2	-11.06	114.77	120.30
1	A	140	ASP	CB-CG-OD1	9.64	126.98	118.30
1	A	63	SER	N-CA-CB	8.79	123.68	110.50
1	A	257	LEU	CB-CA-C	8.61	126.56	110.20
1	A	186	ARG	NE-CZ-NH1	7.92	124.26	120.30
1	A	99	ASP	CB-CG-OD1	7.82	125.34	118.30
1	A	171	TYR	CB-CG-CD1	7.73	125.64	121.00
1	A	150	VAL	CA-CB-CG2	7.27	121.81	110.90
1	A	186	ARG	NE-CZ-NH2	-6.78	116.91	120.30
1	A	28	VAL	CA-CB-CG2	6.74	121.01	110.90
1	A	186	ARG	CD-NE-CZ	6.68	132.95	123.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	244	THR	CA-CB-CG2	6.59	121.63	112.40
1	A	120	ASP	CB-CG-OD2	-6.55	112.40	118.30
1	A	58	PHE	CB-CG-CD2	-6.54	116.22	120.80
1	A	43	LYS	CA-CB-CG	6.45	127.58	113.40
1	A	106	TRP	CA-CB-CG	6.32	125.71	113.70
1	A	140	ASP	CB-CG-OD2	-5.99	112.91	118.30
1	A	90	LEU	CA-CB-CG	5.94	128.97	115.30
1	A	96	LEU	CA-CB-CG	5.94	128.97	115.30
1	A	36	ASP	CB-CG-OD1	5.91	123.62	118.30
1	A	58	PHE	CB-CG-CD1	5.87	124.91	120.80
1	A	60	ASP	CB-CG-OD2	-5.74	113.14	118.30
1	A	139	VAL	CA-CB-CG2	5.74	119.51	110.90
1	A	246	VAL	CA-CB-CG1	5.68	119.42	110.90
1	A	209	LEU	CA-CB-CG	5.52	128.00	115.30
1	A	125	SER	CA-C-O	5.44	131.52	120.10
1	A	171	TYR	CB-CG-CD2	-5.17	117.90	121.00
1	A	180	VAL	CA-CB-CG2	5.17	118.66	110.90
1	A	196	LEU	CA-CB-CG	5.17	127.19	115.30
1	A	241	TRP	CA-CB-CG	5.09	123.36	113.70

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	104	TYR	Mainchain
1	A	112	GLU	Mainchain
1	A	133	ALA	Mainchain
1	A	198	VAL	Mainchain
1	A	264	GLY	Mainchain
1	A	84	VAL	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	1854	0	1808	24	0
2	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	4	0	6	0	0
4	A	160	0	0	2	0
All	All	2019	0	1814	24	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 7.

All (24) 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:31:ILE:HD12	1:A:122:ILE:HG23	1.63	0.79
1:A:6:TYR:HB3	4:A:447:HOH:O	1.96	0.66
1:A:31:ILE:HD12	1:A:122:ILE:CG2	2.27	0.64
1:A:104:TYR:HA	1:A:107:ILE:HD12	1.80	0.63
1:A:115:ILE:HG12	1:A:142:ALA:HA	1.87	0.57
1:A:21:TYR:CZ	1:A:237:LYS:HG3	2.42	0.54
1:A:125:SER:HB3	1:A:221:CSD:HB2	1.89	0.54
1:A:205:ILE:O	1:A:216:ALA:HA	2.08	0.52
1:A:35:ILE:HD12	1:A:92:ALA:HB2	1.91	0.52
1:A:158:THR:HG22	1:A:262:TYR:HE1	1.74	0.51
1:A:170:LYS:HG2	1:A:195:GLU:HG2	1.91	0.51
1:A:129:PRO:HB3	4:A:438:HOH:O	2.10	0.50
1:A:175:ILE:HG12	1:A:247:ARG:HG3	1.95	0.47
1:A:23:GLY:HA2	1:A:236:SER:HB3	1.96	0.46
1:A:67:HIS:CD2	1:A:207:SER:HB3	2.52	0.44
1:A:122:ILE:HD12	1:A:147:VAL:HG11	2.00	0.43
1:A:16:LEU:HD21	1:A:274:ALA:HB3	2.00	0.43
1:A:108:ILE:HG23	1:A:138:ALA:HB2	2.00	0.42
1:A:29:ALA:HB2	1:A:119:MET:HG3	2.02	0.42
1:A:238:HIS:HB3	1:A:241:TRP:CD1	2.56	0.41
1:A:35:ILE:O	1:A:58:PHE:HA	2.20	0.41
1:A:119:MET:O	1:A:147:VAL:HG22	2.21	0.41
1:A:6:TYR:CE1	1:A:182:SER:HB2	2.56	0.40
1:A:136:LYS:HG3	1:A:171:TYR:CD1	2.56	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	260/275 (94%)	251 (96%)	8 (3%)	1 (0%)	34 21

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	63	SER

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	194/204 (95%)	191 (98%)	3 (2%)	65 56

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	107	ILE
1	A	204	SER
1	A	222	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	118	ASN
1	A	275	GLN

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	$\#\|Z\| > 2$
1	CSD	A	221	1	3,7,8	0.78	0	1,8,10	6.52	1 (100%)

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
1	CSD	A	221	1	-	0/2/6/8	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	221	CSD	OD1-SG-CB	-6.52	93.13	105.54

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	221	CSD	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is monoatomic - 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	ACN	A	298	-	3,3,3	0.58	0	3,3,3	0.70	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 (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 is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.