Full wwPDB X-ray Structure Validation Report ⓘ

Oct 23, 2021 – 10:15 AM EDT

PDB ID	:	1SUA
Title	:	SUBTILISIN BPN'
Authors	:	Almog, O.; Gilliland, G.L.
Deposited on	:	1997-01-14
Resolution	:	2.10 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 ⓘ symbol.

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

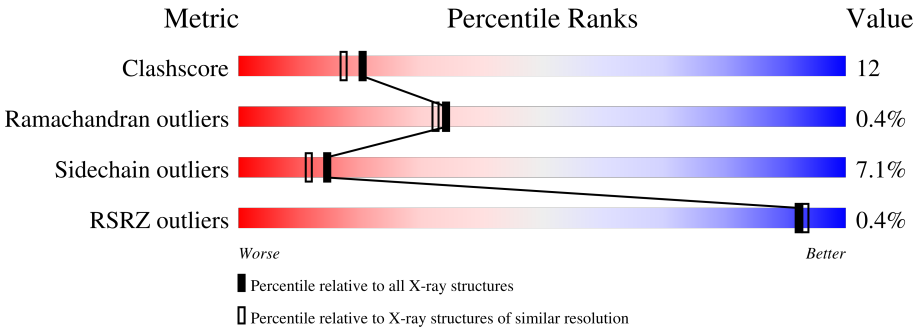
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.10 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	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	A	266	61% 30% 7% •
2	C	4	50% 25% 25%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2018 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'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	263	Total	C	N	O	S	0	1	0
			1852	1153	318	377	4			

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	43	ASN	LYS	engineered mutation	UNP P00782
A	50	PHE	MET	engineered mutation	UNP P00782
A	73	LEU	ALA	engineered mutation	UNP P00782
A	?	-	LEU	deletion	UNP P00782
A	?	-	ASN	deletion	UNP P00782
A	?	-	ASN	deletion	UNP P00782
A	?	-	SER	deletion	UNP P00782
A	?	-	ILE	deletion	UNP P00782
A	?	-	GLY	deletion	UNP P00782
A	?	-	VAL	deletion	UNP P00782
A	?	-	LEU	deletion	UNP P00782
A	?	-	GLY	deletion	UNP P00782
A	206	VAL	GLN	engineered mutation	UNP P00782
A	217	LYS	TYR	engineered mutation	UNP P00782
A	218	SER	ASN	engineered mutation	UNP P00782
A	221	ALA	SER	engineered mutation	UNP P00782
A	271	GLU	GLN	engineered mutation	UNP P00782

- Molecule 2 is a protein called TETRAPEPTIDE ALA-LEU-ALA-LEU.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	4	Total	C	N	O	0	0	0
			27	18	4	5			

- Molecule 3 is water.

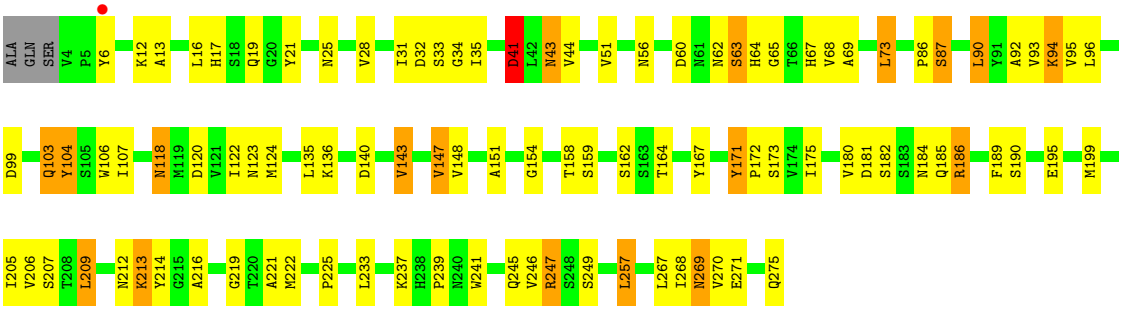
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	137	Total O 137 137	0	0
3	C	2	Total O 2 2	0	0

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'

Chain A:



• Molecule 2: TETRAPEPTIDE ALA-LEU-ALA-LEU

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	53.53Å 60.33Å 83.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.10 7.96 – 2.14	Depositor EDS
% Data completeness (in resolution range)	82.0 (8.00-2.10) 67.3 (7.96-2.14)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	PROLSQ	Depositor
R, R_{free}	0.180 , (Not available) 0.161 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	11.4	Xtriage
Anisotropy	1.096	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 62.5	EDS
L-test for twinning[1]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2018	wwPDB-VP
Average B, all atoms (Å^2)	10.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.05% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]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

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	1.12	1/1895 (0.1%)	1.97	57/2587 (2.2%)
2	C	1.01	0/26	2.33	1/33 (3.0%)
All	All	1.12	1/1921 (0.1%)	1.97	58/2620 (2.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	87	SER	CB-OG	-5.03	1.35	1.42

All (58) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	247	ARG	NE-CZ-NH1	12.78	126.69	120.30
1	A	186	ARG	NE-CZ-NH1	11.42	126.01	120.30
1	A	99	ASP	CB-CG-OD1	9.68	127.01	118.30
1	A	180	VAL	CA-CB-CG2	9.12	124.59	110.90
1	A	247	ARG	NE-CZ-NH2	-8.75	115.92	120.30
1	A	140	ASP	CB-CG-OD1	8.74	126.17	118.30
1	A	257	LEU	CB-CA-C	8.71	126.74	110.20
1	A	186	ARG	CD-NE-CZ	8.57	135.60	123.60
1	A	41	ASP	CA-CB-CG	8.07	131.15	113.40
1	A	104	TYR	CB-CG-CD1	8.02	125.81	121.00
1	A	41	ASP	CB-CG-OD1	7.92	125.43	118.30
1	A	239	PRO	C-N-CA	7.84	141.30	121.70
1	A	28	VAL	CA-CB-CG2	7.75	122.52	110.90
1	A	62	ASN	C-N-CA	7.67	140.86	121.70
2	C	677	LEU	CB-CA-C	7.54	124.52	110.20
1	A	181	ASP	CB-CG-OD1	7.45	125.00	118.30
1	A	43	ASN	CA-CB-CG	6.80	128.36	113.40
1	A	120	ASP	CB-CG-OD1	6.64	124.28	118.30
1	A	209	LEU	CA-CB-CG	6.59	130.45	115.30
1	A	103	GLN	CB-CG-CD	6.24	127.83	111.60
1	A	106	TRP	CA-CB-CG	6.19	125.45	113.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	63	SER	N-CA-CB	6.15	119.73	110.50
1	A	60	ASP	CB-CA-C	6.07	122.53	110.40
1	A	212	ASN	CB-CA-C	6.06	122.51	110.40
1	A	164	THR	C-N-CA	5.99	136.66	121.70
1	A	44	VAL	CA-CB-CG2	5.94	119.81	110.90
1	A	185	GLN	CB-CG-CD	5.92	126.98	111.60
1	A	96	LEU	CA-CB-CG	5.87	128.80	115.30
1	A	186	ARG	NE-CZ-NH2	-5.77	117.42	120.30
1	A	25	ASN	CB-CA-C	5.77	121.93	110.40
1	A	158	THR	CA-CB-CG2	5.76	120.46	112.40
1	A	233	LEU	CA-CB-CG	5.66	128.32	115.30
1	A	28	VAL	CA-CB-CG1	5.63	119.35	110.90
1	A	103	GLN	CA-CB-CG	5.61	125.74	113.40
1	A	16	LEU	CA-CB-CG	5.60	128.19	115.30
1	A	148	VAL	CA-CB-CG2	5.60	119.30	110.90
1	A	199	MET	CA-CB-CG	5.54	122.71	113.30
1	A	13	ALA	N-CA-CB	5.49	117.79	110.10
1	A	73	LEU	CB-CA-C	5.48	120.62	110.20
1	A	270	VAL	CB-CA-C	5.41	121.67	111.40
1	A	73	LEU	CA-CB-CG	5.40	127.72	115.30
1	A	56	ASN	CB-CA-C	5.33	121.06	110.40
1	A	118	ASN	CA-CB-CG	5.29	125.03	113.40
1	A	17	HIS	CA-CB-CG	5.28	122.58	113.60
1	A	41	ASP	N-CA-CB	5.27	120.09	110.60
1	A	185	GLN	CA-CB-CG	5.24	124.94	113.40
1	A	246	VAL	CA-CB-CG1	5.20	118.71	110.90
1	A	271	GLU	CA-CB-CG	5.14	124.72	113.40
1	A	269	ASN	N-CA-CB	5.11	119.80	110.60
1	A	147	VAL	CA-CB-CG2	5.10	118.55	110.90
1	A	123	ASN	N-CA-CB	5.10	119.77	110.60
1	A	195	GLU	CG-CD-OE1	5.09	128.49	118.30
1	A	99	ASP	CB-CG-OD2	-5.09	113.72	118.30
1	A	241	TRP	CA-CB-CG	5.06	123.31	113.70
1	A	32	ASP	CB-CG-OD1	5.05	122.84	118.30
1	A	51	VAL	CA-CB-CG2	5.02	118.43	110.90
1	A	171	TYR	CA-CB-CG	5.01	122.92	113.40
1	A	143	VAL	CA-CB-CG2	5.01	118.41	110.90

There are no chirality outliers.

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	1852	0	1805	46	0
2	C	27	0	31	6	0
3	A	137	0	0	5	0
3	C	2	0	0	0	0
All	All	2018	0	1836	46	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 12.

All (46) 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:103:GLN:HG3	3:A:409:HOH:O	1.70	0.92
1:A:175:ILE:HD11	1:A:247:ARG:HG3	1.56	0.88
1:A:68:VAL:HG13	1:A:225:PRO:HG3	1.66	0.77
1:A:69:ALA:HB1	1:A:90:LEU:HD21	1.72	0.72
1:A:122:ILE:HD12	1:A:147:VAL:HG11	1.71	0.71
1:A:33:SER:O	1:A:94:LYS:HE3	1.91	0.70
1:A:206:VAL:HG13	1:A:216:ALA:HB2	1.83	0.61
1:A:34:GLY:O	1:A:65:GLY:HA3	2.03	0.59
1:A:189:PHE:CE1	1:A:219:GLY:HA2	2.38	0.59
1:A:143:VAL:HG21	1:A:173:SER:HB2	1.87	0.56
1:A:35:ILE:HD12	1:A:92:ALA:HB2	1.87	0.56
1:A:86:PRO:HB3	3:A:422:HOH:O	2.05	0.56
1:A:175:ILE:CD1	1:A:247:ARG:HG3	2.32	0.56
1:A:257:LEU:HD11	1:A:267:LEU:HB2	1.87	0.56
1:A:31:ILE:HD12	1:A:122:ILE:HG23	1.89	0.55
1:A:122:ILE:CD1	1:A:147:VAL:HG11	2.36	0.55
1:A:107:ILE:CD1	2:C:674:ALA:HB3	2.38	0.53
1:A:41:ASP:OD2	1:A:214:TYR:OH	2.23	0.53
1:A:221:ALA:HB2	2:C:677:LEU:C	2.30	0.52
1:A:12:LYS:HD2	1:A:269:ASN:ND2	2.24	0.52
1:A:206:VAL:CG1	1:A:216:ALA:HB2	2.38	0.52
1:A:21:TYR:CZ	1:A:237:LYS:HG3	2.46	0.50
1:A:73:LEU:HG	1:A:90:LEU:HD13	1.93	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:31:ILE:CG2	1:A:95:VAL:HG21	2.42	0.49
1:A:206:VAL:HG22	3:A:369:HOH:O	2.12	0.49
1:A:159:SER:O	1:A:162:SER:HB2	2.12	0.49
1:A:167:TYR:HD2	1:A:171:TYR:CZ	2.31	0.49
1:A:107:ILE:HD11	2:C:674:ALA:HB3	1.95	0.48
1:A:209:LEU:HB2	1:A:213:LYS:HB3	1.96	0.46
1:A:171:TYR:HA	1:A:172:PRO:HD3	1.81	0.46
1:A:124:MET:O	1:A:151:ALA:HA	2.16	0.46
1:A:154:GLY:HA2	2:C:677:LEU:HD13	1.97	0.46
1:A:245:GLN:HG3	3:A:348:HOH:O	2.15	0.46
1:A:31:ILE:HG23	1:A:95:VAL:HG21	1.97	0.45
1:A:205:ILE:O	1:A:216:ALA:HA	2.16	0.45
1:A:104:TYR:CE1	2:C:674:ALA:HB2	2.52	0.44
1:A:67:HIS:CD2	1:A:207:SER:HB3	2.54	0.43
1:A:31:ILE:HG12	1:A:93:VAL:CG1	2.48	0.42
1:A:21:TYR:CE2	1:A:237:LYS:HG3	2.53	0.42
1:A:6:TYR:CE2	1:A:182:SER:HB2	2.55	0.42
1:A:136:LYS:HG2	3:A:337:HOH:O	2.19	0.42
1:A:64:HIS:O	1:A:68:VAL:HG23	2.20	0.41
1:A:104:TYR:CE2	1:A:135:LEU:HD22	2.55	0.41
1:A:184:ASN:HB3	1:A:257:LEU:HD22	2.03	0.41
1:A:104:TYR:HE1	2:C:674:ALA:HB2	1.85	0.41
1:A:31:ILE:HG12	1:A:93:VAL:HG11	2.03	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	262/266 (98%)	249 (95%)	12 (5%)	1 (0%)	34 32
2	C	2/4 (50%)	2 (100%)	0	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	264/270 (98%)	251 (95%)	12 (4%)	1 (0%)	34 32

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	196/198 (99%)	182 (93%)	14 (7%)	14 11
2	C	2/2 (100%)	2 (100%)	0	100 100
All	All	198/200 (99%)	184 (93%)	14 (7%)	14 11

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	19	GLN
1	A	41	ASP
1	A	43	ASN
1	A	87	SER
1	A	90	LEU
1	A	94	LYS
1	A	118	ASN
1	A	186	ARG
1	A	190	SER
1	A	213	LYS
1	A	222	MET
1	A	249	SER
1	A	268	ILE
1	A	275	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	117	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 monosaccharides 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

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, 95^{th} 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	A	263/266 (98%)	-0.46	1 (0%) 92 93	2, 9, 19, 29	0
2	C	4/4 (100%)	0.17	0 100 100	12, 13, 15, 16	4 (100%)
All	All	267/270 (98%)	-0.46	1 (0%) 92 93	2, 9, 19, 29	4 (1%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	6	TYR	2.7

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.