wwPDB X-ray Structure Validation Summary Report (i)

Sep 24, 2023 – 06:55 PM EDT

PDB ID	:	5VAF
Title	:	Crystal structure of accessory secretion protein 1
Authors	:	Chen, Y.; Rapoport, T.A.; Jeffrey, P.D.
Deposited on	:	2017-03-25
Resolution	:	2.77 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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:

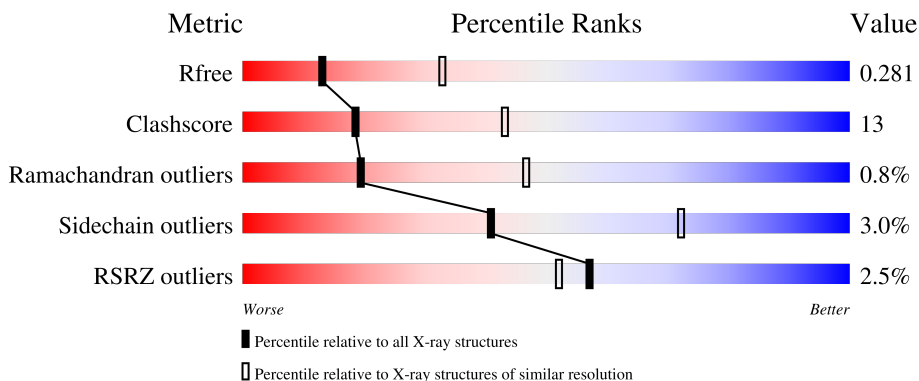|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1   Overall quality at a glance ⓘ

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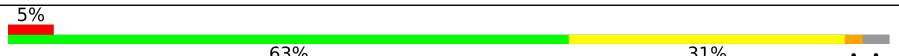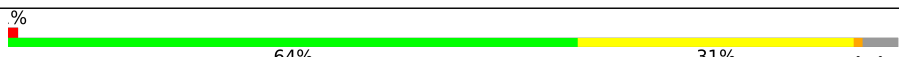*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.77 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	Percentile Ranks	Value
Rfree		0.281
Clashscore		13
Ramachandran outliers		0.8%
Sidechain outliers		3.0%
RSRZ outliers		2.5%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)
RSRZ outliers	127900	4027 (2.80-2.76)

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	533	68% 26% · ·
1	B	533	4% 68% 26% · ·
1	C	533	5% 63% 31% · ·
1	D	533	.% 64% 31% · ·

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 17260 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 Accessory Sec system protein Asp1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	510	Total	C	N	O	S	0	0	0
			4292	2776	705	802	9			
1	B	515	Total	C	N	O	S	0	0	0
			4329	2802	710	808	9			
1	C	516	Total	C	N	O	S	0	0	0
			4338	2807	712	810	9			
1	D	511	Total	C	N	O	S	0	0	0
			4298	2779	706	804	9			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	527	VAL	-	expression tag	UNP Q9AET9
A	528	ASP	-	expression tag	UNP Q9AET9
A	529	LYS	-	expression tag	UNP Q9AET9
A	530	LEU	-	expression tag	UNP Q9AET9
A	531	VAL	-	expression tag	UNP Q9AET9
A	532	PRO	-	expression tag	UNP Q9AET9
A	533	ARG	-	expression tag	UNP Q9AET9
B	527	VAL	-	expression tag	UNP Q9AET9
B	528	ASP	-	expression tag	UNP Q9AET9
B	529	LYS	-	expression tag	UNP Q9AET9
B	530	LEU	-	expression tag	UNP Q9AET9
B	531	VAL	-	expression tag	UNP Q9AET9
B	532	PRO	-	expression tag	UNP Q9AET9
B	533	ARG	-	expression tag	UNP Q9AET9
C	527	VAL	-	expression tag	UNP Q9AET9
C	528	ASP	-	expression tag	UNP Q9AET9
C	529	LYS	-	expression tag	UNP Q9AET9
C	530	LEU	-	expression tag	UNP Q9AET9
C	531	VAL	-	expression tag	UNP Q9AET9
C	532	PRO	-	expression tag	UNP Q9AET9
C	533	ARG	-	expression tag	UNP Q9AET9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	527	VAL	-	expression tag	UNP Q9AET9
D	528	ASP	-	expression tag	UNP Q9AET9
D	529	LYS	-	expression tag	UNP Q9AET9
D	530	LEU	-	expression tag	UNP Q9AET9
D	531	VAL	-	expression tag	UNP Q9AET9
D	532	PRO	-	expression tag	UNP Q9AET9
D	533	ARG	-	expression tag	UNP Q9AET9

- Molecule 2 is water.

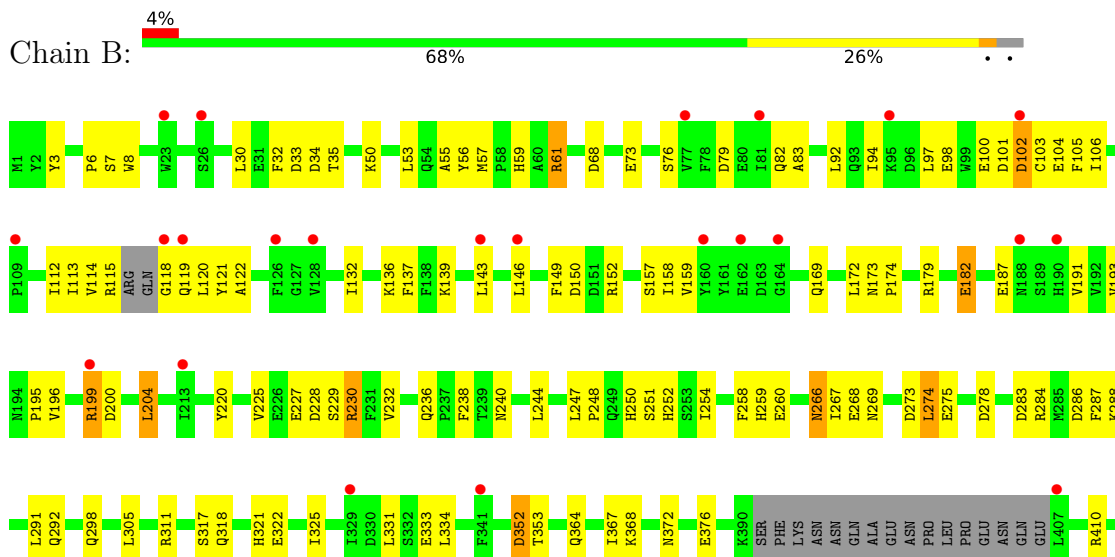| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total 2	O 2	0	0
2	B	1	Total 1	O 1	0	0

3   Residue-property plots ⓘ

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: Accessory Sec system protein Asp1

• Molecule 1: Accessory Sec system protein Asp1

N415 E419 L422 D428 T429 R430 L431 I432 L433 D434 Q438 P439 Y442 I448 G451 I452 P453 Q454 A459 V463 Q467 I471 L472 D473 Q477 L496 I497 Y498 K502 W518 K526 V527 D528 K529 L530 V531 P532 R533

● Molecule 1: Accessory Sec system protein Asp1

Chain C:

5%
63%   31%

M1 Y2 K13 R14 V15 W16 H17 I20 I21 P22 W23 Y24 L30 I39 H43 N46 K50 L51 L52 L53 Y56 M57 A60 R61 Y62 F63 R66 Q67 D68 E71 T72 E73 S76 V77 F78 D79 E80 I81 Q82 Q90 V91 L92 K95 D96 L97 E98

W99 E100 D101 D102 C103 E104 F105 I106 F110 L111 I112 R115 R116 Q117 G118 Q119 L120 Y121 A122 H123 V124 E125 F126 G127 V128 S133 F134 D140 D141 Q142 L146 N147 D151 R152 G153 I158 V159 E162 V167 C168 Q169 D170 Y171 L172 I180 R181 E182 Y183 L184 K185

F186 E187 N188 S189 H190 V196 F197 S198 R199 Y206 E207 C208 M209 L212 I213 K216 L217 G218 S222 D228 S229 R230 F231 Y232 A235 Q236 P237 F238 L246 L247 P248 Q249 H250 S251 I254 L255 S256 F257 N262 N266 I267 E268 N269 L270 K271 E275 R284

M285 D286 F287 K288 L291 Q292 N293 Y294 F295 A299 E300 K301 Y304 L305 S306 R311 G315 Q318 Q319 R320 H321 K324 Y327 Q328 I329 D330 L331 S332 E333 N336 A339 I340 V343 A348 V356 W363 I367 K368 E371 V374 I378

K385 D386 F387 I388 K389 K390 SER PHE LYS ASN ASN GLN ALA GLU ASN PRO LEU PRO GLU ASN GLN GLU L407 R410 N415 D418 E419 L420 D427 D428 T429 R430 L431 I432 I433 D434 Q438 L441 Q444 I448 Q454 V463 T464 H465 L466 Q467 I471 I475

L478 D483 L486 Q487 Y498 I503 K513 R514 W515 W518 D528 P532 ARG

● Molecule 1: Accessory Sec system protein Asp1

Chain D:

%
64%   31%

M1 Y2 Y3 P6 S7 W8 S9 G10 H17 R18 F32 T35 I39 R40 I41 F42 E45 P48 V49 K50 L51 L52 L53 M57 A60 R61 I69 Y74 Y75 S76 V77 V84 M89 L92 Q93 I94 K95 D96 L97 E98 W99 E100 D101 D102 I106

Y107 T108 P109 F110 L111 I112 I113 V114 R115 R116 Q117 A122 H123 V124 I132 S133 F134 F149 D150 D151 R152 G153 F154 V155 Y161 E162 Q165 E166 V167 Y171 P174 N175 I180 K185 F186 E187 N188 S189 H190 V191 D200 Y206 M209 L212 I213 K216

V225 E226 E227 D228 S229 R230 F238 T239 N240 F258 H259 E260 S265 N266 I267 E268 N269 L270 D273 L274 R284 M285 K288 L291 Q298 K301 I302 T310 R311 L312 Q313 S317 R320 F326 D330 L331 S332 E333 F341 F345 Y346 V347

P351 L355 E365 G366 I367 K368 Q369 L377 L382 D383 D386 K389 K390 S391 PHE LYS ASN ASN GLN ALA GLU ASN PRO LEU PRO GLU ASN GLN GLU L407 E408 Y409 R410 F411 R412 I413 R414 E419 L420 S421 L422 I423 Q424 E425 D428 T429 I432 L435 L441

Y442 I445 I448 I452 S460 Q467 D473 S474 I475 S476 L486 Q487 G488 L489 K490 Y498 S499 K502 I503 T507 G508 H509 Q510 R514 W515 E516 K517 W518 L519 K520 I413 A522 I523 D524 E525 K526 VAL ASP LYS LEU VAL PRO ARG

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	41.50Å 99.90Å 179.07Å 100.66° 90.07° 95.78°	Depositor
Resolution (Å)	97.65 – 2.77 97.65 – 2.77	Depositor EDS
% Data completeness (in resolution range)	92.7 (97.65-2.77) 92.7 (97.65-2.77)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.77 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.208 , 0.282 0.208 , 0.281	Depositor DCC
R_{free} test set	3344 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	46.4	Xtriage
Anisotropy	0.613	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 46.5	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	17260	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.38 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.7622e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 ⓘ

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	0.48	0/4400	0.62	0/5962
1	B	0.48	0/4437	0.61	1/6012 (0.0%)
1	C	0.51	0/4447	0.64	2/6027 (0.0%)
1	D	0.51	0/4406	0.64	3/5970 (0.1%)
All	All	0.49	0/17690	0.63	6/23971 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	141	ASP	CB-CG-OD1	6.77	124.39	118.30
1	D	107	TYR	C-N-CA	6.70	138.44	121.70
1	D	97	LEU	N-CA-C	6.45	128.41	111.00
1	D	312	LEU	CA-CB-CG	5.47	127.88	115.30
1	C	528	ASP	CB-CG-OD2	5.34	123.11	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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	4292	0	4164	101	0
1	B	4329	0	4208	110	1
1	C	4338	0	4217	126	1
1	D	4298	0	4169	116	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	2	0	0	1	0
2	B	1	0	0	0	0
All	All	17260	0	16758	448	1

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.

The worst 5 of 448 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:C:236:GLN:HG3	1:C:237:PRO:HD2	1.49	0.94
1:D:414:ARG:NH1	1:D:425:GLU:OE1	2.09	0.85
1:A:108:THR:HG22	1:A:110:PHE:H	1.43	0.83
1:C:257:PHE:HD1	1:C:262:ASN:HD22	1.27	0.82
1:B:100:GLU:HB2	1:B:103:CYS:SG	2.19	0.82

All (1) 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)
1:B:473:ASP:OD2	1:C:390:LYS:NZ[1_445]	2.15	0.05

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	506/533 (95%)	462 (91%)	40 (8%)	4 (1%)	19	47
1	B	509/533 (96%)	461 (91%)	44 (9%)	4 (1%)	19	47
1	C	512/533 (96%)	459 (90%)	47 (9%)	6 (1%)	13	36
1	D	507/533 (95%)	465 (92%)	39 (8%)	3 (1%)	25	54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	2034/2132 (95%)	1847 (91%)	170 (8%)	17 (1%)	19	47

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	187	GLU
1	D	100	GLU
1	A	266	ASN
1	B	352	ASP
1	C	319	GLN

5.3.2 Protein sidechains ⓘ

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	471/493 (96%)	461 (98%)	10 (2%)	53	81
1	B	476/493 (97%)	458 (96%)	18 (4%)	33	64
1	C	477/493 (97%)	458 (96%)	19 (4%)	31	62
1	D	472/493 (96%)	462 (98%)	10 (2%)	53	81
All	All	1896/1972 (96%)	1839 (97%)	57 (3%)	41	72

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	68	ASP
1	D	421	SER
1	C	170	ASP
1	D	317	SER
1	D	101	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	123	HIS
1	C	169	GLN
1	D	444	GLN
1	D	28	GLN
1	A	456	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, 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	A	510/533 (95%)	0.16	2 (0%)	92	92	25, 44, 71, 93	0
1	B	515/533 (96%)	0.33	23 (4%)	33	27	29, 49, 88, 118	0
1	C	516/533 (96%)	0.28	24 (4%)	31	25	32, 49, 89, 116	0
1	D	511/533 (95%)	0.19	3 (0%)	89	88	25, 45, 76, 105	0
All	All	2052/2132 (96%)	0.24	52 (2%)	57	52	25, 46, 84, 118	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	118	GLY	9.8
1	C	162	GLU	4.6
1	B	143	LEU	4.5
1	D	407	LEU	3.9
1	C	103	CYS	3.8

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 (i)

There are no such residues in this entry.