wwPDB X-ray Structure Validation Summary Report (i)

Feb 19, 2016 – 09:50 PM GMT

PDB ID	:	5BV0
Title	:	Crystal Structure of a Complex Between the SNARE Nyv1 and the HOPS Vps33-Vps16 subcomplex from Chaetomium thermophilum
Authors	:	Baker, R.W.; Jeffrey, P.D.; Hughson, F.M.
Deposited on	:	2015-06-04
Resolution	:	3.10 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
http://wwpdb.org/validation/2016/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:

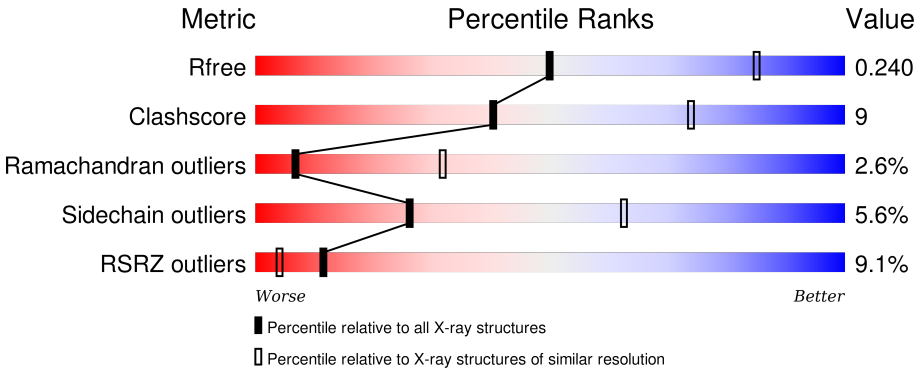
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20026982

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 3.10 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))
R_{free}	91344	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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. 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	669	9%, 68%, 19%, •, 10%
2	B	229	%, 69%, 18%, •, 10%
3	C	45	29%, 58%, 20%, 22%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6659 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 SM (Sec1/Munc18-like) protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	604	Total	C	N	O	S	0	0	0
			4771	3021	841	898	11			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	EXPRESSION TAG	UNP G0SCM5

- Molecule 2 is a protein called Vps16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	207	Total	C	N	O	S	0	0	0
			1657	1046	294	312	5			

- Molecule 3 is a protein called SNARE domain.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	35	Total	C	N	O	0	0	0
			231	141	45	45			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	165	UNK	-	see remark 999	UNP G0S5G3
C	166	UNK	-	see remark 999	UNP G0S5G3
C	167	UNK	-	see remark 999	UNP G0S5G3
C	168	UNK	-	see remark 999	UNP G0S5G3
C	169	UNK	-	see remark 999	UNP G0S5G3
C	170	UNK	-	see remark 999	UNP G0S5G3
C	171	UNK	-	see remark 999	UNP G0S5G3
C	172	UNK	-	see remark 999	UNP G0S5G3
C	173	UNK	-	see remark 999	UNP G0S5G3

Continued on next page...

Continued from previous page...

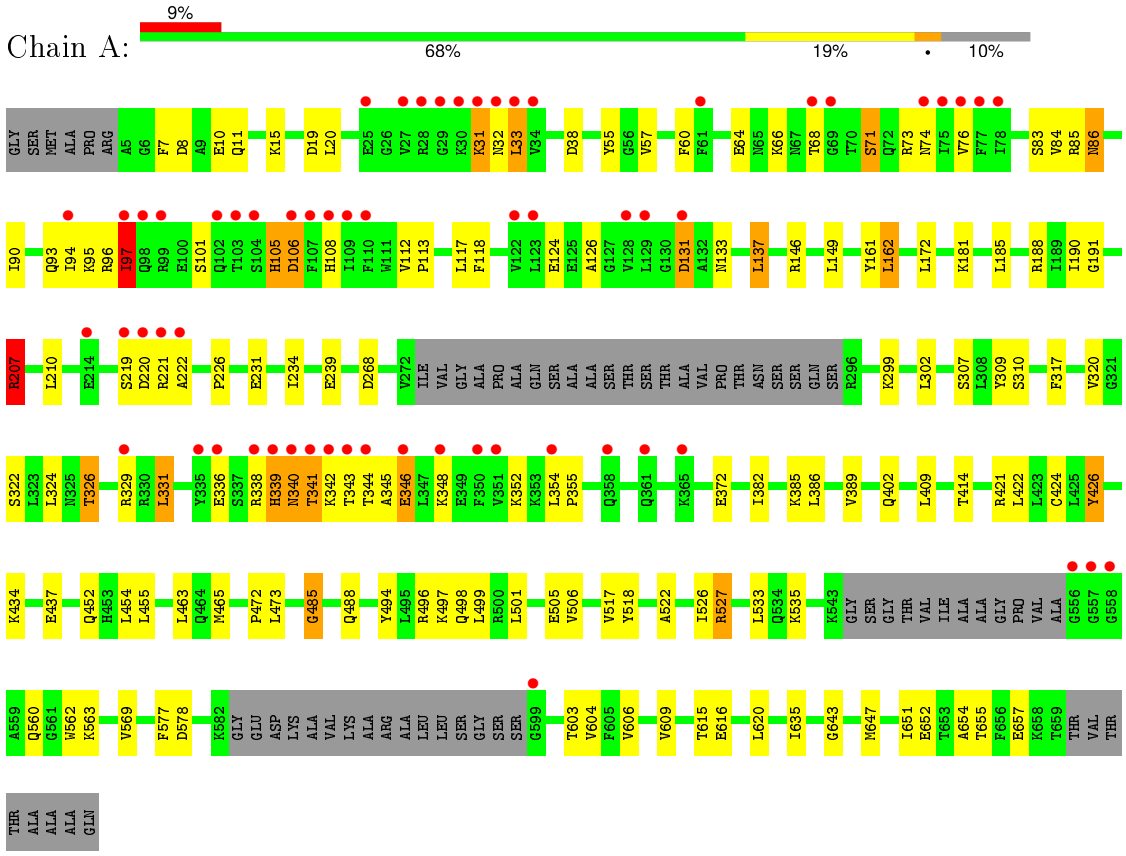
Chain	Residue	Modelled	Actual	Comment	Reference
C	174	UNK	-	see remark 999	UNP G0S5G3
C	175	UNK	-	see remark 999	UNP G0S5G3
C	176	UNK	-	see remark 999	UNP G0S5G3
C	177	UNK	-	see remark 999	UNP G0S5G3
C	178	UNK	-	see remark 999	UNP G0S5G3
C	179	UNK	-	see remark 999	UNP G0S5G3

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 errors 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.

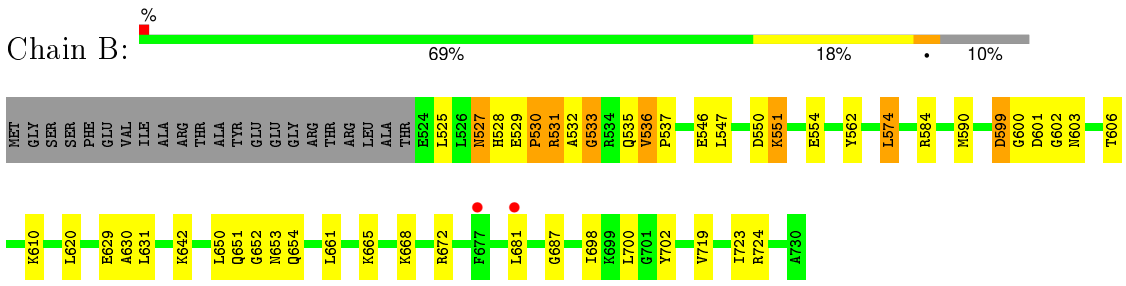
- Molecule 1: SM (Sec1/Munc18-like) protein

Chain A:



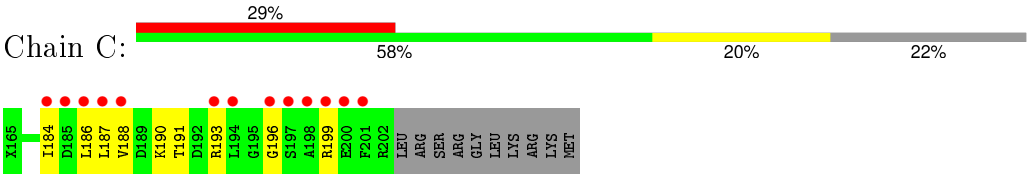
- Molecule 2: Vps16

Chain B:



- Molecule 3: SNARE domain

Chain C:


29%
58%
20%
22%
X165
I184
D185
L186
L187
V188
D189
K190
T191
D192
R193
L194
G195
G196
S197
A198
R199
E200
F201
R202
LEU
ARG
SER
ARG
GLY
LEU
LYS
ARG
LYS
MET

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	101.03Å 258.94Å 75.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.47 – 3.10 49.47 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.47-3.10) 94.1 (49.47-3.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.10 (at 3.19Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.195 , 0.246 0.192 , 0.240	Depositor DCC
R_{free} test set	1610 reflections (5.05%)	DCC
Wilson B-factor (Å^2)	80.6	Xtriage
Anisotropy	0.757	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 82.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Outliers	2 of 33343 reflections (0.006%)	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6659	wwPDB-VP
Average B, all atoms (Å^2)	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.29	0/4846	0.53	1/6536 (0.0%)
2	B	0.31	0/1683	0.65	1/2268 (0.0%)
3	C	0.22	0/156	0.59	0/207
All	All	0.29	0/6685	0.56	2/9011 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	574	LEU	CA-CB-CG	8.09	133.91	115.30
1	A	207	ARG	NE-CZ-NH1	6.52	123.56	120.30

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	4771	0	4829	85	0
2	B	1657	0	1673	33	0
3	C	231	0	173	5	0
All	All	6659	0	6675	116	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 9.

The worst 5 of 116 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)
2:B:599:ASP:H	2:B:600:GLY:HA2	1.39	0.87
1:A:434:LYS:H	1:A:434:LYS:HD2	1.48	0.78
1:A:322:SER:O	1:A:326:THR:OG1	2.08	0.70
1:A:207:ARG:HH11	1:A:207:ARG:HG2	1.56	0.69
1:A:654:ALA:HB1	2:B:631:LEU:HB3	1.75	0.68

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	596/669 (89%)	545 (91%)	39 (6%)	12 (2%)	9	38
2	B	205/229 (90%)	183 (89%)	14 (7%)	8 (4%)	4	22
3	C	18/45 (40%)	12 (67%)	5 (28%)	1 (6%)	2	13
All	All	819/943 (87%)	740 (90%)	58 (7%)	21 (3%)	7	32

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	97	ILE
1	A	340	ASN
1	A	344	THR
1	A	506	VAL
2	B	530	PRO

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	517/564 (92%)	489 (95%)	28 (5%)	27 64
2	B	177/194 (91%)	169 (96%)	8 (4%)	34 70
3	C	16/26 (62%)	12 (75%)	4 (25%)	1 2
All	All	710/784 (91%)	670 (94%)	40 (6%)	26 62

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	339	HIS
1	A	454	LEU
3	C	184	ILE
1	A	426	TYR
1	A	527	ARG

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

Mol	Chain	Res	Type
1	A	74	ASN
1	A	133	ASN
1	A	498	GLN

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	179:UNK	C	183:ARG	N	7.51

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	604/669 (90%)	0.42	61 (10%) 9 3	46, 100, 224, 311	0
2	B	207/229 (90%)	-0.02	2 (0%) 84 69	63, 94, 177, 264	0
3	C	20/45 (44%)	2.53	13 (65%) 0 0	138, 207, 256, 263	0
All	All	831/943 (88%)	0.36	76 (9%) 11 4	46, 98, 224, 311	0

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	75	ILE	6.1
1	A	557	GLY	5.6
1	A	335	TYR	5.1
3	C	199	ARG	4.9
1	A	220	ASP	4.8

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 carbohydrates 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.