wwPDB X-ray Structure Validation Summary Report (i)

Aug 22, 2020 – 11:59 AM BST

PDB ID	:	5VWV
Title	:	Bak core latch dimer in complex with Bim-BH3 - Cubic
Authors	:	Brouwer, J.M.; Colman, P.M.; Czabotar, P.E.
Deposited on	:	2017-05-23
Resolution	:	1.90 Å(reported)

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

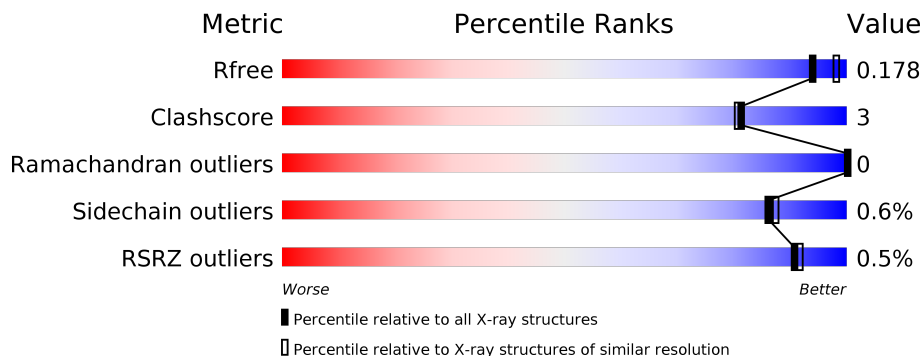|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 1.90 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		0.178
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		0.6%
RSRZ outliers		0.5%

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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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% 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	170	95% 5%
2	B	25	4% 92% 8%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 1828 atoms, of which 44 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 Bcl-2 homologous antagonist/killer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	170	1349	856	236	252	5	0	1	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	17	GLY	-	expression tag	UNP Q16611
A	18	PRO	-	expression tag	UNP Q16611
A	19	LEU	-	expression tag	UNP Q16611
A	20	GLY	-	expression tag	UNP Q16611
A	21	SER	-	expression tag	UNP Q16611
A	22	MET	-	expression tag	UNP Q16611
A	166	SER	CYS	engineered mutation	UNP Q16611

- Molecule 2 is a protein called Bcl-2-like protein 11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	25	211	129	42	39	1	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	147	ARG	TRP	engineered mutation	UNP O43521
B	162	THR	TYR	engineered mutation	UNP O43521

- Molecule 3 is trifluoroacetic acid (three-letter code: TFA) (formula: $C_2HF_3O_2$).

TFA

OXT
HO
F2 F
C1
C2
O O
F
F1
F
F3
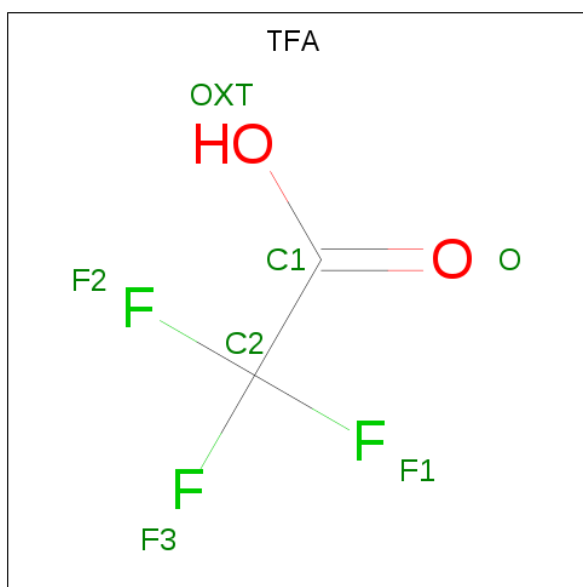
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	F	O		
3	A	1	7	2	3	2	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

EDO

O2
OH
C1
C2
HO
O1
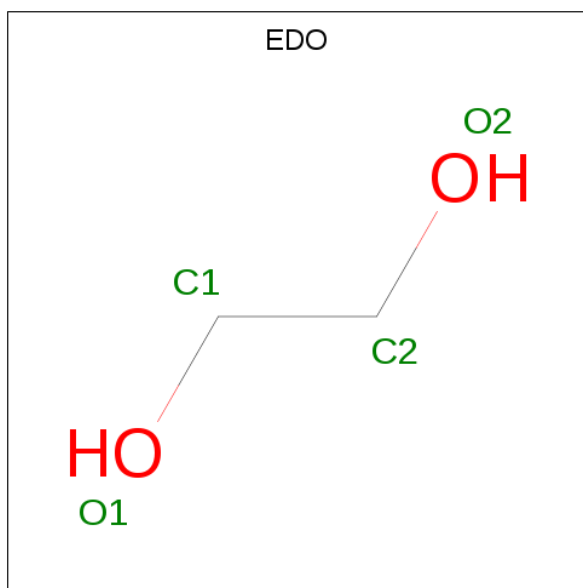
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
4	A	1	10	2	6	2	0	0
			Total	C	H	O		
4	A	1	10	2	6	2	0	0
			Total	C	H	O		
4	A	1	10	2	6	2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 10	C 2	H 6	O 2	0	0
4	A	1	Total 10	C 2	H 6	O 2	0	0
4	B	1	Total 10	C 2	H 6	O 2	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total 1	Cl 1	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
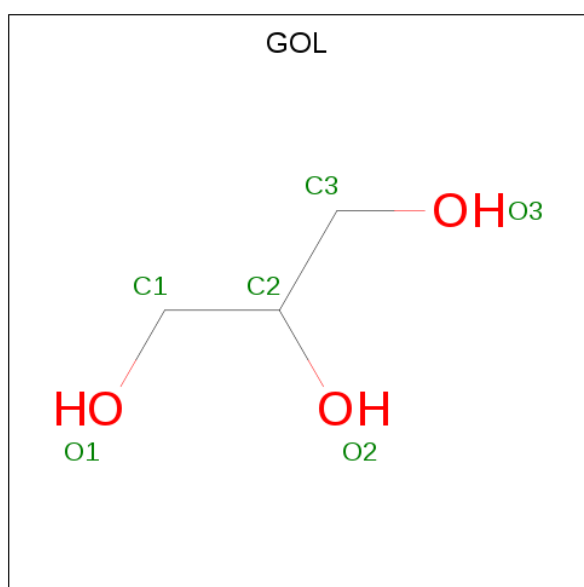
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total 14	C 3	H 8	O 3	0	0

- Molecule 7 is water.

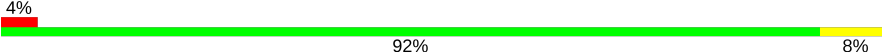| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	159	Total 159	O 159	0	0
7	B	27	Total 27	O 27	0	0

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: Bcl-2 homologous antagonist/killer

Chain A:

95%   5%

G17 G20 S21 M22 V27 Q45 I80 N124 H141 T148 G186

- Molecule 2: Bcl-2-like protein 11

Chain B:

4%

92%   8%

D141 A164 R165

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 3 2	Depositor
Cell constants a, b, c, α, β, γ	139.40Å 139.40Å 139.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.03 – 1.90 49.29 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (42.03-1.90) 94.4 (49.29-1.90)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.14 (at 1.90Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.157 , 0.177 0.157 , 0.178	Depositor DCC
R_{free} test set	1997 reflections (5.37%)	wwPDB-VP
Wilson B-factor (Å^2)	28.5	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 51.0	EDS
L-test for twinning[2]	$<|L|> = 0.48$, $<L^2> = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	1828	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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: GOL, TFA, EDO, CL

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	0.39	0/1382	0.51	0/1875
2	B	0.33	0/213	0.46	0/284
All	All	0.38	0/1595	0.50	0/2159

There are no bond length outliers.

There are no bond angle outliers.

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	1349	0	1291	7	0
2	B	211	0	207	1	0
3	A	7	0	0	0	0
4	A	20	30	30	1	0
4	B	4	6	6	0	0
5	A	1	0	0	0	0
6	A	6	8	8	1	0
7	A	159	0	0	0	1
7	B	27	0	0	0	0
All	All	1784	44	1542	8	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 3.

The worst 5 of 8 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:22:MET:HG2	1:A:27:VAL:CG2	2.37	0.55
1:A:22:MET:HG2	1:A:27:VAL:HG23	1.95	0.48
1:A:20:GLY:HA3	1:A:22:MET:CE	2.44	0.48
1:A:22:MET:CG	1:A:27:VAL:HG23	2.44	0.47
1:A:124:ASN:HD22	4:A:204:EDO:H22	1.79	0.47

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)
7:A:336:HOH:O	7:A:415:HOH:O[13_455]	2.18	0.02

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	169/170 (99%)	168 (99%)	1 (1%)	0	100	100
2	B	23/25 (92%)	22 (96%)	1 (4%)	0	100	100
All	All	192/195 (98%)	190 (99%)	2 (1%)	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	139/138 (101%)	138 (99%)	1 (1%)	84	84
2	B	21/21 (100%)	21 (100%)	0	100	100
All	All	160/159 (101%)	159 (99%)	1 (1%)	86	87

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	148	THR

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

Mol	Chain	Res	Type
1	A	45	GLN
1	A	164	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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	$\#|Z| > 2$
4	EDO	A	203	-	3,3,3	0.51	0	2,2,2	0.08	0
3	TFA	A	201	-	3,6,6	0.46	0	3,9,9	1.29	0
6	GOL	A	208	-	5,5,5	0.34	0	5,5,5	1.22	0
4	EDO	A	204	-	3,3,3	0.50	0	2,2,2	0.28	0
4	EDO	A	205	-	3,3,3	0.51	0	2,2,2	0.26	0
4	EDO	A	202	-	3,3,3	0.49	0	2,2,2	0.23	0
4	EDO	A	206	-	3,3,3	0.50	0	2,2,2	0.29	0
4	EDO	B	201	-	3,3,3	0.48	0	2,2,2	0.24	0

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
4	EDO	A	203	-	-	0/1/1/1	-
3	TFA	A	201	-	-	0/0/6/6	-
6	GOL	A	208	-	-	<mark>2/4/4/4</mark>	-
4	EDO	A	204	-	-	0/1/1/1	-
4	EDO	A	205	-	-	1/1/1/1	-
4	EDO	A	202	-	-	0/1/1/1	-
4	EDO	A	206	-	-	0/1/1/1	-
4	EDO	B	201	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	208	GOL	O1-C1-C2-C3
6	A	208	GOL	O1-C1-C2-O2
4	B	201	EDO	O1-C1-C2-O2
4	A	205	EDO	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	208	GOL	1	0
4	A	204	EDO	1	0

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)

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	170/170 (100%)	-0.41	0	100	100	21, 33, 64, 110	0
2	B	25/25 (100%)	-0.23	1 (4%)	38	41	29, 45, 65, 115	0
All	All	195/195 (100%)	-0.39	1 (0%)	91	92	21, 34, 65, 115	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	165	ARG	2.4

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EDO	A	203	4/4	0.70	0.23	75,91,99,99	0
4	EDO	A	206	4/4	0.74	0.15	70,84,89,92	0
4	EDO	B	201	4/4	0.74	0.15	63,78,90,94	0
4	EDO	A	205	4/4	0.80	0.17	59,74,89,89	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\mathring{A}^2$)	Q<0.9
4	EDO	A	202	4/4	0.80	0.21	83,100,104,107	0
4	EDO	A	204	4/4	0.86	0.12	64,77,85,97	0
6	GOL	A	208	6/6	0.87	0.20	43,62,74,81	0
3	TFA	A	201	7/7	0.95	0.23	47,53,86,88	0
5	CL	A	207	1/1	0.99	0.04	39,39,39,39	0

6.5   Other polymers (i)

There are no such residues in this entry.