



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 06:26 PM GMT

PDB ID : 1ASW
Title : AVIAN SARCOMA VIRUS INTEGRASE CATALYTIC CORE DOMAIN
CRYSTALLIZED FROM 20% PEG 4000, 10% ISOPROPANOL, HEPES
PH 7.5 USING SELENOMETHIONINE SUBSTITUTED PROTEIN; DATA
COLLECTED AT-165 DEGREES C
Authors : Bujacz, G.; Jaskolski, M.; Alexandratos, J.; Wlodawer, A.
Deposited on : 1995-08-25
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
<http://wwpdb.org/validation/2016/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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

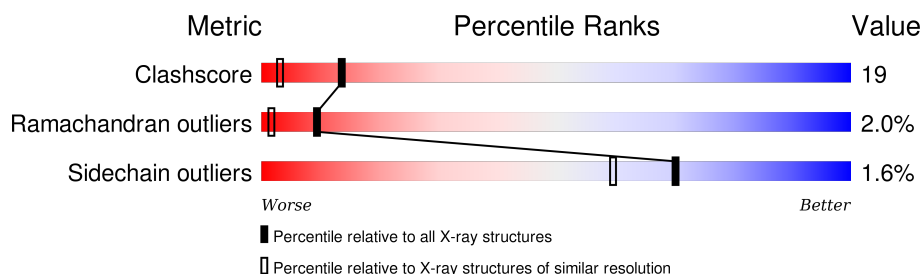
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (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 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 $\leq 5\%$. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	162	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 1388 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.

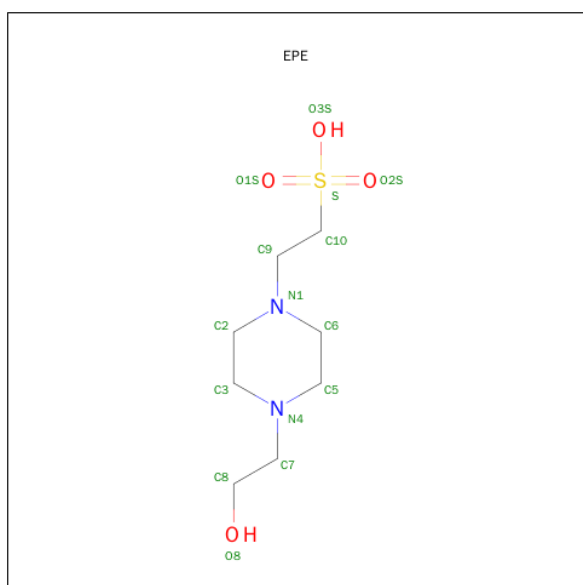
- Molecule 1 is a protein called AVIAN SARCOMA VIRUS INTEGRASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	152	1182	742	225	210	1	4	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

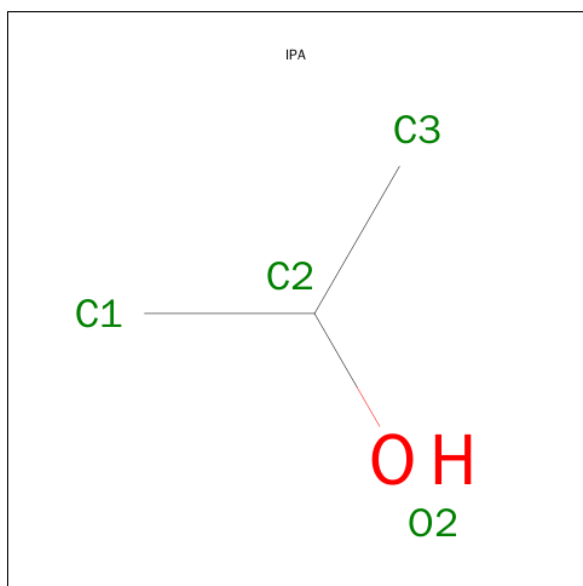
Chain	Residue	Modelled	Actual	Comment	Reference
A	71	MSE	MET	CONFLICT	UNP P03354
A	101	ALA	VAL	CONFLICT	UNP P03354
A	155	MSE	MET	CONFLICT	UNP P03354
A	166	LYS	ARG	CONFLICT	UNP P03354
A	177	MSE	MET	CONFLICT	UNP P03354
A	193	MSE	MET	CONFLICT	UNP P03354

- Molecule 2 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C₈H₁₈N₂O₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 3 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C_3H_8O).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	187	Total	O	0	0
			187	187		

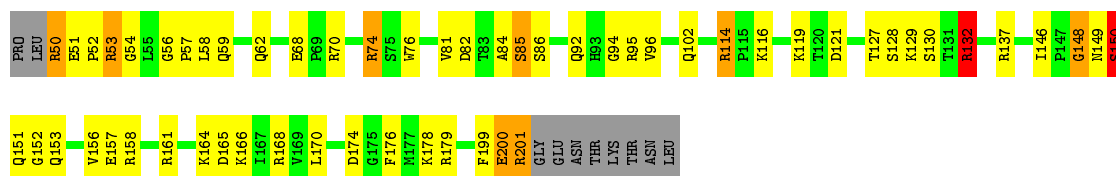
3 Residue-property plots [i](#)

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 ($\text{RSRZ} > 2$). 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.

Note EDS was not executed.

- Molecule 1: AVIAN SARCOMA VIRUS INTEGRASE

Chain A: 



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	65.40 Å 65.40 Å 80.41 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 1.80	Depositor
% Data completeness (in resolution range)	72.3 (6.00-1.80)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.139 , 0.208	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1388	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: EPE, IPA

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	0.71	0/1204	2.40	42/1624 (2.6%)

There are no bond length outliers.

All (42) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	74	ARG	NE-CZ-NH1	39.50	140.05	120.30
1	A	50	ARG	NE-CZ-NH1	-22.54	109.03	120.30
1	A	179	ARG	NE-CZ-NH1	21.98	131.29	120.30
1	A	74	ARG	NE-CZ-NH2	-18.63	110.98	120.30
1	A	137	ARG	NE-CZ-NH1	17.32	128.96	120.30
1	A	179	ARG	NE-CZ-NH2	-15.32	112.64	120.30
1	A	70	ARG	NE-CZ-NH2	11.94	126.27	120.30
1	A	137	ARG	NE-CZ-NH2	-9.94	115.33	120.30
1	A	201	ARG	NE-CZ-NH1	-9.88	115.36	120.30
1	A	168	ARG	NE-CZ-NH2	9.80	125.20	120.30
1	A	50	ARG	NH1-CZ-NH2	9.61	129.97	119.40
1	A	74	ARG	NH1-CZ-NH2	-9.60	108.84	119.40
1	A	70	ARG	NE-CZ-NH1	-8.15	116.23	120.30
1	A	121	ASP	CB-CG-OD2	-8.11	111.00	118.30
1	A	148	GLY	N-CA-C	-7.89	93.38	113.10
1	A	82	ASP	CB-CG-OD1	6.90	124.51	118.30
1	A	201	ARG	NH1-CZ-NH2	6.88	126.97	119.40
1	A	157	GLU	OE1-CD-OE2	6.79	131.45	123.30
1	A	95	ARG	NE-CZ-NH1	-6.43	117.09	120.30
1	A	85	SER	N-CA-CB	-6.28	101.08	110.50
1	A	132	ARG	NE-CZ-NH2	6.14	123.37	120.30
1	A	53	ARG	CD-NE-CZ	6.08	132.12	123.60
1	A	81	VAL	CG1-CB-CG2	-6.08	101.17	110.90
1	A	92	GLN	CG-CD-OE1	-6.00	109.60	121.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	119	LYS	CA-CB-CG	-5.92	100.38	113.40
1	A	178	LYS	CD-CE-NZ	5.87	125.21	111.70
1	A	176	PHE	CA-C-O	5.75	132.16	120.10
1	A	127	THR	CA-CB-CG2	-5.73	104.38	112.40
1	A	96	VAL	CA-CB-CG1	5.71	119.47	110.90
1	A	161	ARG	CG-CD-NE	5.56	123.47	111.80
1	A	114	ARG	NE-CZ-NH1	-5.51	117.54	120.30
1	A	130	SER	N-CA-CB	-5.45	102.32	110.50
1	A	150	SER	CA-C-O	5.42	131.49	120.10
1	A	96	VAL	CG1-CB-CG2	-5.38	102.29	110.90
1	A	158	ARG	NE-CZ-NH1	5.34	122.97	120.30
1	A	157	GLU	CA-CB-CG	-5.30	101.75	113.40
1	A	62	GLN	N-CA-CB	5.29	120.12	110.60
1	A	201	ARG	NE-CZ-NH2	-5.27	117.66	120.30
1	A	174	ASP	CB-CG-OD2	-5.26	113.57	118.30
1	A	114	ARG	CD-NE-CZ	-5.16	116.37	123.60
1	A	102	GLN	N-CA-CB	-5.15	101.33	110.60
1	A	76	TRP	CH2-CZ2-CE2	5.11	122.50	117.40

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	1182	0	1197	45	0
2	A	15	0	17	0	0
3	A	4	0	8	0	0
4	A	187	0	0	11	0
All	All	1388	0	1222	45	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 19.

All (45) 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:149:ASN:O	1:A:150:SER:HB3	1.78	0.82
1:A:146:ILE:O	1:A:149:ASN:HB2	1.79	0.81
1:A:51:GLU:HG2	1:A:52:PRO:HD2	1.62	0.81
1:A:52:PRO:HB3	1:A:58:LEU:HD23	1.64	0.79
1:A:50:ARG:N	1:A:116:LYS:HZ1	1.81	0.78
1:A:132:ARG:HG3	1:A:132:ARG:HH11	1.50	0.76
1:A:57:PRO:HB2	1:A:84:ALA:CB	2.18	0.73
1:A:132:ARG:HG3	1:A:132:ARG:NH1	2.03	0.72
1:A:51:GLU:OE1	1:A:54:GLY:N	2.23	0.71
1:A:200:GLU:C	1:A:201:ARG:HG3	2.11	0.70
1:A:52:PRO:HB2	1:A:56:GLY:O	1.91	0.70
1:A:59:GLN:HB3	4:A:427:HOH:O	1.95	0.67
1:A:52:PRO:O	1:A:53:ARG:HB3	1.97	0.65
1:A:51:GLU:CG	1:A:52:PRO:HD2	2.26	0.64
1:A:165:ASP:HB3	4:A:378:HOH:O	1.97	0.63
1:A:57:PRO:HB2	1:A:84:ALA:HB2	1.80	0.63
1:A:114:ARG:NE	4:A:391:HOH:O	2.31	0.63
1:A:51:GLU:CD	1:A:54:GLY:H	2.03	0.62
1:A:59:GLN:HA	4:A:427:HOH:O	1.99	0.62
1:A:200:GLU:O	1:A:201:ARG:HG3	2.01	0.61
1:A:128:SER:O	1:A:132:ARG:HG2	2.01	0.60
1:A:51:GLU:HG2	1:A:52:PRO:CD	2.30	0.60
1:A:132:ARG:CG	1:A:132:ARG:HH11	2.16	0.59
1:A:199:PHE:O	1:A:201:ARG:N	2.34	0.58
1:A:51:GLU:HG3	4:A:354:HOH:O	2.04	0.57
1:A:57:PRO:HB2	1:A:84:ALA:HB1	1.89	0.55
1:A:165:ASP:HB3	4:A:443:HOH:O	2.06	0.55
1:A:129:LYS:HG3	4:A:371:HOH:O	2.09	0.52
1:A:132:ARG:HD2	4:A:355:HOH:O	2.10	0.52
1:A:165:ASP:CB	4:A:443:HOH:O	2.58	0.51
1:A:59:GLN:CB	4:A:427:HOH:O	2.54	0.51
1:A:53:ARG:HG3	1:A:53:ARG:O	2.11	0.50
1:A:146:ILE:HG22	1:A:148:GLY:O	2.14	0.48
1:A:59:GLN:CA	4:A:427:HOH:O	2.58	0.48
1:A:68:GLU:OE1	1:A:164:LYS:HG2	2.13	0.47
1:A:149:ASN:OD1	1:A:152:GLY:N	2.48	0.47
1:A:153:GLN:HB2	1:A:156:VAL:HG12	1.96	0.46
1:A:51:GLU:CD	1:A:54:GLY:N	2.66	0.46
1:A:74:ARG:HA	1:A:94:GLY:HA2	1.99	0.45
1:A:52:PRO:CB	1:A:56:GLY:O	2.62	0.44
1:A:166:LYS:O	1:A:170:LEU:HG	2.18	0.43
1:A:146:ILE:O	1:A:149:ASN:CB	2.58	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:146:ILE:O	1:A:149:ASN:N	2.54	0.41
1:A:85:SER:O	1:A:86:SER:HB2	2.21	0.41
1:A:114:ARG:HH11	1:A:114:ARG:HD3	1.62	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	150/162 (93%)	139 (93%)	8 (5%)	3 (2%)	9 2

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	150	SER
1	A	151	GLN
1	A	200	GLU

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	122/127 (96%)	120 (98%)	2 (2%)	70 59

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	132	ARG
1	A	150	SER

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

Mol	Chain	Res	Type
1	A	102	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](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

2 ligands are 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	$\# Z > 2$
2	EPE	A	252	-	14,15,15	0.80	0	18,20,20	2.11	6 (33%)
3	IPA	A	253	-	3,3,3	0.67	0	3,3,3	0.57	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
2	EPE	A	252	-	-	0/9/19/19	0/1/1/1
3	IPA	A	253	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	A	252	EPE	O3S-S-O1S	-2.77	105.16	111.61
2	A	252	EPE	C5-N4-C3	2.32	113.92	108.90
2	A	252	EPE	C6-C5-N4	2.90	115.83	110.63
2	A	252	EPE	C5-C6-N1	2.95	115.92	110.63
2	A	252	EPE	O1S-S-C10	3.09	109.54	106.91
2	A	252	EPE	O2S-S-C10	5.22	111.36	106.91

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.