wwPDB X-ray Structure Validation Summary Report (i)

Mar 6, 2026 – 11:10 AM UTC

PDB ID	:	3KXS / pdb_00003kxs
Title	:	Crystal structure of HBV capsid mutant dimer (oxy form), strain adyw
Authors	:	Packianathan, C.; Katen, S.P.; Zlotnick, A.
Deposited on	:	2009-12-03
Resolution	:	2.25 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:

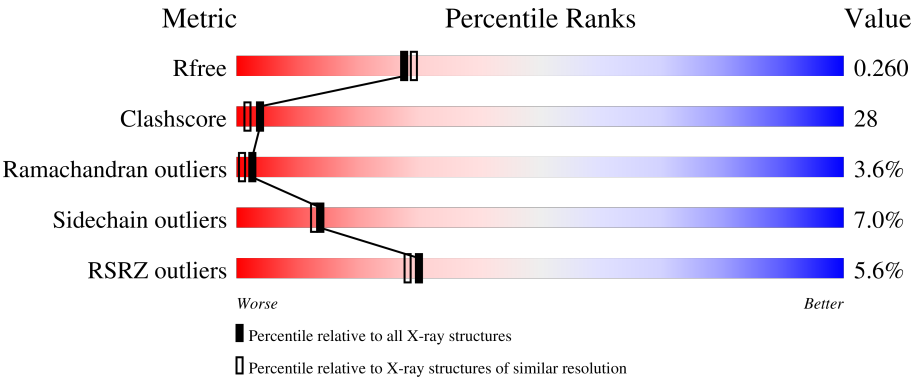
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 2.25 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.



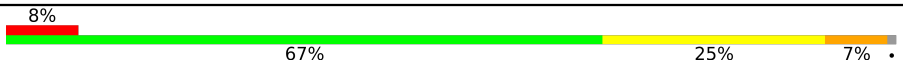
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	1898 (2.26-2.26)
Clashscore	190562	2005 (2.26-2.26)
Ramachandran outliers	187476	1965 (2.26-2.26)
Sidechain outliers	187428	1966 (2.26-2.26)
RSRZ outliers	180081	1898 (2.26-2.26)

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	143	9% 51% 36% 9%
1	B	143	5% 55% 34% 8%
1	C	143	4% 62% 35%
1	D	143	6% 54% 31% 11%
1	E	143	% 69% 24% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	143	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6932 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 Capsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	F	142	Total	C	N	O	S	0	0	0
			1125	729	186	205	5			
1	E	143	Total	C	N	O	S	0	0	0
			1134	735	187	207	5			
1	C	142	Total	C	N	O	S	0	0	0
			1119	726	183	205	5			
1	D	141	Total	C	N	O	S	0	0	0
			1109	721	182	201	5			
1	A	140	Total	C	N	O	S	0	0	0
			1112	722	184	201	5			
1	B	142	Total	C	N	O	S	0	0	0
			1125	729	186	205	5			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	132	ALA	TYR	engineered mutation	UNP P03147
E	132	ALA	TYR	engineered mutation	UNP P03147
C	132	ALA	TYR	engineered mutation	UNP P03147
D	132	ALA	TYR	engineered mutation	UNP P03147
A	132	ALA	TYR	engineered mutation	UNP P03147
B	132	ALA	TYR	engineered mutation	UNP P03147

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	F	47	Total	O	0	0
			47	47		
2	E	58	Total	O	0	0
			58	58		
2	C	24	Total	O	0	0
			24	24		

Continued on next page...

Continued from previous page...

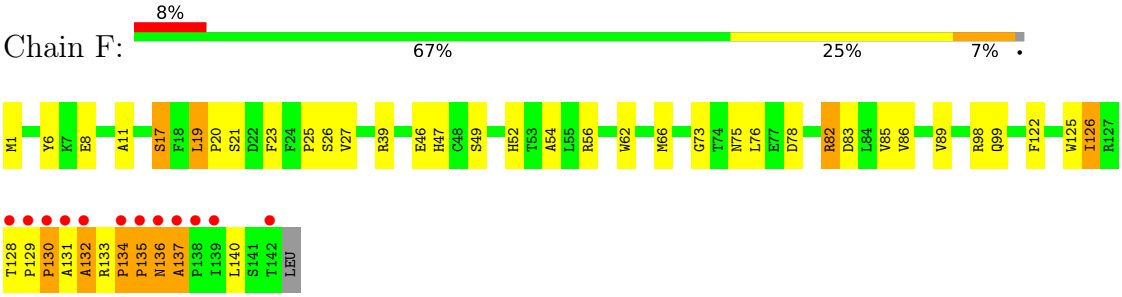
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	20	Total O 20 20	0	0
2	A	26	Total O 26 26	0	0
2	B	33	Total O 33 33	0	0

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.

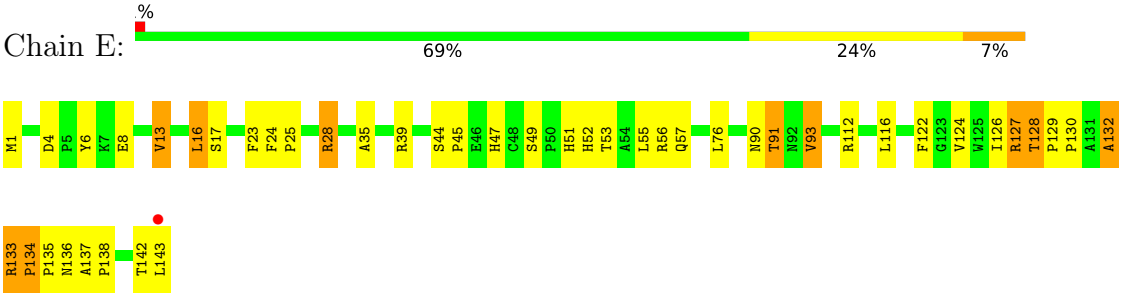
• Molecule 1: Capsid protein

Chain F:



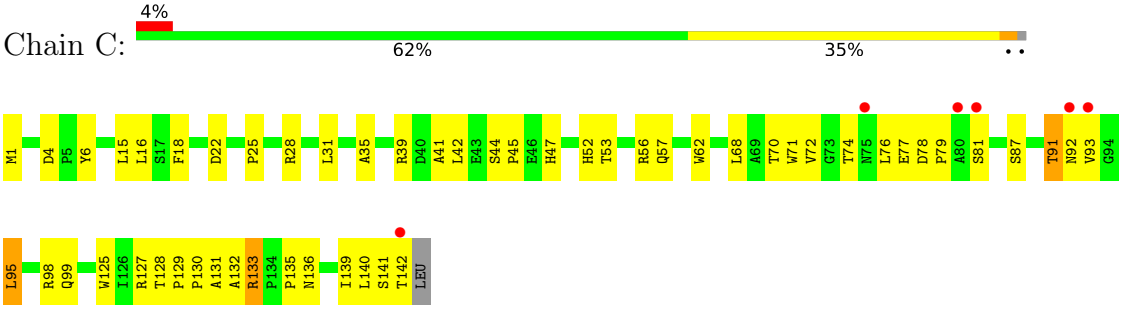
• Molecule 1: Capsid protein

Chain E:



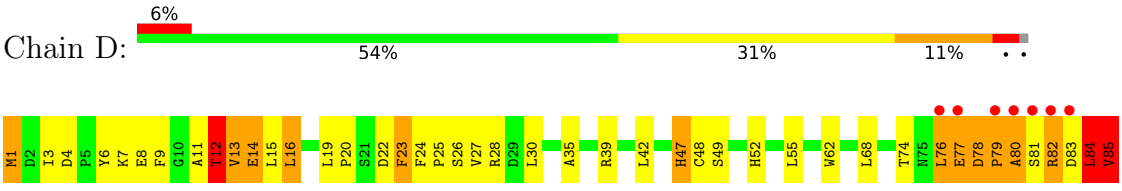
• Molecule 1: Capsid protein

Chain C:



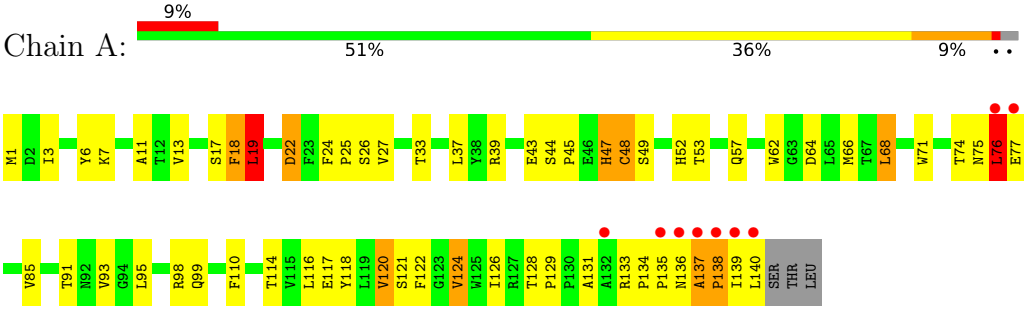
• Molecule 1: Capsid protein

Chain D:

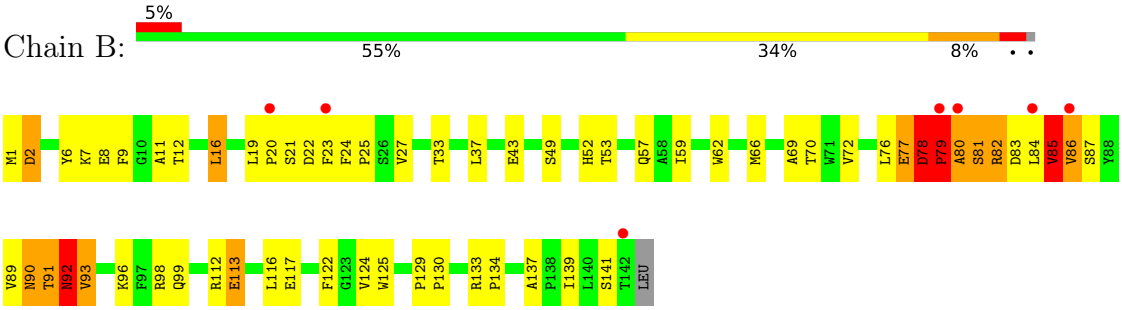




• Molecule 1: Capsid protein



• Molecule 1: Capsid protein



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	104.16Å 104.16Å 86.92Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.32 – 2.25 25.32 – 2.25	Depositor EDS
% Data completeness (in resolution range)	100.0 (25.32-2.25) 99.9 (25.32-2.25)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.04 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.0, CNS	Depositor
R, R_{free}	0.217 , 0.264 0.215 , 0.260	Depositor DCC
R_{free} test set	2547 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	42.7	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 28.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.013 for -h,-k,l 0.033 for h,-h-k,-l 0.017 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6932	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.68% 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

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	1.17	0/1146	1.14	1/1571 (0.1%)
1	B	1.24	1/1159 (0.1%)	1.21	7/1589 (0.4%)
1	C	1.12	0/1153	1.11	4/1582 (0.3%)
1	D	1.14	0/1143	1.18	8/1568 (0.5%)
1	E	1.30	1/1168 (0.1%)	1.13	3/1600 (0.2%)
1	F	1.21	1/1159 (0.1%)	1.16	8/1589 (0.5%)
All	All	1.20	3/6928 (0.0%)	1.15	31/9499 (0.3%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	78	ASP	CA-C	-6.05	1.45	1.52
1	F	54	ALA	CA-CB	5.60	1.62	1.53
1	E	13	VAL	CA-CB	5.00	1.61	1.54

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	18	PHE	N-CA-C	-11.16	95.81	110.43
1	B	137	ALA	CA-C-N	-8.81	111.30	120.03
1	B	137	ALA	C-N-CA	-8.81	111.30	120.03
1	E	93	VAL	CB-CA-C	-8.73	100.35	112.14
1	D	47	HIS	CA-C-N	-8.50	106.88	122.09

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	1112	0	1093	88	3
1	B	1125	0	1105	82	0
1	C	1119	0	1094	58	0
1	D	1109	0	1085	95	0
1	E	1134	0	1116	36	3
1	F	1125	0	1105	51	0
2	A	26	0	0	1	0
2	B	33	0	0	0	0
2	C	24	0	0	0	0
2	D	20	0	0	1	0
2	E	58	0	0	2	0
2	F	47	0	0	5	0
All	All	6932	0	6598	379	3

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 28.

The worst 5 of 379 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:B:91:THR:HG22	1:B:92:ASN:ND2	1.50	1.27
1:D:76:LEU:HD23	1:D:76:LEU:O	1.38	1.23
1:D:84:LEU:O	1:D:86:VAL:HG23	1.46	1.16
1:A:18:PHE:O	1:A:19:LEU:HB2	1.38	1.12
1:B:91:THR:CG2	1:B:92:ASN:HD22	1.63	1.10

All (3) 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:E:127:ARG:NH2	1:A:81:SER:OG[3_544]	0.72	1.48
1:E:127:ARG:CZ	1:A:81:SER:OG[3_544]	1.77	0.43
1:E:127:ARG:NH2	1:A:81:SER:CB[3_544]	2.11	0.09

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	138/143 (96%)	116 (84%)	13 (9%)	9 (6%)	1 0
1	B	140/143 (98%)	125 (89%)	8 (6%)	7 (5%)	1 0
1	C	140/143 (98%)	131 (94%)	9 (6%)	0	100 100
1	D	139/143 (97%)	120 (86%)	11 (8%)	8 (6%)	1 0
1	E	141/143 (99%)	130 (92%)	8 (6%)	3 (2%)	5 2
1	F	140/143 (98%)	134 (96%)	3 (2%)	3 (2%)	5 2
All	All	838/858 (98%)	756 (90%)	52 (6%)	30 (4%)	2 1

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	136	ASN
1	E	134	PRO
1	D	78	ASP
1	D	79	PRO
1	D	84	LEU

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	122/125 (98%)	111 (91%)	11 (9%)	9 7
1	B	124/125 (99%)	112 (90%)	12 (10%)	8 6
1	C	123/125 (98%)	118 (96%)	5 (4%)	27 33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	D	121/125 (97%)	109 (90%)	12 (10%)	7 5
1	E	125/125 (100%)	116 (93%)	9 (7%)	13 12
1	F	124/125 (99%)	121 (98%)	3 (2%)	43 54
All	All	739/750 (98%)	687 (93%)	52 (7%)	14 13

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	113	GLU
1	A	78	ASP
1	B	124	VAL
1	A	13	VAL
1	A	47	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	52	HIS
1	B	92	ASN
1	E	57	GLN
1	C	52	HIS
1	C	90	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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

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	140/143 (97%)	0.41	13 (9%)	14	13	30, 48, 92, 98	0
1	B	142/143 (99%)	0.15	7 (4%)	35	34	30, 43, 69, 77	0
1	C	142/143 (99%)	0.39	6 (4%)	40	40	29, 51, 74, 78	0
1	D	141/143 (98%)	0.48	9 (6%)	25	24	36, 55, 79, 96	0
1	E	143/143 (100%)	-0.23	1 (0%)	84	85	27, 36, 65, 76	0
1	F	142/143 (99%)	0.11	12 (8%)	16	15	27, 39, 87, 99	0
All	All	850/858 (99%)	0.22	48 (5%)	30	28	27, 45, 78, 99	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	137	ALA	6.0
1	A	81	SER	5.2
1	B	79	PRO	4.3
1	D	79	PRO	4.2
1	F	132	ALA	4.0

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 oligosaccharides 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.