



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 03:53 pm BST

PDB ID : 2ISI
Title : Crystal structure of Ape1 from Homo sapiens in a new crystal form complexed with a ligand
Authors : Agarwal, R.; Naidu, M.D.
Deposited on : 2006-10-17
Resolution : 2.76 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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 : 2.11
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) : 2.11

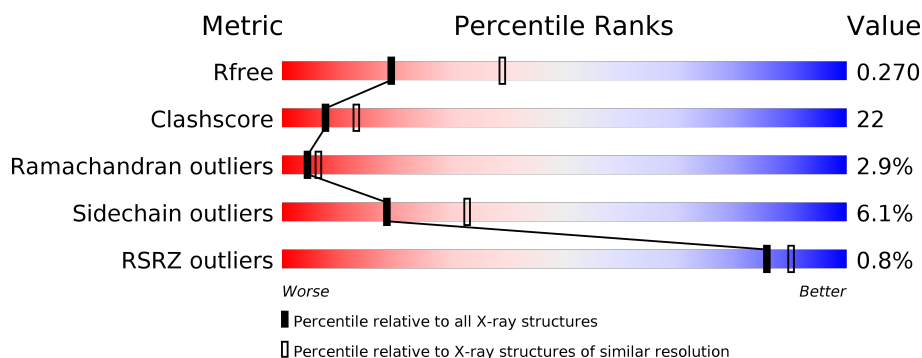
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.76 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}	130704	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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 $\leq 5\%$. 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.

Mol	Chain	Length	Quality of chain
1	A	317	<div> <div>56%</div> <div>28%</div> <div>•</div> <div>13%</div> </div>
1	B	317	<div> <div>50%</div> <div>33%</div> <div>•</div> <div>13%</div> </div>
1	C	317	<div> <div>2%</div> <div>42%</div> <div>39%</div> <div>•</div> <div>14%</div> </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6610 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 DNA-(apurinic or apyrimidinic site) lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2198	1404	380	405	9			
1	B	276	Total	C	N	O	S	0	0	0
			2187	1398	379	401	9			
1	C	273	Total	C	N	O	S	0	0	0
			2175	1391	376	399	9			

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	P	0	0
			5	4	1		
2	B	1	Total	O	P	0	0
			5	4	1		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	Mg 1	0	0
3	A	1	Total 1	Mg 1	0	0
3	C	1	Total 1	Mg 1	0	0

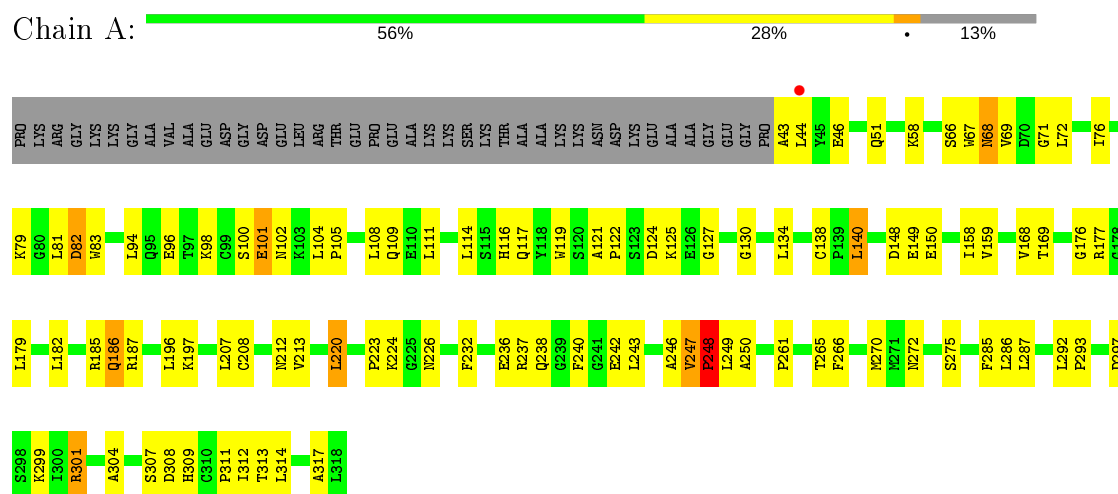
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	16	Total 16	O 16	0	0
4	B	7	Total 7	O 7	0	0
4	C	14	Total 14	O 14	0	0

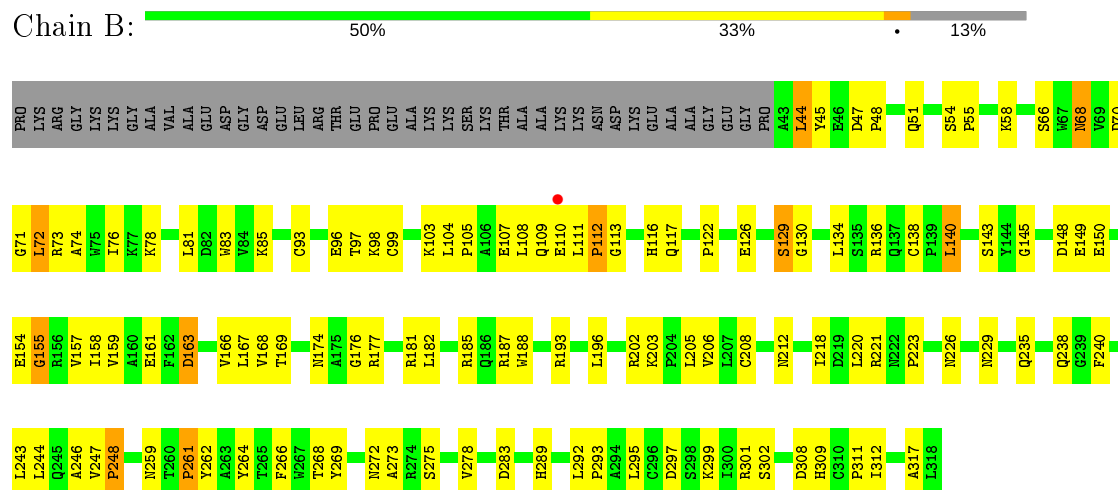
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: DNA-(apurinic or apyrimidinic site) lyase



- Molecule 1: DNA-(apurinic or apyrimidinic site) lyase



- Molecule 1: DNA-(apurinic or apyrimidinic site) lyase



A304	P223	L140	H68	PR0
H309	K227	I146	H69	LYS
C310	K228		D70	ARG
P311	N229	E149	G71	GLY
I312			L72	LYS
	F232	Q153	H73	LYS
L316		E154	A74	GLY
A317	Q235		H75	ALA
L318		I158	I76	VAL
	F240		K77	ALA
			K78	GLU
	L243	A160		ASP
	A246	E161	L81	GLY
	V247	F162	H82	ASP
	P248	D163	H83	GLU
	L249	S164		LEU
		F165	I91	ARG
		V166		THR
	S252	L167	I94	GLU
	F253	V168	H95	PRO
	R254		E96	GLU
	H255	V172	T97	ALA
	L256	P173	K98	LYS
	V257		C99	LYS
	P258	R177	S100	SER
	N259	G178	E101	LYS
	P261	L179	H102	THR
			K103	ALA
	T265	L182	L104	ALA
	F266	E183	P105	LYS
	H267	R185	A106	LYS
		Q186	E107	ASN
	M270		L108	ASP
	N271	D189	Q109	LYS
	N272	E190	E110	GLU
	A273		L111	ALA
	R274	R193	P112	ALA
		K194		GLY
	D283	F195	H116	GLU
	F284	L196	Q117	GLY
	L285		V118	PRO
	L286		W119	A43
	L287	L198	S120	L44
	S288		A121	
	H289	K203	P122	P48
	S290	P204	SER	P49
	L291	L205	ASP	D50
	L292	V206	LYS	Q51
	P293	L207	E126	K52
	A294	C208	G127	T53
	L295		Y128	S54
	C296	N212	S129	P55
		V213	G130	
			V131	K58
	K299	E216		P59
	I300	E217	L134	A60
	R301	T218	S135	T61
	S302	D219		
	K303	L220	C138	S66
			P139	W67

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	167.35Å 92.39Å 94.01Å 90.00° 121.38° 90.00°	Depositor
Resolution (Å)	32.13 – 2.76 42.33 – 2.76	Depositor EDS
% Data completeness (in resolution range)	82.9 (32.13-2.76) 82.9 (42.33-2.76)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.11 (at 2.77Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.228 , 0.284 0.224 , 0.270	Depositor DCC
R_{free} test set	753 reflections (2.64%)	wwPDB-VP
Wilson B-factor (Å ²)	50.9	Xtriage
Anisotropy	0.402	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 43.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.024 for -h+k-l,-l,-k 0.005 for -h-k-l,l,k 0.027 for -h-2*l,-k,l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	6610	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.89% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: PO4, MG

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	0.44	0/2256	0.71	0/3058
1	B	0.43	0/2245	0.70	0/3045
1	C	0.48	0/2232	0.70	0/3025
All	All	0.45	0/6733	0.70	0/9128

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	2198	0	2172	82	0
1	B	2187	0	2157	93	0
1	C	2175	0	2151	118	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	16	0	0	1	0
4	B	7	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	14	0	0	3	0
All	All	6610	0	6480	286	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 22.

The worst 5 of 286 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:105:PRO:HG2	1:B:108:LEU:HD13	1.56	0.86
1:A:111:LEU:HD13	1:A:114:LEU:HD12	1.58	0.86
1:C:105:PRO:HD2	1:C:108:LEU:HD22	1.59	0.85
1:A:177:ARG:HG3	1:A:177:ARG:HH11	1.41	0.84
1:A:58:LYS:HD2	1:A:317:ALA:HB1	1.60	0.84

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	274/317 (86%)	240 (88%)	26 (10%)	8 (3%)	4	6
1	B	274/317 (86%)	236 (86%)	30 (11%)	8 (3%)	4	6
1	C	269/317 (85%)	228 (85%)	33 (12%)	8 (3%)	4	6
All	All	817/951 (86%)	704 (86%)	89 (11%)	24 (3%)	4	6

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	149	GLU
1	A	248	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	112	PRO
1	C	49	PRO
1	C	301	ARG

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	236/265 (89%)	221 (94%)	15 (6%)	17	31
1	B	233/265 (88%)	221 (95%)	12 (5%)	23	39
1	C	233/265 (88%)	217 (93%)	16 (7%)	15	27
All	All	702/795 (88%)	659 (94%)	43 (6%)	18	33

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	140	LEU
1	B	248	PRO
1	C	287	LEU
1	B	143	SER
1	B	163	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	116	HIS
1	B	117	GLN
1	C	116	HIS
1	B	68	ASN
1	B	95	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](#)

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 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	$\# Z > 2$
2	PO4	B	401	-	4,4,4	1.56	0	6,6,6	0.70	0
2	PO4	A	400	-	4,4,4	1.63	0	6,6,6	0.65	0

There are no bond length outliers.

There are no bond angle outliers.

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 [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	276/317 (87%)	-0.19	1 (0%) 92 95	33, 53, 68, 79	0
1	B	276/317 (87%)	-0.19	1 (0%) 92 95	38, 57, 74, 84	0
1	C	273/317 (86%)	0.04	5 (1%) 68 76	47, 65, 77, 82	0
All	All	825/951 (86%)	-0.12	7 (0%) 86 90	33, 59, 75, 84	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	43	ALA	3.8
1	A	44	LEU	2.9
1	C	44	LEU	2.9
1	C	199	LEU	2.8
1	C	109	GLN	2.3

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 carbohydrates 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(\AA^2)	Q<0.9
3	MG	B	403	1/1	0.83	0.45	53,53,53,53	0
3	MG	A	402	1/1	0.89	0.25	46,46,46,46	0
3	MG	C	404	1/1	0.93	0.35	66,66,66,66	0
2	PO4	B	401	5/5	0.93	0.20	90,91,92,92	0
2	PO4	A	400	5/5	0.96	0.18	78,80,82,82	0

6.5 Other polymers [i](#)

There are no such residues in this entry.