wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 9, 2023 – 10:05 AM EDT

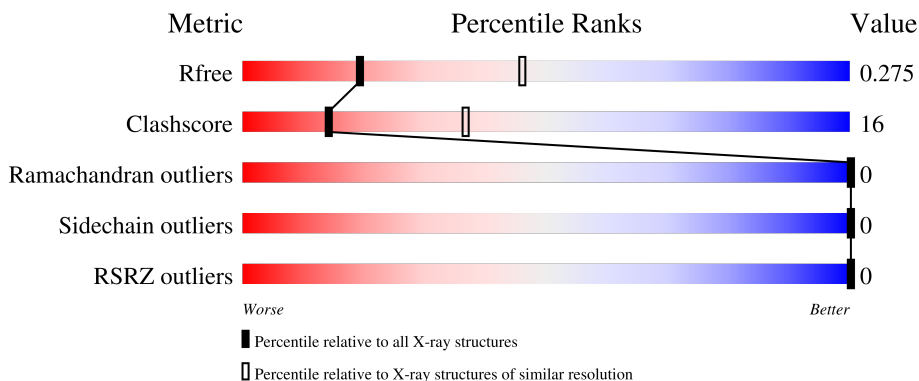|		
PDB ID	:	6W4T
Title	:	APE1 Y269A phosphorothioate substrate complex with abasic DNA
Authors	:	Freudenthal, B.D.; Hoitsma, N.M.
Deposited on	:	2020-03-11
Resolution	:	2.77 Å(reported)

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 2.77 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		0.275
Clashscore		16
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		0

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	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)
RSRZ outliers	127900	4027 (2.80-2.76)

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	276	73% 27%
1	B	276	68% 32%
2	C	21	48% 52%
2	P	21	33% 67%
3	F	21	29% 71%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	V	21	52% 48%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 6132 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
			2190	1398	380	404	8			
1	B	276	Total	C	N	O	S	0	0	0
			2190	1398	380	404	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	138	ALA	CYS	engineered mutation	UNP P27695
A	269	ALA	TYR	engineered mutation	UNP P27695
B	138	ALA	CYS	engineered mutation	UNP P27695
B	269	ALA	TYR	engineered mutation	UNP P27695

- Molecule 2 is a DNA chain called DNA (5'-D(P*GP*CP*TP*GP*AP*TP*GP*CP*GP*TP*(48Z)P*CP*GP*AP*CP*GP*GP*AP*TP*CP*C)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	P	21	Total	C	N	O	P	S	0	1	0
			433	204	76	129	22	2			
2	C	21	Total	C	N	O	P	S	0	1	0
			428	203	76	126	21	2			

- Molecule 3 is a DNA chain called DNA (5'-D(*GP*GP*AP*TP*CP*CP*GP*TP*CP*GP*GP*AP*CP*GP*CP*AP*TP*CP*AP*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	V	21	Total	C	N	O	P	0	0	0
			426	202	82	122	20			
3	F	21	Total	C	N	O	P	0	0	0
			426	202	82	122	20			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	16	Total 16	O 16	0	0
4	P	4	Total 4	O 4	0	0
4	V	1	Total 1	O 1	0	0
4	B	18	Total 18	O 18	0	0

3 Residue-property plots ⓘ

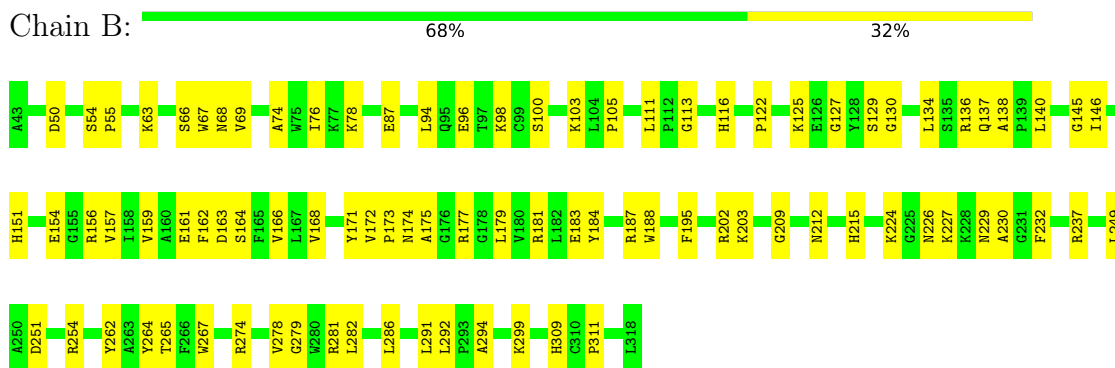These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: DNA-(apurinic or apyrimidinic site) lyase
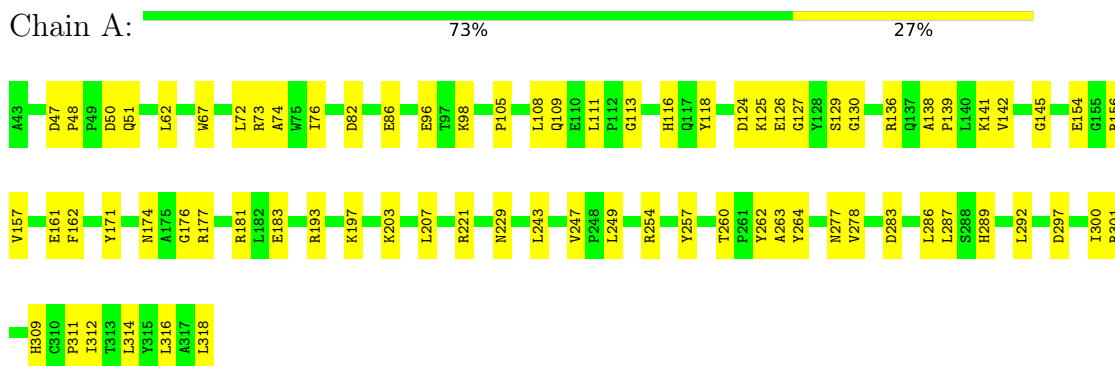
Chain A:

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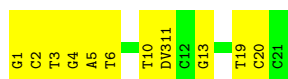
Chain B:

• Molecule 2: DNA (5'-D(P*GP*CP*TP*GP*AP*TP*GP*CP*GP*TP*(48Z)P*CP*GP*AP*C P*GP*GP*AP*TP*CP*C)-3')

Chain P:

• Molecule 2: DNA (5'-D(P*GP*CP*TP*GP*AP*TP*GP*CP*GP*TP*(48Z)P*CP*GP*AP*C P*GP*GP*AP*TP*CP*C)-3')

Chain C:

48%   52%

G1 C2 T3 G4 A5 T6 T10 DV311 C12 G13 T19 C20 C21

• Molecule 3: DNA (5'-D(*GP*GP*AP*TP*CP*CP*GP*TP*CP*GP*GP*AP*CP*GP*CP*AP
*TP*CP*AP*GP*C)-3')

Chain V:

52%   48%

G1 G2 C5 G10 G11 A12 C13 G14 C15 A16 T17 C21

• Molecule 3: DNA (5'-D(*GP*GP*AP*TP*CP*CP*GP*TP*CP*GP*GP*AP*CP*GP*CP*AP
*TP*CP*AP*GP*C)-3')

Chain F:

29%   71%

G1 G2 A3 T4 C5 C6 G7 T8 C9 G10 G11 A12 C13 G14 C15 A16 T17 C18 A19 G20 C21
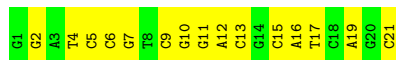
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	45.66Å 148.31Å 153.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.29 – 2.77 48.29 – 2.77	Depositor EDS
% Data completeness (in resolution range)	98.7 (48.29-2.77) 98.7 (48.29-2.77)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.25 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.221 , 0.274 0.222 , 0.275	Depositor DCC
R_{free} test set	1267 reflections (4.69%)	wwPDB-VP
Wilson B-factor (Å^2)	56.9	Xtriage
Anisotropy	0.572	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 20.4	EDS
L-test for twinning[2]	$< |L| > = 0.44, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.074 for -h,l,k	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6132	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.35% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: DV3

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	0.60	0/2247	0.69	0/3046
1	B	0.69	0/2247	0.71	0/3046
2	C	0.92	0/454	0.91	0/697
2	P	0.92	0/459	0.95	0/704
3	F	0.93	0/478	0.98	0/736
3	V	0.81	0/478	0.92	0/736
All	All	0.73	0/6363	0.79	0/8965

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 ⓘ

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	2190	0	2170	61	0
1	B	2190	0	2170	73	0
2	C	428	0	222	18	0
2	P	433	0	225	16	0
3	F	426	0	232	21	0
3	V	426	0	232	12	0
4	A	16	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	18	0	0	0	0
4	P	4	0	0	0	0
4	V	1	0	0	0	0
All	All	6132	0	5251	175	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 16.

The worst 5 of 175 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:174:ASN:HB2	2:C:11[A]:DV3:SP3	2.22	0.78
1:A:221:ARG:HG2	1:A:278:VAL:HG13	1.65	0.78
2:C:3:DT:H3	3:F:19:DA:H62	1.32	0.78
1:B:156:ARG:NH1	1:B:171:TYR:HD2	1.83	0.75
1:B:156:ARG:NH1	1:B:171:TYR:CD2	2.56	0.74

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	274/276 (99%)	253 (92%)	21 (8%)	0	100	100
1	B	274/276 (99%)	249 (91%)	25 (9%)	0	100	100
All	All	548/552 (99%)	502 (92%)	46 (8%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	234/234 (100%)	234 (100%)	0	100	100
1	B	234/234 (100%)	234 (100%)	0	100	100
All	All	468/468 (100%)	468 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides 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 (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(Å²)	Q<0.9
1	A	276/276 (100%)	-0.06	0	100	100	28, 44, 60, 77	0
1	B	276/276 (100%)	-0.07	0	100	100	32, 46, 64, 79	0
2	C	20/21 (95%)	-0.70	0	100	100	52, 80, 94, 94	0
2	P	20/21 (95%)	-0.83	0	100	100	52, 80, 89, 94	0
3	F	21/21 (100%)	-0.79	0	100	100	53, 73, 88, 91	0
3	V	21/21 (100%)	-0.79	0	100	100	52, 76, 87, 91	0
All	All	634/636 (99%)	-0.16	0	100	100	28, 46, 81, 94	0

There are no RSRZ outliers to report.

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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK