W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Dec 14, 2023 – 05:26 pm GMT

|		
PDB ID	:	2XO7
Title	:	Crystal structure of a dA:O-allylhydroxylamine-dC basepair in complex with fragment DNA polymerase I from Bacillus stearothermophilus
Authors	:	Muenzel, M.; Lercher, L.; Mueller, M.; Carell, T.
Deposited on	:	2010-08-10
Resolution	:	2.85 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 2.85 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.
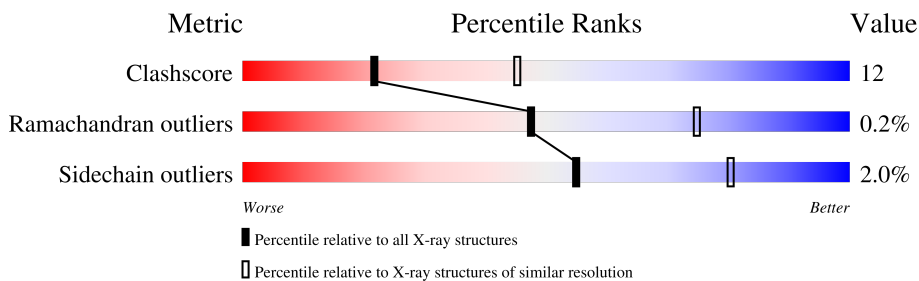
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 5088 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 POLYMERASE I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	580	Total	C	N	O	S	0	0	0
			4633	2945	808	863	17			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	512	ARG	GLY	SEE REMARK 999	UNP C3IXT2
A	550	SER	THR	SEE REMARK 999	UNP C3IXT2

- Molecule 2 is a DNA chain called 5'-D(*GP*AP*CP*CP*AP*TP*47C*CP*CP*T)-3'.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	10	Total	C	N	O	P	0	0	0
			200	98	34	59	9			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	11	Total	C	N	O	P	0	0	0
			228	109	47	62	10			

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	7	Total 7	O 7	0	0

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics ⓘ

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	87.48Å 93.64Å 105.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.74 – 2.85	Depositor
% Data completeness (in resolution range)	99.4 (43.74-2.85)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.02 (at 2.86Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.215 , 0.245	Depositor
Wilson B-factor (Å^2)	51.1	Xtriage
Anisotropy	0.145	Xtriage
L-test for twinning[2]	$<\lvert L \rvert> = 0.49$, $<L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5088	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.67% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $<\lvert L \rvert>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

4 Model quality ⓘ

4.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 47C, SO4

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	0.68	2/4715 (0.0%)	0.69	0/6371
2	B	1.07	0/196	1.84	5/297 (1.7%)
3	C	1.24	0/257	1.78	6/396 (1.5%)
All	All	0.74	2/5168 (0.0%)	0.87	11/7064 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	309	GLU	CB-CG	6.57	1.64	1.52
1	A	309	GLU	CG-CD	5.94	1.60	1.51

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	13	DC	P-O3'-C3'	10.19	131.92	119.70
2	B	20	DG	O4'-C4'-C3'	-9.89	100.07	106.00
3	C	11	DG	O4'-C1'-N9	-7.48	102.77	108.00
3	C	13	DC	O4'-C1'-N1	6.83	112.78	108.00
3	C	14	DA	O4'-C1'-N9	6.68	112.67	108.00

There are no chirality outliers.

There are no planarity outliers.

4.2 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	4633	0	4688	108	0
2	B	200	0	118	5	0
3	C	228	0	125	7	0
4	A	20	0	0	1	0
5	A	7	0	0	0	0
All	All	5088	0	4931	117	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 12.

The worst 5 of 117 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:308:THR:H	1:A:311:MET:CE	1.52	1.21
1:A:308:THR:H	1:A:311:MET:HE1	1.01	1.11
1:A:308:THR:N	1:A:311:MET:HE1	1.84	0.92
1:A:308:THR:N	1:A:311:MET:CE	2.34	0.91
1:A:363:ASP:OD1	1:A:365:THR:OG1	1.88	0.89

There are no symmetry-related clashes.

4.3 Torsion angles ⓘ

4.3.1 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	578/580 (100%)	551 (95%)	26 (4%)	1 (0%)	47	75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	727	ILE

4.3.2 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	491/496 (99%)	481 (98%)	10 (2%)	55	80

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	727	ILE
1	A	823	ARG
1	A	827	GLN
1	A	645	SER
1	A	665	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	529	ASN
1	A	543	GLN
1	A	579	GLN
1	A	608	GLN

4.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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.

4.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

4.6 Ligand geometry ⓘ

4 ligands are modelled in this entry.

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.

4.7 Other polymers ⓘ

There are no such residues in this entry.

4.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

5 Fit of model and data ⓘ

5.1 Protein, DNA and RNA chains ⓘ

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates ⓘ

EDS failed to run properly - this section is therefore empty.

5.4 Ligands ⓘ

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers ⓘ

EDS failed to run properly - this section is therefore empty.