



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 12:17 PM UTC

PDB ID : 4Q5O / pdb_00004q5o
Title : Crystal structure of EctD from *S. alaskensis* with 2-oxoglutarate and 5-hydroxyectoine
Authors : Hoepfner, A.; Widderich, N.; Bremer, E.; Smits, S.H.
Deposited on : 2014-04-17
Resolution : 2.64 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
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

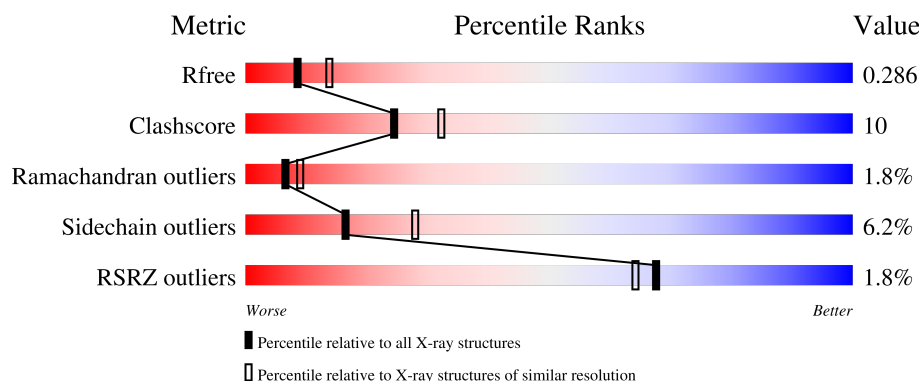
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.64 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	2053 (2.66-2.62)
Clashscore	190562	2097 (2.66-2.62)
Ramachandran outliers	187476	2066 (2.66-2.62)
Sidechain outliers	187428	2066 (2.66-2.62)
RSRZ outliers	180081	2052 (2.66-2.62)

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 $\leq 5\%$. 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	314	68% 19% • 10%
1	B	314	66% 21% •• 10%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	AKG	A	402	-	X	-	-
4	6CS	B	403	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4469 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 Ectoine hydroxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	282	Total	C	N	O	S	0	0	0
			2198	1389	397	403	9			
1	B	282	Total	C	N	O	S	0	0	0
			2198	1389	397	403	9			

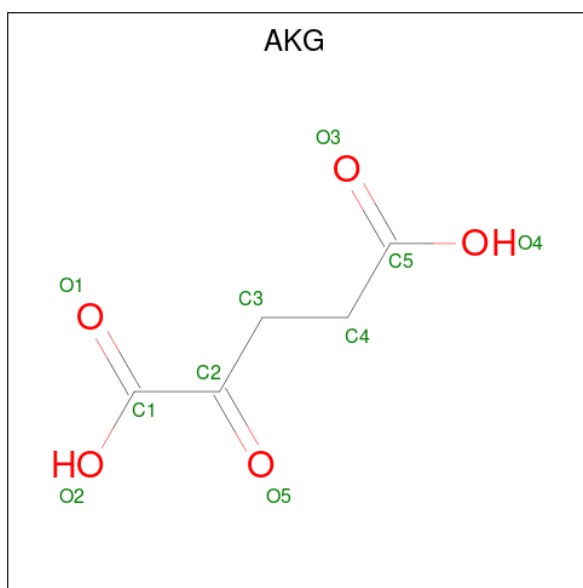
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	HIS	-	expression tag	UNP Q1GNW5
A	-6	HIS	-	expression tag	UNP Q1GNW5
A	-5	HIS	-	expression tag	UNP Q1GNW5
A	-4	HIS	-	expression tag	UNP Q1GNW5
A	-3	HIS	-	expression tag	UNP Q1GNW5
A	-2	HIS	-	expression tag	UNP Q1GNW5
A	-1	SER	-	expression tag	UNP Q1GNW5
A	0	GLU	-	expression tag	UNP Q1GNW5
B	-7	HIS	-	expression tag	UNP Q1GNW5
B	-6	HIS	-	expression tag	UNP Q1GNW5
B	-5	HIS	-	expression tag	UNP Q1GNW5
B	-4	HIS	-	expression tag	UNP Q1GNW5
B	-3	HIS	-	expression tag	UNP Q1GNW5
B	-2	HIS	-	expression tag	UNP Q1GNW5
B	-1	SER	-	expression tag	UNP Q1GNW5
B	0	GLU	-	expression tag	UNP Q1GNW5

- Molecule 2 is FE (III) ION (CCD ID: FE) (formula: Fe).

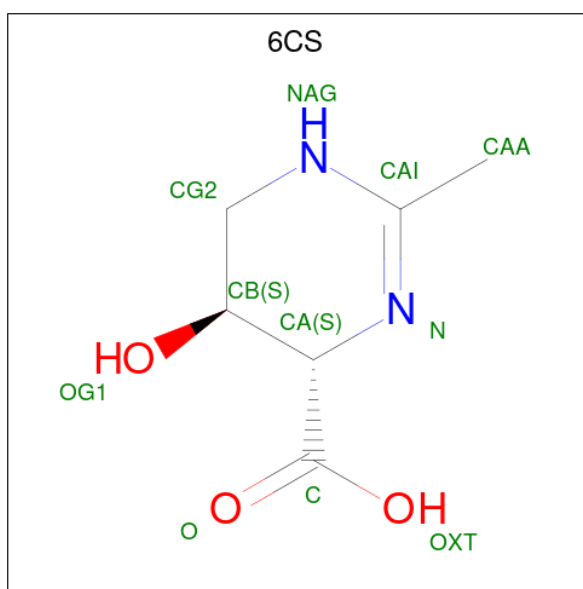
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Fe	0	0
			1	1		
2	B	1	Total	Fe	0	0
			1	1		

- Molecule 3 is 2-OXOGLUTARIC ACID (CCD ID: AKG) (formula: $C_5H_6O_5$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			10	5	5		
3	B	1	Total	C	O	0	0
			10	5	5		

- Molecule 4 is (4S,5S)-5-HYDROXY-2-METHYL-1,4,5,6-TETRAHYDOPYRIMIDINE-4-CARBOXYLIC ACID (CCD ID: 6CS) (formula: $C_6H_{10}N_2O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			11	6	2	3		
4	B	1	Total	C	N	O	0	0
			11	6	2	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	14	Total	O	0	0
			14	14		
5	B	15	Total	O	0	0
			15	15		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	81.00Å 87.08Å 94.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.64 30.00 – 2.64	Depositor EDS
% Data completeness (in resolution range)	99.8 (30.00-2.64) 99.7 (30.00-2.64)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.81 (at 2.65Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.206 , 0.279 0.212 , 0.286	Depositor DCC
R_{free} test set	1035 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å ²)	57.6	Xtriage
Anisotropy	0.231	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 40.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4469	wwPDB-VP
Average B, all atoms (Å ²)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.52% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: AKG, 6CS, FE

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	0.57	0/2255	0.93	6/3070 (0.2%)
1	B	0.59	0/2255	0.96	6/3070 (0.2%)
All	All	0.58	0/4510	0.94	12/6140 (0.2%)

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	83	GLU	CA-C-N	9.83	132.12	119.84
1	B	83	GLU	C-N-CA	9.83	132.12	119.84
1	A	279	PRO	N-CA-CB	7.47	110.81	102.60
1	A	3	ASP	N-CA-C	7.11	119.39	109.15
1	B	279	PRO	N-CA-CB	7.10	110.41	102.60

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	2198	0	2123	42	0
1	B	2198	0	2123	48	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	10	0	4	0	0
3	B	10	0	4	2	0
4	A	11	0	8	2	0
4	B	11	0	9	8	0
5	A	14	0	0	0	0
5	B	15	0	0	0	0
All	All	4469	0	4271	86	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 10.

The worst 5 of 86 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:146:ASP:OD1	4:A:403:6CS:OG1	1.75	1.05
1:B:146:ASP:OD1	4:B:403:6CS:OG1	1.89	0.90
1:A:252:THR:HG22	1:B:179:MET:HE1	1.60	0.83
1:B:127:GLN:HE22	4:B:403:6CS:HAA1	1.45	0.80
1:A:227:VAL:HG12	1:A:229:PRO:HD3	1.67	0.75

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	278/314 (88%)	261 (94%)	12 (4%)	5 (2%)	6	9
1	B	278/314 (88%)	253 (91%)	20 (7%)	5 (2%)	6	9
All	All	556/628 (88%)	514 (92%)	32 (6%)	10 (2%)	6	9

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	85	GLN
1	A	277	GLU
1	A	279	PRO
1	B	83	GLU
1	B	84	PRO

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	227/262 (87%)	219 (96%)	8 (4%)	32	51
1	B	227/262 (87%)	207 (91%)	20 (9%)	9	14
All	All	454/524 (87%)	426 (94%)	28 (6%)	16	28

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	129	ARG
1	B	301	VAL
1	B	149	THR
1	B	267	SER
1	B	139	ARG

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	224	HIS
1	A	250	ASN
1	B	171	ASN
1	B	258	ASN

5.3.3 RNA ⓘ

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](#)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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$
3	AKG	B	402	2	9,9,9	0.99	0	11,11,11	1.92	3 (27%)
4	6CS	A	403	2	10,11,11	2.52	4 (40%)	10,15,15	4.42	6 (60%)
4	6CS	B	403	-	10,11,11	3.08	5 (50%)	10,15,15	2.87	2 (20%)
3	AKG	A	402	2	9,9,9	1.01	0	11,11,11	2.11	6 (54%)

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
3	AKG	B	402	2	-	6/9/9/9	-
4	6CS	A	403	2	-	0/4/17/17	0/1/1/1
4	6CS	B	403	-	-	0/4/17/17	0/1/1/1
3	AKG	A	402	2	-	7/9/9/9	-

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	403	6CS	CG2-NAG	-6.64	1.39	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	403	6CS	CG2-NAG	-4.93	1.41	1.46
4	B	403	6CS	CA-N	-4.44	1.42	1.46
4	A	403	6CS	CAI-NAG	3.59	1.43	1.35
4	B	403	6CS	CAI-NAG	3.56	1.43	1.35

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	403	6CS	CB-CG2-NAG	11.24	121.28	109.83
4	B	403	6CS	CB-CG2-NAG	8.23	118.22	109.83
4	A	403	6CS	CB-CA-C	5.88	120.36	111.66
3	B	402	AKG	C3-C4-C5	-3.67	103.93	113.67
4	A	403	6CS	CA-N-CAI	3.00	127.10	117.28

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	402	AKG	O1-C1-C2-O5
3	A	402	AKG	O1-C1-C2-C3
3	A	402	AKG	O2-C1-C2-C3
3	A	402	AKG	C1-C2-C3-C4
3	B	402	AKG	O2-C1-C2-C3

There are no ring outliers.

3 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	402	AKG	2	0
4	A	403	6CS	2	0
4	B	403	6CS	8	0

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	282/314 (89%)	-0.24	3 (1%) 78 75	32, 49, 75, 100	0
1	B	282/314 (89%)	-0.14	7 (2%) 58 54	36, 55, 81, 103	0
All	All	564/628 (89%)	-0.19	10 (1%) 67 64	32, 52, 80, 103	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	301	VAL	3.4
1	B	84	PRO	3.2
1	B	85	GLN	3.2
1	A	288	GLY	2.9
1	B	276	VAL	2.7

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
4	6CS	A	403	11/11	0.66	0.20	69,77,80,82	0
4	6CS	B	403	11/11	0.68	0.18	66,70,76,77	0
3	AKG	A	402	10/10	0.69	0.19	66,78,86,90	0
3	AKG	B	402	10/10	0.77	0.13	54,59,71,72	0
2	FE	A	401	1/1	0.94	0.05	61,61,61,61	0
2	FE	B	401	1/1	0.98	0.03	52,52,52,52	0

6.5 Other polymers [i](#)

There are no such residues in this entry.