

wwPDB X-ray Structure Validation Summary Report (i)

Jan 3, 2024 - 08:55 am GMT

PDB ID	:	4UTT
Title	:	Structural characterisation of NanE, ManNac6P C2 epimerase, from Clostrid-
		ium perfingens
Authors	:	Pelissier, M.C.; Sebban-Kreuzer, C.; Guerlesquin, F.; Brannigan, J.A.; Davies,
		G.J.; Bourne, Y.; Vincent, F.
Deposited on		
Resolution	:	1.71 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 (1)) were used in the production of this report:

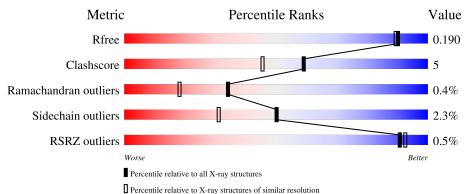
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as 541 be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 1.71 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	130704	5722(1.74-1.70)
Clashscore	141614	6152(1.74-1.70)
Ramachandran outliers	138981	6051 (1.74-1.70)
Sidechain outliers	138945	6051 (1.74-1.70)
RSRZ outliers	127900	5629(1.74-1.70)

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	А	229	90%	9%	-
1	В	229	90%	9%	•
2	С	229	88%	11%	•
2	D	229	% 91%	8%	•

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	\mathbf{Res}	Chirality	Geometry	Clashes	Electron density
4	ACT	D	1223	-	Х	-	-



2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8252 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 PUTATIVE N-ACETYLMANNOSAMINE-6-PHOSPHATE 2-EPIMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Δ	229	Total	С	Ν	0	S	0	5	0
T	11	225	1782	1117	308	347	10	0	5	
1	Р	229	Total	С	Ν	0	\mathbf{S}	0	Б	0
	D	229	1784	1123	307	344	10	U	0	0

Chain	Residue	Modelled	Actual	Comment	Reference
А	-8	GLY	-	expression tag	UNP Q0TUP9
А	-7	SER	-	expression tag	UNP Q0TUP9
А	-6	HIS	-	expression tag	UNP Q0TUP9
A	-5	HIS	-	expression tag	UNP Q0TUP9
A	-4	HIS	-	expression tag	UNP Q0TUP9
A	-3	HIS	-	expression tag	UNP Q0TUP9
A	-2	HIS	-	expression tag	UNP Q0TUP9
А	-1	HIS	-	expression tag	UNP Q0TUP9
A	0	HIS	-	expression tag	UNP Q0TUP9
В	-8	GLY	-	expression tag	UNP Q0TUP9
В	-7	SER	-	expression tag	UNP Q0TUP9
В	-6	HIS	-	expression tag	UNP Q0TUP9
В	-5	HIS	-	expression tag	UNP Q0TUP9
В	-4	HIS	-	expression tag	UNP Q0TUP9
В	-3	HIS	-	expression tag	UNP Q0TUP9
В	-2	HIS	-	expression tag	UNP Q0TUP9
В	-1	HIS	-	expression tag	UNP Q0TUP9
В	0	HIS	-	expression tag	UNP Q0TUP9

There are 18 discrepancies between the modelled and reference sequences:

• Molecule 2 is a protein called PUTATIVE N-ACETYLMANNOSAMINE-6-PHOSPHATE 2-EPIMERASE.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
0	C	229	Total	С	Ν	0	\mathbf{S}	0	2	0
	U	229	1769	1111	304	342	12	0	5	0
0	л	229	Total	С	Ν	0	S	0	1	0
	D	229	1778	1117	307	342	12	0	4	0

There are 18 discrepancies between the modelled and reference sequences:

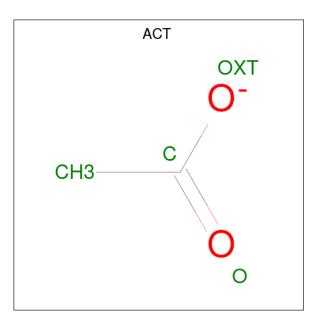
Chain	Residue	Modelled	Actual	Comment	Reference
С	-8	GLY	-	expression tag	UNP Q0TUP9
С	-7	SER	-	expression tag	UNP Q0TUP9
С	-6	SER	-	expression tag	UNP Q0TUP9
С	-5	HIS	-	expression tag	UNP Q0TUP9
С	-4	HIS	-	expression tag	UNP Q0TUP9
С	-3	HIS	-	expression tag	UNP Q0TUP9
С	-2	HIS	-	expression tag	UNP Q0TUP9
С	-1	HIS	-	expression tag	UNP Q0TUP9
С	0	HIS	-	expression tag	UNP Q0TUP9
D	-8	GLY	-	expression tag	UNP Q0TUP9
D	-7	SER	-	expression tag	UNP Q0TUP9
D	-6	SER	-	expression tag	UNP Q0TUP9
D	-5	HIS	-	expression tag	UNP Q0TUP9
D	-4	HIS	-	expression tag	UNP Q0TUP9
D	-3	HIS	-	expression tag	UNP Q0TUP9
D	-2	HIS	-	expression tag	UNP Q0TUP9
D	-1	HIS	-	expression tag	UNP Q0TUP9
D	0	HIS	-	expression tag	UNP Q0TUP9

• Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	2	Total Cl 2 2	0	0
3	В	2	Total Cl 2 2	0	0
3	С	2	Total Cl 2 2	0	0
3	D	2	Total Cl 2 2	0	0

• Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 4 & 2 & 2 \end{array}$	0	0
4	В	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 4 & 2 & 2 \end{array}$	0	0
4	С	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 4 & 2 & 2 \end{array}$	0	0
4	D	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 4 & 2 & 2 \end{array}$	0	0

• Molecule 5 is water.

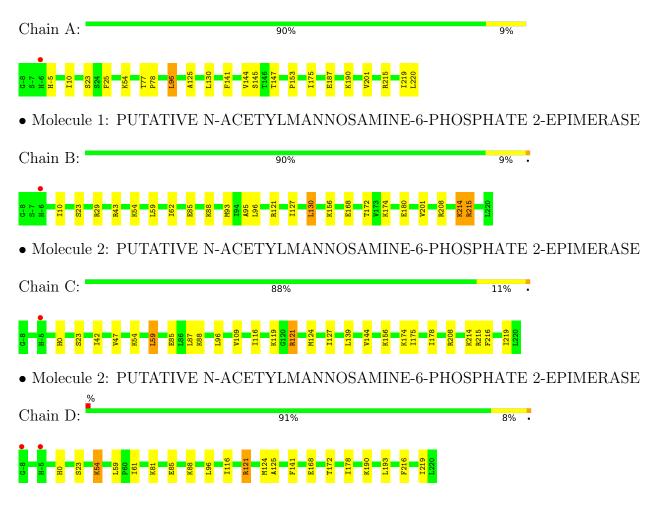
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	307	Total O 308 308	0	1
5	В	284	Total O 284 284	0	0
5	С	266	Total O 266 266	0	0
5	D	257	Total O 257 257	0	0



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: PUTATIVE N-ACETYLMANNOSAMINE-6-PHOSPHATE 2-EPIMERASE





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants	36.99Å 75.50Å 82.23Å	Deneiten
a, b, c, α , β , γ	90.22° 90.01° 92.83°	Depositor
Resolution (Å)	82.20 - 1.71	Depositor
Resolution (A)	41.12 - 1.71	EDS
% Data completeness	95.5 (82.20-1.71)	Depositor
(in resolution range)	94.9 (41.12-1.71)	EDS
R _{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.94 (at 1.71 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
D D	0.151 , 0.185	Depositor
R, R_{free}	0.180 , 0.190	DCC
R_{free} test set	4573 reflections $(4.99%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	15.1	Xtriage
Anisotropy	0.083	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.35, 25.6	EDS
L-test for twinning ²	$< L > = 0.49, < L^2 > = 0.32$	Xtriage
	0.032 for h,-k,-l	
Estimated twinning fraction	0.031 for -h,k,-l	Xtriage
	0.476 for -h,-k,l	
F_o, F_c correlation	0.95	EDS
Total number of atoms	8252	wwPDB-VP
Average B, all atoms $(Å^2)$	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.98% of the height of the origin peak. No significant pseudotranslation is detected.

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.



¹Intensities estimated from amplitudes.

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: ACT, CL

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	Bo	Bond lengths		nd angles
	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	А	0.73	1/1819~(0.1%)	0.78	0/2456
1	В	0.73	0/1824	0.80	4/2461~(0.2%)
2	С	0.70	0/1802	0.76	3/2431~(0.1%)
2	D	0.67	0/1814	0.74	2/2446~(0.1%)
All	All	0.71	1/7259~(0.0%)	0.77	9/9794~(0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	А	187	GLU	CG-CD	5.20	1.59	1.51

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

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	С	208	ARG	NE-CZ-NH1	6.75	123.68	120.30
2	С	121	ARG	NE-CZ-NH2	-6.58	117.01	120.30
1	В	208	ARG	NE-CZ-NH2	-6.44	117.08	120.30
2	С	121	ARG	NE-CZ-NH1	5.95	123.28	120.30
2	D	121	ARG	NE-CZ-NH1	5.73	123.16	120.30

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	А	1782	0	1822	16	0
1	В	1784	0	1840	26	0
2	С	1769	0	1819	18	0
2	D	1778	0	1834	16	0
3	А	2	0	0	0	0
3	В	2	0	0	0	0
3	С	2	0	0	0	0
3	D	2	0	0	0	0
4	А	4	0	3	0	0
4	В	4	0	3	0	0
4	С	4	0	3	0	0
4	D	4	0	3	0	0
5	А	308	0	0	6	0
5	В	284	0	0	5	0
5	С	266	0	0	6	0
5	D	257	0	0	1	0
All	All	8252	0	7327	68	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 5.

The worst 5 of 68 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)
5:A:2080:HOH:O	1:B:172:THR:HG21	1.58	1.03
1:A:220:LEU:HD22	2:D:193[B]:LEU:HD22	1.60	0.84
5:A:2081:HOH:O	1:B:130[A]:LEU:HD11	1.77	0.82
5:A:2083:HOH:O	1:B:130[A]:LEU:HD12	1.81	0.80
1:A:220:LEU:HD22	2:D:193[A]:LEU:CD1	2.16	0.76

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	Perce	ntiles
1	А	232/229~(101%)	227~(98%)	4(2%)	1 (0%)	34	18
1	В	232/229~(101%)	228~(98%)	3~(1%)	1 (0%)	34	18
2	\mathbf{C}	230/229~(100%)	225~(98%)	4(2%)	1 (0%)	34	18
2	D	231/229~(101%)	229~(99%)	1 (0%)	1 (0%)	34	18
All	All	925/916~(101%)	909~(98%)	12 (1%)	4 (0%)	34	18

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	23	SER
2	С	23	SER
2	D	23	SER
1	В	23	SER

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain 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	А	199/194~(103%)	195~(98%)	4 (2%)	55 37	
1	В	199/194~(103%)	191~(96%)	8 (4%)	31 12	
2	С	197/194~(102%)	193~(98%)	4 (2%)	55 37	
2	D	198/194~(102%)	195~(98%)	3~(2%)	65 49	
All	All	793/776~(102%)	774~(98%)	19 (2%)	50 29	

5 of 19 residues with a non-rotameric side chain are listed below:

Mol	Chain	Res	Type
2	С	96	LEU
2	D	59	LEU
2	D	96	LEU
2	D	54	LYS
1	В	130[B]	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11



such sidechains are listed below:

Mol	Chain	\mathbf{Res}	Type
2	D	8	ASN
2	D	108	ASN
2	D	184	ASN
2	D	157	GLN
2	С	184	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 8 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).

Mal	Mol Type Chain Res		Chain Res Link		B	Bond lengths			Bond angles		
INIOI	Type Chain Re	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2		
4	ACT	D	1223	-	$3,\!3,\!3$	1.57	1 (33%)	$3,\!3,\!3$	2.73	2 (66%)	
4	ACT	С	1223	-	$3,\!3,\!3$	1.90	1 (33%)	$3,\!3,\!3$	2.54	1 (33%)	
4	ACT	В	1223	-	$3,\!3,\!3$	1.10	0	$3,\!3,\!3$	3.05	2 (66%)	
4	ACT	А	1223	-	$3,\!3,\!3$	0.91	0	$3,\!3,\!3$	1.88	2 (66%)	

All (2) bond length outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	С	1223	ACT	OXT-C	-3.24	1.15	1.30
4	D	1223	ACT	OXT-C	-2.71	1.17	1.30

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
4	D	1223	ACT	OXT-C-O	-3.93	107.58	122.05
4	В	1223	ACT	OXT-C-CH3	3.92	131.37	115.18
4	С	1223	ACT	OXT-C-O	-3.64	108.63	122.05
4	В	1223	ACT	OXT-C-O	-3.44	109.38	122.05
4	А	1223	ACT	OXT-C-O	-2.53	112.74	122.05

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, 95^{th} 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	$\langle RSRZ \rangle$	> #RSRZ>2		$OWAB(Å^2)$	Q<0.9
1	А	229/229~(100%)	-0.27	1 (0%) 92	93	13, 19, 32, 42	0
1	В	229/229~(100%)	-0.28	1 (0%) 92	93	12, 19, 32, 43	0
2	С	229/229~(100%)	-0.22	1 (0%) 92	93	14, 22, 34, 46	0
2	D	229/229~(100%)	-0.26	2 (0%) 84	87	13, 22, 33, 43	0
All	All	916/916 (100%)	-0.26	5 (0%) 91	92	12, 20, 33, 46	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	-8	GLY	4.4
1	А	-6	HIS	3.9
1	В	-6	HIS	3.4
2	D	-5	HIS	2.4
2	С	-5	HIS	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 monosaccharides 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,



Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	ACT	А	1223	4/4	0.90	0.19	$25,\!31,\!32,\!33$	0
4	ACT	С	1223	4/4	0.92	0.13	24,26,28,31	0
4	ACT	D	1223	4/4	0.92	0.17	25,31,31,32	0
4	ACT	В	1223	4/4	0.94	0.16	21,26,28,28	0
3	CL	С	1221	1/1	0.98	0.09	27,27,27,27	0
3	CL	В	1222	1/1	0.99	0.04	16,16,16,16	0
3	CL	В	1221	1/1	0.99	0.04	23,23,23,23	0
3	CL	D	1221	1/1	0.99	0.05	26,26,26,26	0
3	CL	D	1222	1/1	0.99	0.05	16,16,16,16	0
3	CL	А	1222	1/1	1.00	0.04	16,16,16,16	0
3	CL	А	1221	1/1	1.00	0.03	23,23,23,23	0
3	CL	С	1222	1/1	1.00	0.04	$15,\!15,\!15,\!15$	0

median, 95^{th} 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.

6.5 Other polymers (i)

There are no such residues in this entry.

