



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:33 AM GMT

PDB ID : 2R8V
Title : Native structure of N-acetylglutamate synthase from *Neisseria gonorrhoeae*
Authors : Shi, D.; Sagar, V.; Jin, Z.; Yu, X.; Caldovic, L.; Morizono, H.; Allewell, N.M.;
Tuchman, M.
Deposited on : 2007-09-11
Resolution : 2.50 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

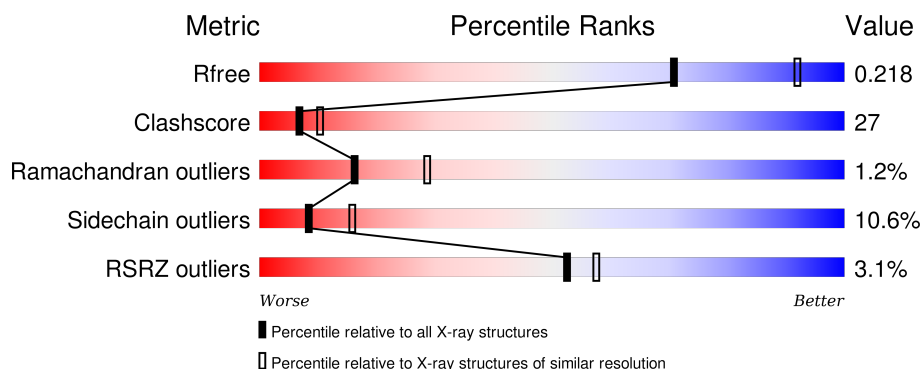
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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}	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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. 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	456	<div> <div>3%</div> <div>57%</div> <div>30%</div> <div>5%</div> <div>7%</div> </div>

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
2	ACO	A	601	X	-	-	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3395 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.

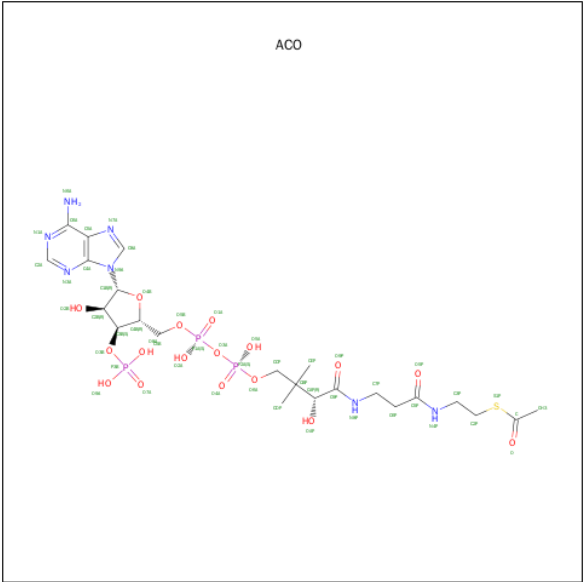
- Molecule 1 is a protein called Putative acetylglutamate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	424	Total	C	N	O	S	0	0	0
			3227	2012	595	611	9			

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q5FAK7
A	-18	GLY	-	EXPRESSION TAG	UNP Q5FAK7
A	-17	SER	-	EXPRESSION TAG	UNP Q5FAK7
A	-16	SER	-	EXPRESSION TAG	UNP Q5FAK7
A	-15	HIS	-	EXPRESSION TAG	UNP Q5FAK7
A	-14	HIS	-	EXPRESSION TAG	UNP Q5FAK7
A	-13	HIS	-	EXPRESSION TAG	UNP Q5FAK7
A	-12	HIS	-	EXPRESSION TAG	UNP Q5FAK7
A	-11	HIS	-	EXPRESSION TAG	UNP Q5FAK7
A	-10	HIS	-	EXPRESSION TAG	UNP Q5FAK7
A	-9	SER	-	EXPRESSION TAG	UNP Q5FAK7
A	-8	SER	-	EXPRESSION TAG	UNP Q5FAK7
A	-7	GLY	-	EXPRESSION TAG	UNP Q5FAK7
A	-6	LEU	-	EXPRESSION TAG	UNP Q5FAK7
A	-5	VAL	-	EXPRESSION TAG	UNP Q5FAK7
A	-4	PRO	-	EXPRESSION TAG	UNP Q5FAK7
A	-3	ARG	-	EXPRESSION TAG	UNP Q5FAK7
A	-2	GLY	-	EXPRESSION TAG	UNP Q5FAK7
A	-1	SER	-	EXPRESSION TAG	UNP Q5FAK7
A	0	HIS	-	EXPRESSION TAG	UNP Q5FAK7
A	312	ILE	VAL	ENGINEERED	UNP Q5FAK7
A	336	ASN	ASP	ENGINEERED	UNP Q5FAK7
A	427	SER	PRO	ENGINEERED	UNP Q5FAK7

- Molecule 2 is ACETYL COENZYME *A (three-letter code: ACO) (formula: C₂₃H₃₈N₇O₁₇P₃S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	51	23	7	17	3	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	117	Total	O	0	0
			117	117		

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ($\text{RSRZ} > 2$). 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.

Chain A:

The visualization displays amino acid conservation across four protein chains (A, B, C, D) and a corresponding sequence logo. The top bar chart shows the overall conservation levels: 57% (green), 30% (yellow), 5% (orange), and 7% (red). The subsequent rows show specific amino acids and their conservation status, indicated by colored dots. The bottom row shows the sequence logo for Chain A, with letters colored by their information content.

Amino Acid	Conservation Level
MET	Low
GLY	High
SER	High
SER	High
HIS	High
HIS	High
HIS	High
HIS	High
HIS	High
HIS	High
HIS	High
SER	High
SER	High
GLY	High
LEU	High
VAL	High
PRO	High
ARG	High
GLY	High
SER	High
HIS	High
MET	High
ASN	High
ALA	High
PRO	High
P5	High
V6	High
F11	High
A12	High
E13	High
A14	High
A15	High
P16	High
G11	High
G12	High
S113	High
VAL	High
SER	High
GLY	High
PHE	High
ALA	High
ALA	High
ARG	High
ALA	High
PRO	High
V123	High
G32	High
R33	High
I34	High
L35	High
G38	High
T39	High
L40	High
K41	High
K42	High
L43	High
G48	High
L49	High
L50	High
S51	High
Q52	High
L53	High
I60	High
H61	High
Y64	High
H65	High
F66	High
L67	High
D68	High
R69	High
L70	High
Q74	High
G75	High
R76	High
H79	High
Y80	High
C81	High
R82	High
G83	High
L84	High
S91	High
L92	High
T101	High
V102	High
R103	High
L110	High
G111	High
G112	High
S113	High
VAL	High
SER	High
GLY	High
PHE	High
ALA	High
ALA	High
ARG	High
ALA	High
PRO	High
V123	High
S122	High
P124	High
S127	High
I136	High
M144	High
E145	High
Y146	High
G148	High
R151	High
L152	High
T153	High
D154	High
T155	High
L158	High
R159	High
N166	High
M170	High
P171	High
P172	High
L173	High
S176	High
F182	High
D185	High
Q188	High
K200	High
E201	High
L202	High
L205	High
T206	High
L207	High
R213	High
P214	High
D215	High
G216	High
T217	High
E220	High
S223	High
A224	High
Q225	High
E226	High
A227	High
Q228	High
A231	High
E232	High
S236	High
E237	High
T238	High
R239	High
S243	High
V246	High
L249	High
E250	High
V253	High
H254	High
R255	High
V256	High
L258	High
L259	High
G265	High
S266	High
L267	High
L268	High
Q269	High
E270	High
R274	High
V275	High
G276	High
L277	High
G278	High
T279	High
S280	High
A285	High
F286	High
S288	High
T289	High
R290	High
Q291	High
A292	High
H293	High
S294	High
L297	High
A301	High
L304	High
R305	High
E308	High
G311	High
L312	High
L313	High
L314	High
R316	High
S317	High
R318	High
E319	High
V320	High
L321	High
E322	High
N323	High
H324	High
L330	High
L337	High
Y338	High
G339	High
C340	High
K344	High
D350	High
E353	High
L354	High
S360	High
P361	High
Q362	High
A363	High
G364	High
D365	High
R371	High
L372	High
L373	High
A374	High
H375	High
L376	High
L377	High
D378	High
L3	

4 Data and refinement statistics

Property	Value	Source
Space group	P 3 1 2	Depositor
Cell constants a, b, c, α , β , γ	98.66 Å 98.66 Å 89.75 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.50 19.87 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.9 (20.00-2.50) 99.1 (19.87-2.50)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.55 (at 2.50 Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.192 , 0.252 0.191 , 0.218	Depositor DCC
R_{free} test set	835 reflections (5.08%)	DCC
Wilson B-factor (Å ²)	54.3	Xtriage
Anisotropy	0.346	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 49.5	EDS
Estimated twinning fraction	0.053 for -h,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 17270 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	3395	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.56% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: ACO

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	0/3280	1.01	6/4433 (0.1%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	402	ARG	NE-CZ-NH2	-7.46	116.57	120.30
1	A	259	LEU	CB-CG-CD2	-6.73	99.55	111.00
1	A	110	LEU	CA-CB-CG	5.51	127.97	115.30
1	A	425	ARG	NE-CZ-NH1	-5.40	117.60	120.30
1	A	431	VAL	CB-CA-C	-5.17	101.58	111.40

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	3227	0	3208	172	1
2	A	51	0	34	12	0
3	A	117	0	0	27	1
All	All	3395	0	3242	175	1

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 27.

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:A:304:ILE:HD11	1:A:313:LEU:CD1	1.55	1.35
1:A:304:ILE:CD1	1:A:313:LEU:HD12	1.78	1.12
1:A:316:ARG:HG2	1:A:317:SER:N	1.57	1.11
1:A:82:ARG:HH11	1:A:82:ARG:HG2	0.94	1.06
1:A:239:ARG:NH2	1:A:243:SER:HB2	1.71	1.04

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:316:ARG:NH1	3:A:501:HOH:O[5_655]	2.16	0.04

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	420/456 (92%)	394 (94%)	21 (5%)	5 (1%)	16 29

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	287	VAL
1	A	82	ARG
1	A	145	GLU
1	A	214	PRO
1	A	277	ILE

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	330/355 (93%)	295 (89%)	35 (11%)	8 16

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	238	THR
1	A	286	PHE
1	A	418	LYS
1	A	267	LEU
1	A	268	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	225	GLN
1	A	228	GLN
1	A	364	GLN
1	A	175	HIS
1	A	362	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry

1 ligand is modelled in this entry.

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	ACO	A	601	-	43,53,53	1.05	2 (4%)	55,79,79	2.42	14 (25%)

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
2	ACO	A	601	-	1/1/12/14	0/47/67/67	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	ACO	CH3-C	-2.71	1.38	1.50
2	A	601	ACO	P3B-O7A	3.49	1.62	1.51

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	601	ACO	N3A-C2A-N1A	-10.30	121.01	128.89
2	A	601	ACO	C2B-C1B-N9A	-8.76	100.91	114.29
2	A	601	ACO	C2P-C3P-N4P	-4.14	104.09	112.36
2	A	601	ACO	C7P-C6P-C5P	-4.02	105.69	112.31
2	A	601	ACO	C4B-O4B-C1B	-3.34	106.05	109.72

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	601	ACO	CAP

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	ACO	12	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	424/456 (92%)	0.22	13 (3%) 52 57	8, 23, 55, 71	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	436	ARG	4.9
1	A	113	SER	4.1
1	A	339	GLY	3.7
1	A	274	ARG	3.0
1	A	340	CYS	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 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
2	ACO	A	601	51/51	0.98	0.15	-0.63	20,33,48,52	0

6.5 Other polymers [i](#)

There are no such residues in this entry.