wwPDB X-ray Structure Validation Summary Report (i)

May 13, 2020 – 12:38 am BST

PDB ID	:	4KT6
Title	:	High-resolution crystal structure Streptococcus pyogenes beta-NAD+ glycohydrolase in complex with its endogenous inhibitor IFS reveals a water-rich interface
Authors	:	Yoon, J.Y.; An, D.R.; Yoon, H.-J.; Kim, H.S.; Lee, S.J.; Im, H.N.; Jang, J.Y.; Suh, S.W.
Deposited on	:	2013-05-20
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 following versions of software and data (see references (1)) were used in the production of this report:

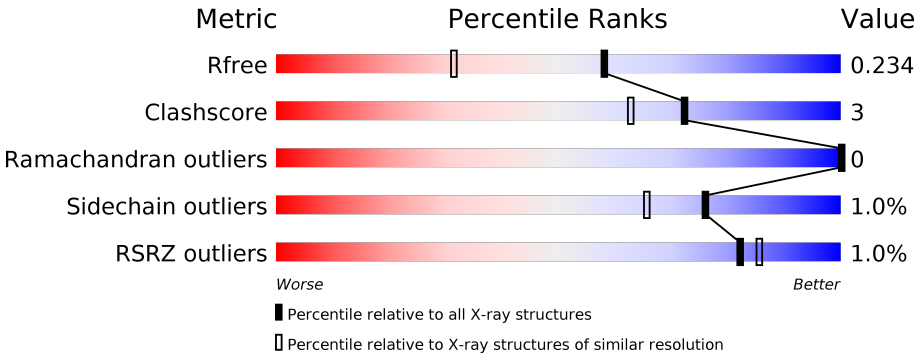
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
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 on 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	259	90% 8%
1	C	259	2% 89% 8%
2	B	161	% 93% 7%
2	D	161	% 96%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7304 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 Nicotine adenine dinucleotide glycohydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	254	Total	C	N	O	S	0	0	0
			2032	1285	342	398	7			
1	C	254	Total	C	N	O	S	0	0	0
			2032	1285	342	398	7			

- Molecule 2 is a protein called Putative uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	161	Total	C	N	O	S	0	0	0
			1322	847	215	252	8			
2	D	161	Total	C	N	O	S	0	0	0
			1322	847	215	252	8			

- Molecule 3 is water.

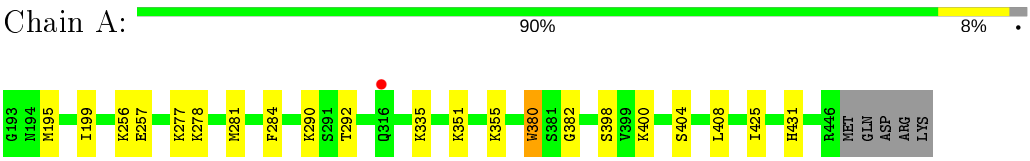
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	196	Total	O	0	0
			196	196		
3	B	128	Total	O	0	0
			128	128		
3	C	154	Total	O	0	0
			154	154		
3	D	118	Total	O	0	0
			118	118		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

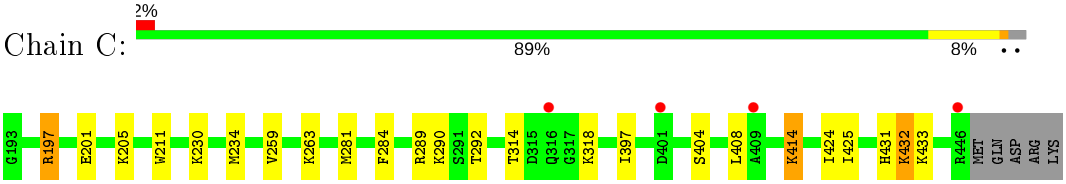
- Molecule 1: Nicotine adenine dinucleotide glycohydrolase

Chain A:



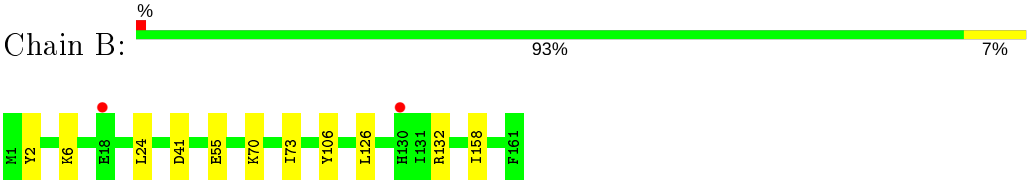
- Molecule 1: Nicotine adenine dinucleotide glycohydrolase

Chain C:



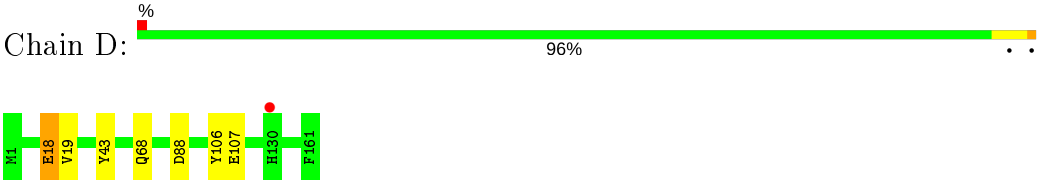
- Molecule 2: Putative uncharacterized protein

Chain B:



- Molecule 2: Putative uncharacterized protein

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	43.20Å 56.88Å 89.98Å 72.96° 90.01° 82.27°	Depositor
Resolution (Å)	20.00 – 1.71 19.98 – 1.71	Depositor EDS
% Data completeness (in resolution range)	93.5 (20.00-1.71) 93.6 (19.98-1.71)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.16 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.197 , 0.235 0.198 , 0.234	Depositor DCC
R_{free} test set	4140 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	24.6	Xtriage
Anisotropy	0.014	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 50.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.009 for -h,-k,-k+l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7304	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.43% 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

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	0.63	1/2068 (0.0%)	0.70	0/2770
1	C	0.60	1/2068 (0.0%)	0.64	0/2770
2	B	0.78	0/1348	0.72	0/1817
2	D	0.76	0/1348	0.65	0/1817
All	All	0.68	2/6832 (0.0%)	0.68	0/9174

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	211	TRP	CD2-CE2	5.38	1.47	1.41
1	A	380	TRP	CD2-CE2	5.24	1.47	1.41

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	2032	0	2021	12	0
1	C	2032	0	2021	21	0
2	B	1322	0	1310	12	0
2	D	1322	0	1310	4	0
3	A	196	0	0	0	0
3	B	128	0	0	0	0
3	C	154	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	118	0	0	0	0
All	All	7304	0	6662	46	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 3.

The worst 5 of 46 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:C:197:ARG:CG	1:C:197:ARG:HH11	1.56	1.15
1:C:197:ARG:HH11	1:C:197:ARG:HG2	1.20	1.03
1:C:197:ARG:CG	1:C:197:ARG:NH1	2.28	0.87
1:C:197:ARG:HH11	1:C:197:ARG:HG3	1.39	0.86
1:C:197:ARG:NH1	1:C:197:ARG:HG2	1.84	0.79

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	252/259 (97%)	249 (99%)	3 (1%)	0	100	100
1	C	252/259 (97%)	250 (99%)	2 (1%)	0	100	100
2	B	159/161 (99%)	157 (99%)	2 (1%)	0	100	100
2	D	159/161 (99%)	155 (98%)	4 (2%)	0	100	100
All	All	822/840 (98%)	811 (99%)	11 (1%)	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	219/224 (98%)	217 (99%)	2 (1%)	78 69
1	C	219/224 (98%)	216 (99%)	3 (1%)	67 52
2	B	142/142 (100%)	142 (100%)	0	100 100
2	D	142/142 (100%)	140 (99%)	2 (1%)	67 52
All	All	722/732 (99%)	715 (99%)	7 (1%)	76 65

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	414	LYS
2	D	88	ASP
1	C	432	LYS
1	A	257	GLU
2	D	18	GLU

Some 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 carbohydrates 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

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, 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	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	254/259 (98%)	-0.10	1 (0%)	92	93	17, 26, 41, 53	0
1	C	254/259 (98%)	-0.01	4 (1%)	72	76	20, 29, 45, 60	0
2	B	161/161 (100%)	-0.12	2 (1%)	79	83	14, 24, 38, 56	0
2	D	161/161 (100%)	-0.05	1 (0%)	89	91	19, 27, 39, 56	0
All	All	830/840 (98%)	-0.07	8 (0%)	82	85	14, 27, 42, 60	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	446	ARG	3.5
1	C	316	GLN	3.2
1	A	316	GLN	2.9
1	C	409	ALA	2.8
2	B	130	HIS	2.7

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 carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.