WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 22, 2024 – 02:58 PM EDT

|		
PDB ID	:	6MG6
Title	:	Crystal structure of carbon-nitrogen hydrolase from Helicobacter pylori G27
Authors	:	Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on	:	2018-09-13
Resolution	:	2.10 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 2.10 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.
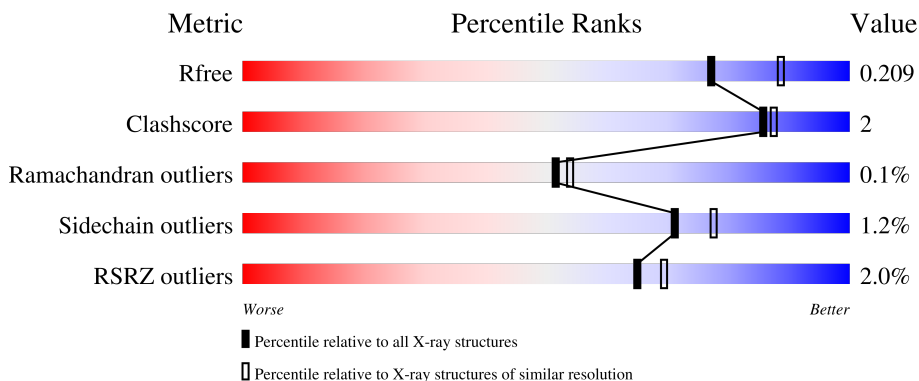
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	A	300	92% 6% ·
1	B	300	91% 7% ·
1	C	300	% 91% 7% ·
1	D	300	7% 91% 6% ·

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9537 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 Carbon-nitrogen hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	294	Total	C	N	O	S	0	0	0
			2338	1514	390	426	8			
1	B	293	Total	C	N	O	S	0	1	0
			2331	1511	385	427	8			
1	C	293	Total	C	N	O	S	0	0	0
			2281	1481	372	421	7			
1	D	292	Total	C	N	O	S	0	1	0
			2189	1406	370	406	7			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	expression tag	UNP B5Z7B9
A	-6	ALA	-	expression tag	UNP B5Z7B9
A	-5	HIS	-	expression tag	UNP B5Z7B9
A	-4	HIS	-	expression tag	UNP B5Z7B9
A	-3	HIS	-	expression tag	UNP B5Z7B9
A	-2	HIS	-	expression tag	UNP B5Z7B9
A	-1	HIS	-	expression tag	UNP B5Z7B9
A	0	HIS	-	expression tag	UNP B5Z7B9
B	-7	MET	-	expression tag	UNP B5Z7B9
B	-6	ALA	-	expression tag	UNP B5Z7B9
B	-5	HIS	-	expression tag	UNP B5Z7B9
B	-4	HIS	-	expression tag	UNP B5Z7B9
B	-3	HIS	-	expression tag	UNP B5Z7B9
B	-2	HIS	-	expression tag	UNP B5Z7B9
B	-1	HIS	-	expression tag	UNP B5Z7B9
B	0	HIS	-	expression tag	UNP B5Z7B9
C	-7	MET	-	expression tag	UNP B5Z7B9
C	-6	ALA	-	expression tag	UNP B5Z7B9
C	-5	HIS	-	expression tag	UNP B5Z7B9
C	-4	HIS	-	expression tag	UNP B5Z7B9
C	-3	HIS	-	expression tag	UNP B5Z7B9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	HIS	-	expression tag	UNP B5Z7B9
C	-1	HIS	-	expression tag	UNP B5Z7B9
C	0	HIS	-	expression tag	UNP B5Z7B9
D	-7	MET	-	expression tag	UNP B5Z7B9
D	-6	ALA	-	expression tag	UNP B5Z7B9
D	-5	HIS	-	expression tag	UNP B5Z7B9
D	-4	HIS	-	expression tag	UNP B5Z7B9
D	-3	HIS	-	expression tag	UNP B5Z7B9
D	-2	HIS	-	expression tag	UNP B5Z7B9
D	-1	HIS	-	expression tag	UNP B5Z7B9
D	0	HIS	-	expression tag	UNP B5Z7B9

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
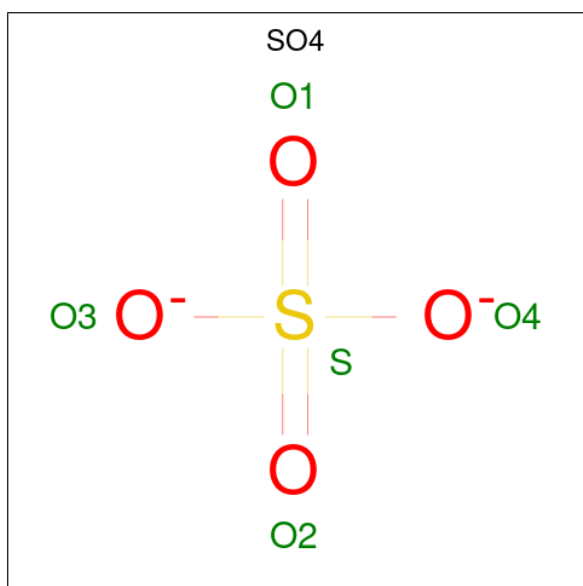
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	140	140	140	0	0
3	B	124	124	124	0	0
3	C	89	89	89	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	40	Total O 40 40		0	0

3 Residue-property plots ⓘ

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: Carbon-nitrogen hydrolase

Chain A:

• Molecule 1: Carbon-nitrogen hydrolase

Chain B:

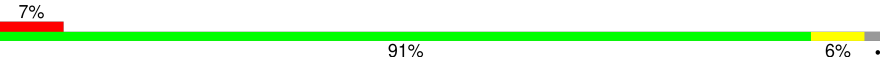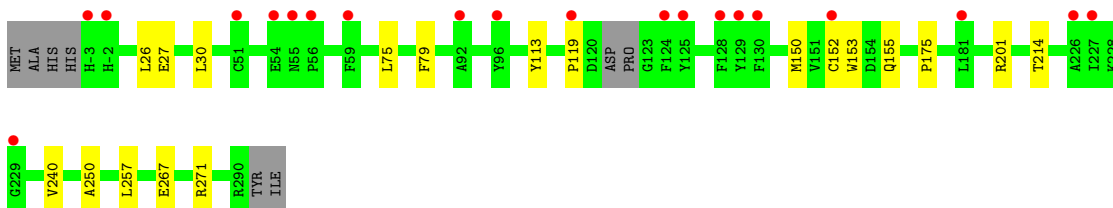• Molecule 1: Carbon-nitrogen hydrolase

Chain C:

• Molecule 1: Carbon-nitrogen hydrolase

Chain D:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	137.58Å 91.38Å 95.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.58 – 2.10 47.58 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.4 (47.58-2.10) 99.4 (47.58-2.10)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.15 (at 2.10Å)	Xtriage
Refinement program	PHENIX 1.14_3247	Depositor
R, R_{free}	0.168 , 0.208 0.171 , 0.209	Depositor DCC
R_{free} test set	2045 reflections (2.91%)	wwPDB-VP
Wilson B-factor (Å^2)	42.2	Xtriage
Anisotropy	0.207	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 47.5	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.016 for -h,l,k	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	9537	wwPDB-VP
Average B, all atoms (Å^2)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.53% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.44	0/2398	0.59	1/3244 (0.0%)
1	B	0.43	0/2393	0.58	0/3238
1	C	0.37	0/2340	0.55	0/3176
1	D	0.31	0/2248	0.50	0/3066
All	All	0.39	0/9379	0.56	1/12724 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	30	LEU	CA-CB-CG	-5.48	102.69	115.30

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	A	2338	0	2270	11	0
1	B	2331	0	2267	12	0
1	C	2281	0	2171	10	0
1	D	2189	0	1978	9	0
2	A	5	0	0	0	0
3	A	140	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	124	0	0	1	0
3	C	89	0	0	0	0
3	D	40	0	0	0	0
All	All	9537	0	8686	40	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 2.

The worst 5 of 40 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:B:292:ILE:OXT	3:B:301:HOH:O	2.01	0.78
1:D:30:LEU:HD22	1:D:79:PHE:HB3	1.74	0.69
1:A:106:ASP:OD2	3:A:501:HOH:O	2.11	0.69
1:B:30:LEU:HD22	1:B:79:PHE:HB3	1.76	0.66
1:D:26:LEU:HD23	1:D:75:LEU:HD22	1.85	0.57

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 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	292/300 (97%)	286 (98%)	6 (2%)	0	100	100
1	B	292/300 (97%)	287 (98%)	5 (2%)	0	100	100
1	C	291/300 (97%)	286 (98%)	5 (2%)	0	100	100
1	D	289/300 (96%)	285 (99%)	3 (1%)	1 (0%)	41	41
All	All	1164/1200 (97%)	1144 (98%)	19 (2%)	1 (0%)	51	54

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	152	CYS

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	238/252 (94%)	235 (99%)	3 (1%)	69	75
1	B	238/252 (94%)	235 (99%)	3 (1%)	69	75
1	C	227/252 (90%)	222 (98%)	5 (2%)	52	57
1	D	203/252 (81%)	202 (100%)	1 (0%)	88	92
All	All	906/1008 (90%)	894 (99%)	12 (1%)	71	75

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	105	LYS
1	C	121	ASP
1	D	155	GLN
1	C	155	GLN
1	B	121[A]	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	77	GLN
1	A	186	ASN
1	C	186	ASN
1	D	155	GLN
1	D	200	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 monosaccharides 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	SO4	A	400	-	4,4,4	0.34	0	6,6,6	0.28	0

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.

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, 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(Å²)	Q<0.9
1	A	294/300 (98%)	-0.48	0	100	100	30, 43, 65, 96	0
1	B	293/300 (97%)	-0.33	0	100	100	29, 46, 69, 87	0
1	C	293/300 (97%)	-0.34	3 (1%)	82	85	35, 57, 87, 106	0
1	D	292/300 (97%)	0.14	20 (6%)	17	21	38, 74, 111, 128	0
All	All	1172/1200 (97%)	-0.25	23 (1%)	65	69	29, 53, 93, 128	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	130	PHE	4.7
1	D	181	LEU	4.3
1	D	229	GLY	3.9
1	D	125	TYR	3.8
1	D	129	TYR	3.7

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 monosaccharides in this entry.

6.4 Ligands ⓘ

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(Å^2)	Q<0.9
2	SO4	A	400	5/5	0.82	0.13	70,84,86,87	0

6.5 Other polymers ⓘ

There are no such residues in this entry.