wwPDB X-ray Structure Validation Summary Report (i)

Dec 5, 2023 – 01:41 am GMT

PDB ID	:	1E9Z
Title	:	Crystal structure of Helicobacter pylori urease
Authors	:	Ha, N.-C.; Oh, S.-T.; Oh, B.-H.
Deposited on	:	2000-11-01
Resolution	:	3.00 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 (i)) were used in the production of this report:

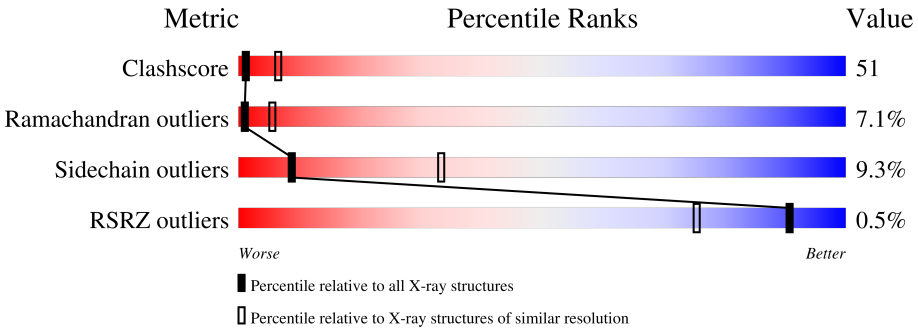
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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))
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	238	32% 59% 8%
2	B	569	% 32% 57% 10%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6278 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 UREASE SUBUNIT ALPHA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	238	Total	C	N	O	S	0	0	0
			1864	1175	333	348	8			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	170	ALA	SER	conflict	UNP P14916

- Molecule 2 is a protein called UREASE SUBUNIT BETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	569	Total	C	N	O	S	0	0	0
			4334	2716	747	850	21			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	219	KCX	LYS	modified residue	UNP P14917
B	355	ALA	ILE	conflict	UNP P69996
B	522	VAL	ILE	conflict	UNP P69996
B	531	ASN	ASP	conflict	UNP P69996

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Ni	0	0
			2	2		

- Molecule 4 is water.

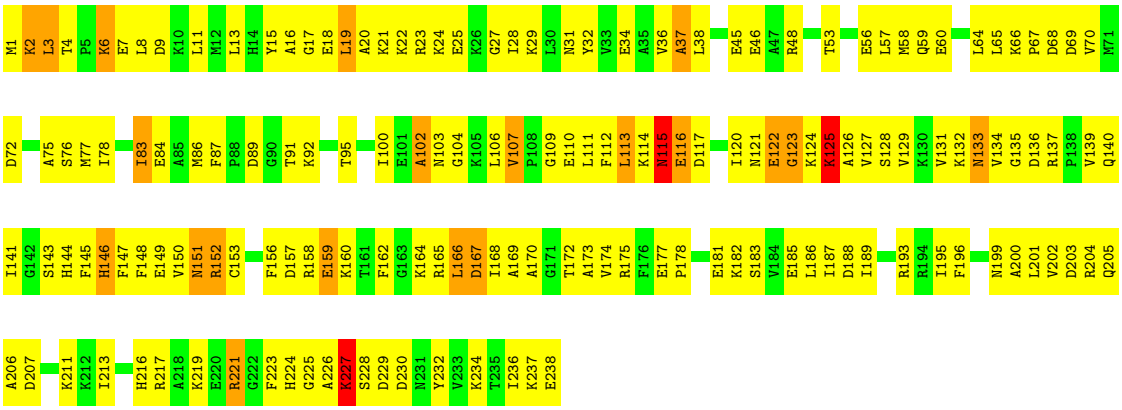
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	8	Total O 8 8	0	0
4	B	70	Total O 70 70	0	0

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.

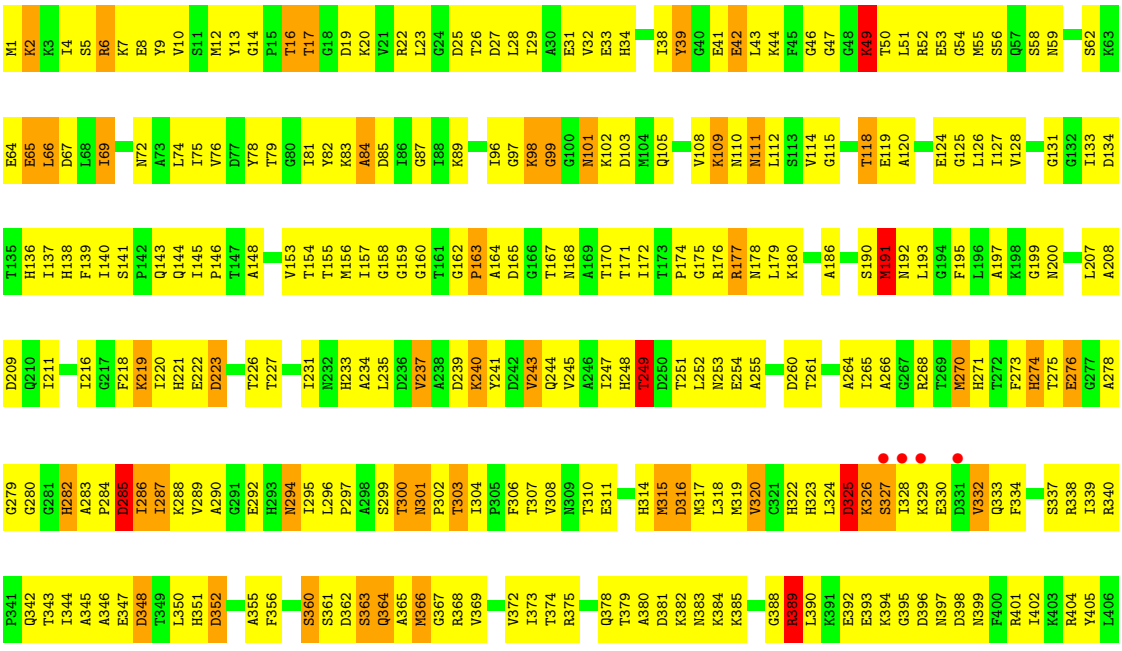
• Molecule 1: UREASE SUBUNIT ALPHA

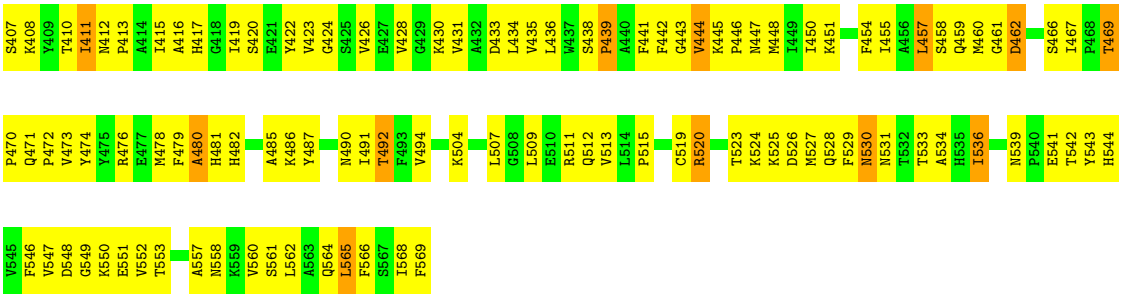
Chain A:



• Molecule 2: UREASE SUBUNIT BETA

Chain B:




S407 K408 Y409 T410 I411 N412 P413 A414 I415 A416 H417 G418 I419 S420 E421 Y422 V423 G424 S425 V426 E427 V428 G429 K430 V431 A432 D433 L434 V435 L436 W437 S438 P439 A440 F441 F442 G443 V444 K445 P446 M447 M448 I449 I450 K451 F454 I455 A456 L457 S458 Q459 M460 G461 D462 S466 I467 P468 T469
P470 Q471 P472 V473 Y474 Y475 R476 E477 M478 F479 A480 H481 H482 A485 K486 Y487 N490 I491 T492 F493 V494 K504 L507 G508 L509 E510 R511 Q512 V513 L514 P515 C519 R520 T523 K524 K525 D526 M527 Q528 F529 M530 N531 T532 T533 A534 H535 I536 N539 P540 E541 T542 Y543 H544
V545 F546 V547 D548 G549 K550 E551 V552 T553 A557 N558 K559 V560 S561 L562 A563 Q564 L565 F566 S567 I568 F569

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α, β, γ	177.91Å 177.91Å 177.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00 19.41 – 2.80	Depositor EDS
% Data completeness (in resolution range)	95.3 (20.00-3.00) 93.2 (19.41-2.80)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$ [1]	2.01 (at 2.79Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.212 , 0.284 0.205 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	24.1	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 18.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.062 for -l,-k,-h	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	6278	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.88% 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

Bond lengths and bond angles in the following residue types are not validated in this section: NI, KCX

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	0/1892	0.63	0/2532
2	B	0.47	0/4405	0.67	0/5958
All	All	0.47	0/6297	0.66	0/8490

There are no bond length outliers.

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	1864	0	1909	165	0
2	B	4334	0	4261	486	0
3	B	2	0	0	0	0
4	A	8	0	0	2	0
4	B	70	0	0	14	0
All	All	6278	0	6170	634	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 51.

The worst 5 of 634 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)
2:B:301:ASN:ND2	2:B:368:ARG:H	1.55	1.02
2:B:109:LYS:H	2:B:109:LYS:HE3	1.23	1.02
2:B:200:ASN:HD21	2:B:221:HIS:H	1.01	1.00
2:B:140:ILE:HD12	2:B:364:GLN:HB3	1.42	0.99
2:B:244:GLN:HE21	2:B:245:VAL:H	1.07	0.96

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	236/238 (99%)	183 (78%)	37 (16%)	16 (7%)	1	6
2	B	566/569 (100%)	428 (76%)	97 (17%)	41 (7%)	1	5
All	All	802/807 (99%)	611 (76%)	134 (17%)	57 (7%)	1	5

5 of 57 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	102	ALA
1	A	116	GLU
1	A	125	LYS
1	A	172	THR
1	A	227	LYS

5.3.2 Protein sidechains (i)

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	197/197 (100%)	181 (92%)	16 (8%)	11 40
2	B	458/458 (100%)	413 (90%)	45 (10%)	8 30
All	All	655/655 (100%)	594 (91%)	61 (9%)	9 33

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	118	THR
2	B	469	THR
2	B	240	LYS
2	B	457	LEU
2	B	536	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	244	GLN
2	B	496	GLN
2	B	333	GLN
2	B	518	ASN
2	B	417	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

1 non-standard protein/DNA/RNA residue 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	KCX	B	219	2,3	9,11,12	1.14	1 (11%)	5,12,14	1.15	1 (20%)

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	KCX	B	219	2,3	-	3/9/10/12	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	219	KCX	OQ1-CX	2.47	1.26	1.21

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	219	KCX	OQ1-CX-NZ	-2.06	121.77	124.96

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	219	KCX	OQ1-CX-NZ-CE
2	B	219	KCX	OQ2-CX-NZ-CE
2	B	219	KCX	CE-CD-CG-CB

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	219	KCX	4	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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 (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

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	238/238 (100%)	-0.79	0 100 100	6, 26, 47, 79	0
2	B	568/569 (99%)	-0.82	4 (0%) 87 69	1, 13, 55, 100	0
All	All	806/807 (99%)	-0.81	4 (0%) 91 75	1, 17, 50, 100	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	331	ASP	3.2
2	B	327	SER	2.2
2	B	329	LYS	2.1
2	B	328	ILE	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

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, 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.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	KCX	B	219	12/13	0.92	0.17	18,24,39,41	0

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, 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.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NI	B	3001	1/1	0.98	0.09	56,56,56,56	0
3	NI	B	3002	1/1	0.98	0.08	74,74,74,74	0

6.5 Other polymers

There are no such residues in this entry.