



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:46 PM GMT

PDB ID : 3UX4  
Title : Crystal structure of the urea channel from the human gastric pathogen Helicobacter pylori  
Authors : McNulty, R.; Luecke, H.  
Deposited on : 2011-12-03  
Resolution : 3.26 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](mailto: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

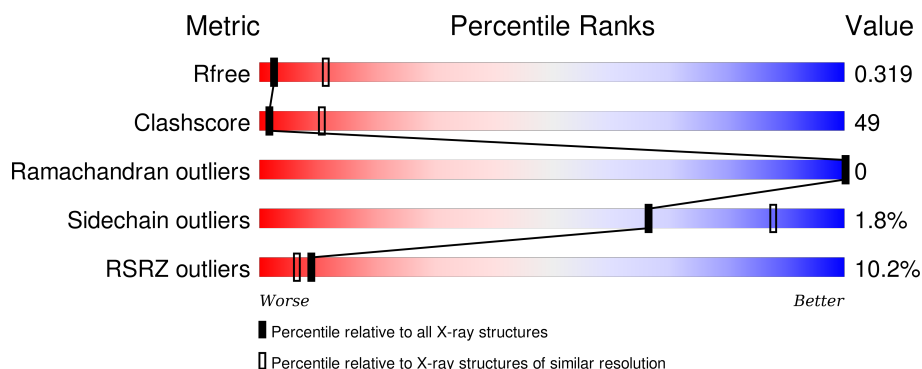
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.26 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	1624 (3.32-3.20)
Clashscore	102246	1806 (3.32-3.20)
Ramachandran outliers	100387	1773 (3.32-3.20)
Sidechain outliers	100360	1771 (3.32-3.20)
RSRZ outliers	91569	1632 (3.32-3.20)

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  $\geq 3$ , 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	201	<div> <div>8%</div> <div>36%</div> <div>51%</div> <div>10%</div> </div>
1	B	201	<div> <div>10%</div> <div>38%</div> <div>49%</div> <div>10%</div> </div>
1	C	201	<div> <div>8%</div> <div>38%</div> <div>50%</div> <div>10%</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	XP4	B	301	-	-	-	X
2	XP4	B	305	-	-	-	X
2	XP4	C	301	-	-	-	X
2	XP4	C	302	-	-	-	X

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4482 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.

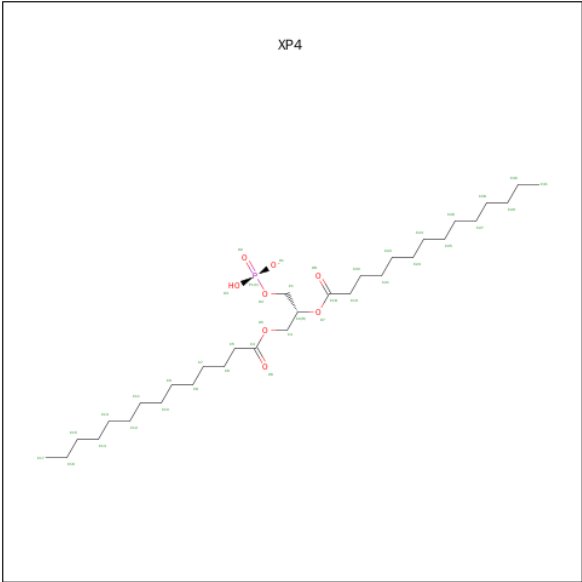
- Molecule 1 is a protein called Acid-activated urea channel.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	0	0	0
			1428	972	216	235	5			
1	B	180	Total	C	N	O	S	0	0	0
			1428	972	216	235	5			
1	C	180	Total	C	N	O	S	0	0	0
			1428	972	216	235	5			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	61A	HIS	-	EXPRESSION TAG	UNP P56874
A	61B	HIS	-	EXPRESSION TAG	UNP P56874
A	61C	HIS	-	EXPRESSION TAG	UNP P56874
A	61D	HIS	-	EXPRESSION TAG	UNP P56874
A	61E	HIS	-	EXPRESSION TAG	UNP P56874
A	61F	HIS	-	EXPRESSION TAG	UNP P56874
B	61A	HIS	-	EXPRESSION TAG	UNP P56874
B	61B	HIS	-	EXPRESSION TAG	UNP P56874
B	61C	HIS	-	EXPRESSION TAG	UNP P56874
B	61D	HIS	-	EXPRESSION TAG	UNP P56874
B	61E	HIS	-	EXPRESSION TAG	UNP P56874
B	61F	HIS	-	EXPRESSION TAG	UNP P56874
C	61A	HIS	-	EXPRESSION TAG	UNP P56874
C	61B	HIS	-	EXPRESSION TAG	UNP P56874
C	61C	HIS	-	EXPRESSION TAG	UNP P56874
C	61D	HIS	-	EXPRESSION TAG	UNP P56874
C	61E	HIS	-	EXPRESSION TAG	UNP P56874
C	61F	HIS	-	EXPRESSION TAG	UNP P56874

- Molecule 2 is 1,2-DIMYRISTOYL-SN-GLYCERO-3-PHOSPHATE (three-letter code: XP4) (formula: C<sub>31</sub>H<sub>60</sub>O<sub>8</sub>P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total C O P 40 31 8 1	0	0
2	B	1	Total C O 36 31 5	0	0
2	B	1	Total C O P 40 31 8 1	0	0
2	B	1	Total C O 18 16 2	0	0
2	B	1	Total C O 17 15 2	0	0
2	A	1	Total C O 17 15 2	0	0
2	C	1	Total C O 10 8 2	0	0
2	A	1	Total C O 10 8 2	0	0
2	B	1	Total C O 10 8 2	0	0



I144	I145	W146	W149	G150	V151	L152	W153	L154	T155	A156	F157	L158	E159	N160	I161	L162	K163	L164	P165	L166	G167	K168	F169	T170	P171	W172	L173	A174	I175	I176	E177	G178	I179	L180	T181	A182	W183	I184	P185	A186	W187	L188	L189	Q192	H193	W194	V195
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 4 <sub>2</sub> 2 <sub>1</sub> 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	122.90Å 122.90Å 141.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.35 – 3.26 39.35 – 3.26	Depositor EDS
% Data completeness (in resolution range)	98.2 (39.35-3.26) 98.0 (39.35-3.26)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.63 (at 3.25Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.239 , 0.299 0.244 , 0.319	Depositor DCC
$R_{free}$ test set	872 reflections (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	133.8	Xtriage
Anisotropy	0.404	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 152.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	3 of 17190 reflections (0.017%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	4482	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	170.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.54% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: XP4

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	0.71	4/1480 (0.3%)	0.65	0/2035
1	B	0.72	4/1480 (0.3%)	0.64	0/2035
1	C	0.70	3/1480 (0.2%)	0.63	0/2035
All	All	0.71	11/4440 (0.2%)	0.64	0/6105

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	149	TRP	CD2-CE2	5.50	1.48	1.41
1	C	141	TRP	CD2-CE2	5.46	1.48	1.41
1	A	141	TRP	CD2-CE2	5.35	1.47	1.41
1	B	141	TRP	CD2-CE2	5.30	1.47	1.41
1	C	146	TRP	CD2-CE2	5.21	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 [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	1428	0	1440	173	1
1	B	1428	0	1440	154	0
1	C	1428	0	1440	155	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	27	0	37	1	0
2	B	121	0	185	27	0
2	C	50	0	70	2	0
All	All	4482	0	4612	447	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 49.

The worst 5 of 447 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:53:LEU:HD23	1:A:54:HIS:CE1	1.51	1.45
1:B:53:LEU:HD23	1:B:54:HIS:CE1	1.54	1.42
1:C:53:LEU:HD23	1:C:54:HIS:CE1	1.54	1.42
1:A:53:LEU:CD2	1:A:54:HIS:CE1	2.19	1.24
1:C:19:CYS:O	1:C:23:LYS:HA	1.37	1.23

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:25:ASP:OD2	1:C:25:ASP:OD2[7_555]	2.06	0.14

## 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	176/201 (88%)	153 (87%)	23 (13%)	0	100	100
1	B	176/201 (88%)	153 (87%)	23 (13%)	0	100	100
1	C	176/201 (88%)	154 (88%)	22 (12%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	528/603 (88%)	460 (87%)	68 (13%)	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	150/169 (89%)	147 (98%)	3 (2%)	63	87
1	B	150/169 (89%)	148 (99%)	2 (1%)	76	91
1	C	150/169 (89%)	147 (98%)	3 (2%)	63	87
All	All	450/507 (89%)	442 (98%)	8 (2%)	66	88

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	90	TYR
1	C	142	TRP
1	C	90	TYR
1	A	142	TRP
1	B	142	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	95	HIS
1	C	193	HIS
1	C	54	HIS
1	A	193	HIS
1	B	193	HIS

### 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 ⓘ

9 ligands are 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	XP4	A	201	-	16,16,39	1.31	1 (6%)	16,16,44	1.16	1 (6%)
2	XP4	A	202	-	9,9,39	1.67	1 (11%)	9,9,44	1.42	1 (11%)
2	XP4	B	301	-	35,35,39	1.20	2 (5%)	37,37,44	1.92	6 (16%)
2	XP4	B	302	-	39,39,39	1.35	2 (5%)	43,44,44	1.91	9 (20%)
2	XP4	B	303	-	17,17,39	1.50	1 (5%)	17,17,44	1.46	3 (17%)
2	XP4	B	304	-	16,16,39	1.28	1 (6%)	16,16,44	1.28	2 (12%)
2	XP4	B	305	-	9,9,39	1.82	1 (11%)	9,9,44	1.05	1 (11%)
2	XP4	C	301	-	39,39,39	1.41	3 (7%)	43,44,44	2.38	11 (25%)
2	XP4	C	302	-	9,9,39	1.58	1 (11%)	9,9,44	1.08	1 (11%)

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	XP4	A	201	-	-	0/15/15/41	0/0/0/0
2	XP4	A	202	-	-	0/8/8/41	0/0/0/0
2	XP4	B	301	-	-	0/37/37/41	0/0/0/0
2	XP4	B	302	-	-	0/41/41/41	0/0/0/0
2	XP4	B	303	-	-	0/16/16/41	0/0/0/0
2	XP4	B	304	-	-	0/15/15/41	0/0/0/0
2	XP4	B	305	-	-	0/8/8/41	0/0/0/0
2	XP4	C	301	-	-	0/41/41/41	0/0/0/0
2	XP4	C	302	-	-	0/8/8/41	0/0/0/0

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	301	XP4	P1-O4	2.81	1.68	1.60
2	B	301	XP4	O5-C4	4.17	1.45	1.33
2	C	301	XP4	O5-C4	4.29	1.46	1.33
2	C	302	XP4	O5-C4	4.42	1.47	1.32
2	B	301	XP4	O7-C18	4.73	1.48	1.34

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	301	XP4	C3-C2-C1	-6.43	97.03	112.07
2	C	301	XP4	O6-C4-C5	-4.43	105.98	123.72
2	B	302	XP4	O6-C4-C5	-3.84	108.34	123.72
2	B	301	XP4	O5-C4-O6	-3.72	113.89	123.49
2	B	301	XP4	O6-C4-C5	-2.66	113.08	123.72

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 29 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	201	XP4	1	0
2	B	301	XP4	12	0
2	B	302	XP4	20	0
2	B	303	XP4	3	0
2	B	304	XP4	5	0
2	C	301	XP4	2	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	180/201 (89%)	0.34	17 (9%)	11 8	81, 155, 259, 304	0
1	B	180/201 (89%)	0.47	21 (11%)	6 4	75, 178, 277, 312	0
1	C	180/201 (89%)	0.25	17 (9%)	11 8	73, 170, 257, 286	0
All	All	540/603 (89%)	0.35	55 (10%)	9 6	73, 163, 261, 312	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	126	ASP	11.7
1	B	129	ASP	8.8
1	A	127	MET	7.4
1	B	137	THR	7.0
1	A	58	PRO	7.0

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	XP4	B	301	36/40	0.75	0.40	5.71	79,134,186,192	0
2	XP4	B	305	10/40	0.84	0.39	5.59	96,176,231,244	0
2	XP4	C	301	40/40	0.85	0.33	3.10	63,127,258,273	0
2	XP4	C	302	10/40	0.85	0.38	3.01	150,163,217,238	0
2	XP4	B	302	40/40	0.85	0.29	1.31	72,136,217,291	0
2	XP4	B	304	17/40	0.71	0.40	-	83,148,243,243	0
2	XP4	A	201	17/40	0.69	0.45	-	110,191,275,281	0
2	XP4	B	303	18/40	0.73	0.39	-	112,144,186,252	0
2	XP4	A	202	10/40	0.89	0.51	-	105,174,230,267	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.