



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 30, 2017 – 06:57 PM EDT

PDB ID : 3MVK  
Title : The Crystal Structure of FucU from Bifidobacterium longum to 1.65Å  
Authors : Stein, A.J.; Xu, X.; Cui, H.; Savchenko, A.; Joachimiak, A.; Midwest Center  
for Structural Genomics (MCSG)  
Deposited on : unknown  
Resolution : 1.65 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : **FAILED**  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : rb-20030345

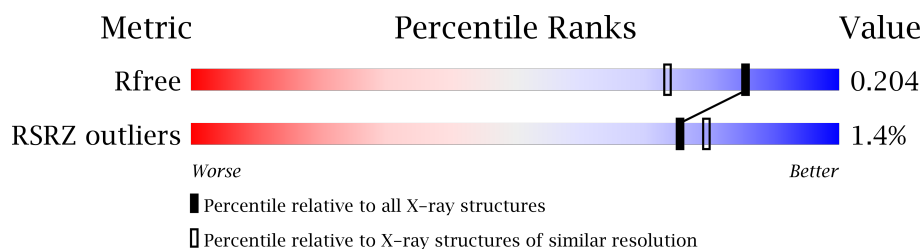
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.65 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))
$R_{free}$	100719	1368 (1.66-1.66)
RSRZ outliers	101464	1371 (1.66-1.66)

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

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	NA	A	149	-	-	-	X
2	NA	B	149	-	-	-	X
3	PGE	C	150	-	-	-	X
3	PGE	C	151	-	-	-	X
3	PGE	E	151	-	-	-	X
3	PGE	E	152	-	-	-	X
3	PGE	G	150	-	-	-	X
3	PGE	H	151	-	-	-	X
3	PGE	J	150	-	-	-	X
4	GOL	A	152	-	-	-	X
4	GOL	B	152	-	-	-	X
4	GOL	E	153	-	-	-	X
4	GOL	F	152	-	-	-	X
4	GOL	H	152	-	-	-	X

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 12068 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 protein FucU.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	140	Total	C	N	O	S	Se	0	2	0
			1054	679	171	198	3	3			
1	B	140	Total	C	N	O	S	Se	0	1	0
			1053	676	171	200	3	3			
1	C	141	Total	C	N	O	S	Se	0	1	0
			1064	685	172	201	3	3			
1	D	140	Total	C	N	O	S	Se	0	1	0
			1055	681	172	196	3	3			
1	E	141	Total	C	N	O	S	Se	0	1	0
			1061	684	172	199	3	3			
1	F	140	Total	C	N	O	S	Se	0	1	0
			1049	674	171	198	3	3			
1	G	143	Total	C	N	O	S	Se	0	2	0
			1068	687	173	202	3	3			
1	H	141	Total	C	N	O	S	Se	0	2	0
			1063	685	172	200	3	3			
1	I	141	Total	C	N	O	S	Se	0	2	0
			1059	683	171	199	3	3			
1	J	141	Total	C	N	O	S	Se	0	1	0
			1064	686	173	199	3	3			

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

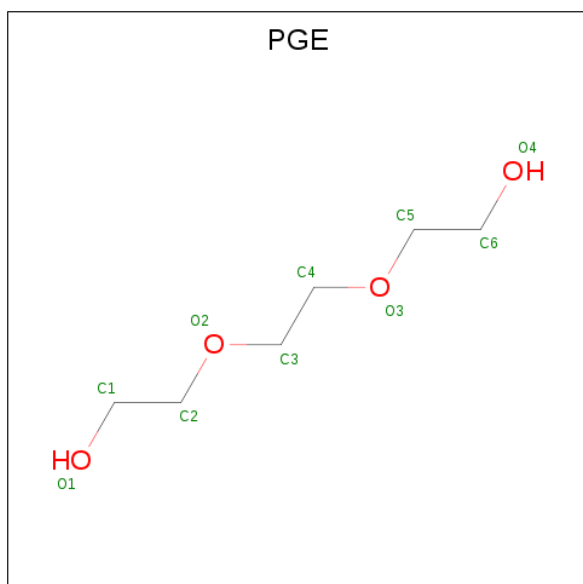
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	1	Total	Na	0	0
			1	1		
2	J	1	Total	Na	0	0
			1	1		
2	D	1	Total	Na	0	0
			1	1		
2	E	2	Total	Na	0	0
			2	2		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	H	2	Total	Na	0	0
			2	2		
2	B	2	Total	Na	0	0
			2	2		
2	I	1	Total	Na	0	0
			1	1		
2	C	1	Total	Na	0	0
			1	1		
2	A	2	Total	Na	0	0
			2	2		
2	F	2	Total	Na	0	0
			2	2		

- Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			10	6	4		
3	B	1	Total	C	O	0	0
			10	6	4		
3	C	1	Total	C	O	0	0
			10	6	4		
3	C	1	Total	C	O	0	0
			10	6	4		
3	E	1	Total	C	O	0	0
			10	6	4		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total	C	O	0	0
			10	6	4		
3	F	1	Total	C	O	0	0
			10	6	4		
3	G	1	Total	C	O	0	0
			10	6	4		
3	H	1	Total	C	O	0	0
			10	6	4		
3	J	1	Total	C	O	0	0
			10	6	4		
3	J	1	Total	C	O	0	0
			10	6	4		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		
4	F	1	Total	C	O	0	0
			6	3	3		
4	G	1	Total	C	O	0	0
			6	3	3		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	H	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	137	Total	O	0	0
			137	137		
5	B	139	Total	O	0	0
			139	139		
5	C	116	Total	O	0	0
			116	116		
5	D	121	Total	O	0	0
			121	121		
5	E	136	Total	O	0	0
			136	136		
5	F	120	Total	O	0	0
			120	120		
5	G	156	Total	O	0	0
			156	156		
5	H	138	Total	O	0	0
			138	138		
5	I	121	Total	O	0	0
			121	121		
5	J	133	Total	O	0	0
			133	133		

MolProbity failed to run properly - this section is therefore empty.

### 3 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.05Å 83.86Å 144.81Å 90.00° 95.11° 90.00°	Depositor
Resolution (Å)	40.72 – 1.65 40.72 – 1.65	Depositor EDS
% Data completeness (in resolution range)	99.8 (40.72-1.65) 99.8 (40.72-1.65)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.14 (at 1.65Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.182 , 0.205 0.181 , 0.204	Depositor DCC
$R_{free}$ test set	9474 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	20.2	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 51.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12068	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 4 Model quality [i](#)

### 4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.3 Torsion angles [i](#)

#### 4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 4.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 4.6 Ligand geometry [i](#)

Of 32 ligands modelled in this entry, 15 are monoatomic - leaving 17 for Mogul analysis.

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	$\# Z  > 2$
3	PGE	A	151	-	9,9,9	0.50	0	8,8,8	0.16	0
4	GOL	A	152	2	5,5,5	0.39	0	5,5,5	0.71	0
3	PGE	B	151	-	9,9,9	0.44	0	8,8,8	0.41	0
4	GOL	B	152	2	5,5,5	0.34	0	5,5,5	0.61	0
3	PGE	C	150	-	9,9,9	0.47	0	8,8,8	0.29	0
3	PGE	C	151	-	9,9,9	0.41	0	8,8,8	0.48	0
3	PGE	E	151	-	9,9,9	0.49	0	8,8,8	0.22	0
3	PGE	E	152	-	9,9,9	0.47	0	8,8,8	0.29	0
4	GOL	E	153	2	5,5,5	0.40	0	5,5,5	0.83	0
3	PGE	F	151	-	9,9,9	0.47	0	8,8,8	0.28	0
4	GOL	F	152	2	5,5,5	0.42	0	5,5,5	0.52	0
3	PGE	G	150	-	9,9,9	0.44	0	8,8,8	0.35	0
4	GOL	G	151	-	5,5,5	0.33	0	5,5,5	0.40	0
3	PGE	H	151	-	9,9,9	0.56	0	8,8,8	0.34	0
4	GOL	H	152	2	5,5,5	0.53	0	5,5,5	1.27	1 (20%)
3	PGE	J	150	-	9,9,9	0.62	0	8,8,8	0.89	0
3	PGE	J	151	-	9,9,9	0.44	0	8,8,8	0.40	0

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
3	PGE	A	151	-	-	0/7/7/7	0/0/0/0
4	GOL	A	152	2	-	0/4/4/4	0/0/0/0
3	PGE	B	151	-	-	0/7/7/7	0/0/0/0
4	GOL	B	152	2	-	0/4/4/4	0/0/0/0
3	PGE	C	150	-	-	0/7/7/7	0/0/0/0
3	PGE	C	151	-	-	0/7/7/7	0/0/0/0
3	PGE	E	151	-	-	0/7/7/7	0/0/0/0
3	PGE	E	152	-	-	0/7/7/7	0/0/0/0
4	GOL	E	153	2	-	0/4/4/4	0/0/0/0
3	PGE	F	151	-	-	0/7/7/7	0/0/0/0
4	GOL	F	152	2	-	0/4/4/4	0/0/0/0
3	PGE	G	150	-	-	0/7/7/7	0/0/0/0
4	GOL	G	151	-	-	0/4/4/4	0/0/0/0
3	PGE	H	151	-	-	0/7/7/7	0/0/0/0
4	GOL	H	152	2	-	0/4/4/4	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PGE	J	150	-	-	0/7/7/7	0/0/0/0
3	PGE	J	151	-	-	0/7/7/7	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	152	GOL	O3-C3-C2	-2.01	99.97	110.07

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 4.7 Other polymers [i](#)

There are no such residues in this entry.

## 4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 5 Fit of model and data ⓘ

### 5.1 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<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	137/148 (92%)	-0.27	1 (0%) 87 89	13, 18, 29, 32	0
1	B	137/148 (92%)	-0.36	0 100 100	13, 18, 30, 36	0
1	C	138/148 (93%)	-0.17	6 (4%) 36 36	14, 21, 35, 40	0
1	D	137/148 (92%)	-0.17	2 (1%) 74 78	12, 19, 34, 43	0
1	E	138/148 (93%)	-0.13	2 (1%) 75 79	13, 19, 33, 42	0
1	F	137/148 (92%)	-0.25	1 (0%) 87 89	14, 20, 32, 39	0
1	G	140/148 (94%)	-0.43	1 (0%) 87 89	13, 18, 29, 43	0
1	H	138/148 (93%)	-0.25	1 (0%) 87 89	13, 18, 29, 38	0
1	I	138/148 (93%)	-0.16	5 (3%) 43 45	14, 21, 35, 38	0
1	J	138/148 (93%)	-0.25	0 100 100	13, 19, 29, 34	0
All	All	1378/1480 (93%)	-0.24	19 (1%) 75 79	12, 19, 32, 43	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	145	PHE	5.6
1	C	81	GLY	4.7
1	I	134	TYR	4.3
1	C	83	THR	3.7
1	H	145	PHE	3.6
1	I	85	ALA	3.5
1	I	83	THR	3.4
1	D	80	PRO	3.1
1	C	85	ALA	3.1
1	I	81	GLY	2.9
1	C	145	PHE	2.9
1	E	101	GLU	2.9
1	C	79	VAL	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	134	TYR	2.7
1	D	81	GLY	2.5
1	A	85	ALA	2.4
1	I	113	PHE	2.3
1	G	4	GLY	2.1
1	F	105	ASP	2.0

## 5.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.4 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(Å <sup>2</sup> )	Q<0.9
3	PGE	C	151	10/10	0.78	0.38	15.11	38,41,47,48	0
4	GOL	H	152	6/6	0.79	0.25	13.53	31,39,39,40	0
3	PGE	J	150	10/10	0.79	0.29	11.22	33,35,38,40	0
2	NA	B	149	1/1	0.98	0.15	10.84	22,22,22,22	0
3	PGE	E	152	10/10	0.84	0.16	7.83	44,46,48,49	0
4	GOL	E	153	6/6	0.75	0.20	7.06	34,35,37,39	0
4	GOL	A	152	6/6	0.86	0.16	6.32	27,33,35,39	0
4	GOL	F	152	6/6	0.79	0.19	4.37	29,36,39,42	0
4	GOL	B	152	6/6	0.87	0.18	4.13	34,39,40,41	0
3	PGE	H	151	10/10	0.78	0.12	3.96	43,47,49,49	0
3	PGE	C	150	10/10	0.83	0.11	3.58	40,41,42,43	0
2	NA	A	149	1/1	0.97	0.13	3.27	23,23,23,23	0
3	PGE	G	150	10/10	0.89	0.10	2.65	43,44,44,45	0
3	PGE	E	151	10/10	0.87	0.11	2.55	44,45,45,45	0
3	PGE	B	151	10/10	0.85	0.12	1.95	47,47,49,49	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	PGE	A	151	10/10	0.84	0.13	1.45	48,49,50,51	0
2	NA	F	149	1/1	0.97	0.10	1.29	27,27,27,27	0
2	NA	F	150	1/1	0.96	0.08	0.68	28,28,28,28	0
3	PGE	J	151	10/10	0.92	0.08	0.54	48,50,52,53	0
4	GOL	G	151	6/6	0.95	0.08	0.45	20,25,25,27	0
2	NA	E	149	1/1	0.97	0.11	0.17	24,24,24,24	0
2	NA	G	149	1/1	0.97	0.07	0.01	22,22,22,22	0
2	NA	D	149	1/1	0.97	0.07	-0.16	20,20,20,20	0
3	PGE	F	151	10/10	0.89	0.10	-0.27	51,51,53,54	0
2	NA	H	149	1/1	0.98	0.08	-0.50	23,23,23,23	0
2	NA	C	149	1/1	0.99	0.05	-0.76	24,24,24,24	0
2	NA	I	149	1/1	0.99	0.05	-0.98	23,23,23,23	0
2	NA	J	149	1/1	0.99	0.06	-1.16	22,22,22,22	0
2	NA	H	150	1/1	0.94	0.06	-1.62	25,25,25,25	0
2	NA	A	150	1/1	0.99	0.05	-4.58	27,27,27,27	0
2	NA	E	150	1/1	0.97	0.09	-	27,27,27,27	0
2	NA	B	150	1/1	0.99	0.08	-	26,26,26,26	0

## 5.5 Other polymers [i](#)

There are no such residues in this entry.