



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 06:55 am BST

PDB ID : 6AZQ  
Title : Structural and biochemical characterization of a non-canonical biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminosorum* bv. *viciae* 3841  
Authors : Peat, T.S.; Esquirol, L.; Newman, J.; Scott, C.  
Deposited on : 2017-09-11  
Resolution : 2.22 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

<https://www.wwpdb.org/validation/2017/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	:	1.8.5 (274361), CSD as541be (2020)
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

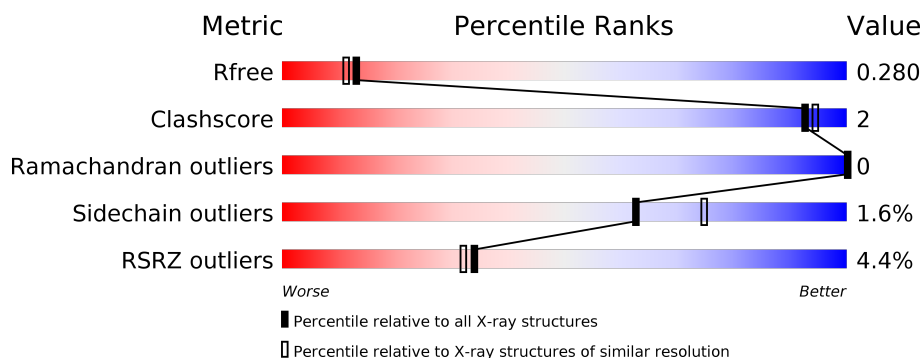
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.22 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	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RSRZ outliers	127900	5797 (2.24-2.20)

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 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  $\leq 5\%$ . 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	262	<div> <div>4%</div> <div> <div></div> <div>82%</div> <div>•</div> <div>14%</div> </div> </div>
1	B	262	<div> <div>%</div> <div> <div></div> <div>84%</div> <div>•</div> <div>14%</div> </div> </div>
1	C	262	<div> <div>2%</div> <div> <div></div> <div>81%</div> <div>• •</div> <div>14%</div> </div> </div>
1	D	262	<div> <div>2%</div> <div> <div></div> <div>82%</div> <div>•</div> <div>14%</div> </div> </div>
1	E	262	<div> <div>7%</div> <div> <div></div> <div>82%</div> <div>•</div> <div>14%</div> </div> </div>
1	F	262	<div> <div>8%</div> <div> <div></div> <div>81%</div> <div>• •</div> <div>14%</div> </div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	262	
1	H	262	

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	C5J	C	301	-	X	-	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 14376 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 Putative amidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	226	Total	C	N	O	S	0	2	0
			1749	1098	308	330	13			
1	H	226	Total	C	N	O	S	0	2	0
			1748	1099	309	327	13			
1	A	226	Total	C	N	O	S	0	0	0
			1732	1090	305	325	12			
1	B	226	Total	C	N	O	S	0	1	0
			1738	1093	306	326	13			
1	C	225	Total	C	N	O	S	0	2	0
			1743	1096	308	326	13			
1	D	226	Total	C	N	O	S	0	0	0
			1732	1090	305	325	12			
1	E	226	Total	C	N	O	S	0	1	0
			1738	1093	306	326	13			
1	F	225	Total	C	N	O	S	0	1	0
			1733	1090	305	325	13			

There are 240 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	-19	MET	-	initiating methionine	UNP Q1M7F4
G	-18	GLY	-	expression tag	UNP Q1M7F4
G	-17	SER	-	expression tag	UNP Q1M7F4
G	-16	SER	-	expression tag	UNP Q1M7F4
G	-15	HIS	-	expression tag	UNP Q1M7F4
G	-14	HIS	-	expression tag	UNP Q1M7F4
G	-13	HIS	-	expression tag	UNP Q1M7F4
G	-12	HIS	-	expression tag	UNP Q1M7F4
G	-11	HIS	-	expression tag	UNP Q1M7F4
G	-10	HIS	-	expression tag	UNP Q1M7F4
G	-9	SER	-	expression tag	UNP Q1M7F4
G	-8	SER	-	expression tag	UNP Q1M7F4
G	-7	GLY	-	expression tag	UNP Q1M7F4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	-6	LEU	-	expression tag	UNP Q1M7F4
G	-5	VAL	-	expression tag	UNP Q1M7F4
G	-4	PRO	-	expression tag	UNP Q1M7F4
G	-3	ARG	-	expression tag	UNP Q1M7F4
G	-2	GLY	-	expression tag	UNP Q1M7F4
G	-1	SER	-	expression tag	UNP Q1M7F4
G	0	HIS	-	expression tag	UNP Q1M7F4
G	175	SER	CYS	conflict	UNP Q1M7F4
G	234	GLY	-	expression tag	UNP Q1M7F4
G	235	LEU	-	expression tag	UNP Q1M7F4
G	236	VAL	-	expression tag	UNP Q1M7F4
G	237	PRO	-	expression tag	UNP Q1M7F4
G	238	ARG	-	expression tag	UNP Q1M7F4
G	239	GLY	-	expression tag	UNP Q1M7F4
G	240	SER	-	expression tag	UNP Q1M7F4
G	241	ILE	-	expression tag	UNP Q1M7F4
G	242	GLU	-	expression tag	UNP Q1M7F4
H	-19	MET	-	initiating methionine	UNP Q1M7F4
H	-18	GLY	-	expression tag	UNP Q1M7F4
H	-17	SER	-	expression tag	UNP Q1M7F4
H	-16	SER	-	expression tag	UNP Q1M7F4
H	-15	HIS	-	expression tag	UNP Q1M7F4
H	-14	HIS	-	expression tag	UNP Q1M7F4
H	-13	HIS	-	expression tag	UNP Q1M7F4
H	-12	HIS	-	expression tag	UNP Q1M7F4
H	-11	HIS	-	expression tag	UNP Q1M7F4
H	-10	HIS	-	expression tag	UNP Q1M7F4
H	-9	SER	-	expression tag	UNP Q1M7F4
H	-8	SER	-	expression tag	UNP Q1M7F4
H	-7	GLY	-	expression tag	UNP Q1M7F4
H	-6	LEU	-	expression tag	UNP Q1M7F4
H	-5	VAL	-	expression tag	UNP Q1M7F4
H	-4	PRO	-	expression tag	UNP Q1M7F4
H	-3	ARG	-	expression tag	UNP Q1M7F4
H	-2	GLY	-	expression tag	UNP Q1M7F4
H	-1	SER	-	expression tag	UNP Q1M7F4
H	0	HIS	-	expression tag	UNP Q1M7F4
H	175	SER	CYS	conflict	UNP Q1M7F4
H	234	GLY	-	expression tag	UNP Q1M7F4
H	235	LEU	-	expression tag	UNP Q1M7F4
H	236	VAL	-	expression tag	UNP Q1M7F4
H	237	PRO	-	expression tag	UNP Q1M7F4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	238	ARG	-	expression tag	UNP Q1M7F4
H	239	GLY	-	expression tag	UNP Q1M7F4
H	240	SER	-	expression tag	UNP Q1M7F4
H	241	ILE	-	expression tag	UNP Q1M7F4
H	242	GLU	-	expression tag	UNP Q1M7F4
A	-19	MET	-	initiating methionine	UNP Q1M7F4
A	-18	GLY	-	expression tag	UNP Q1M7F4
A	-17	SER	-	expression tag	UNP Q1M7F4
A	-16	SER	-	expression tag	UNP Q1M7F4
A	-15	HIS	-	expression tag	UNP Q1M7F4
A	-14	HIS	-	expression tag	UNP Q1M7F4
A	-13	HIS	-	expression tag	UNP Q1M7F4
A	-12	HIS	-	expression tag	UNP Q1M7F4
A	-11	HIS	-	expression tag	UNP Q1M7F4
A	-10	HIS	-	expression tag	UNP Q1M7F4
A	-9	SER	-	expression tag	UNP Q1M7F4
A	-8	SER	-	expression tag	UNP Q1M7F4
A	-7	GLY	-	expression tag	UNP Q1M7F4
A	-6	LEU	-	expression tag	UNP Q1M7F4
A	-5	VAL	-	expression tag	UNP Q1M7F4
A	-4	PRO	-	expression tag	UNP Q1M7F4
A	-3	ARG	-	expression tag	UNP Q1M7F4
A	-2	GLY	-	expression tag	UNP Q1M7F4
A	-1	SER	-	expression tag	UNP Q1M7F4
A	0	HIS	-	expression tag	UNP Q1M7F4
A	175	SER	CYS	conflict	UNP Q1M7F4
A	234	GLY	-	expression tag	UNP Q1M7F4
A	235	LEU	-	expression tag	UNP Q1M7F4
A	236	VAL	-	expression tag	UNP Q1M7F4
A	237	PRO	-	expression tag	UNP Q1M7F4
A	238	ARG	-	expression tag	UNP Q1M7F4
A	239	GLY	-	expression tag	UNP Q1M7F4
A	240	SER	-	expression tag	UNP Q1M7F4
A	241	ILE	-	expression tag	UNP Q1M7F4
A	242	GLU	-	expression tag	UNP Q1M7F4
B	-19	MET	-	initiating methionine	UNP Q1M7F4
B	-18	GLY	-	expression tag	UNP Q1M7F4
B	-17	SER	-	expression tag	UNP Q1M7F4
B	-16	SER	-	expression tag	UNP Q1M7F4
B	-15	HIS	-	expression tag	UNP Q1M7F4
B	-14	HIS	-	expression tag	UNP Q1M7F4
B	-13	HIS	-	expression tag	UNP Q1M7F4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-12	HIS	-	expression tag	UNP Q1M7F4
B	-11	HIS	-	expression tag	UNP Q1M7F4
B	-10	HIS	-	expression tag	UNP Q1M7F4
B	-9	SER	-	expression tag	UNP Q1M7F4
B	-8	SER	-	expression tag	UNP Q1M7F4
B	-7	GLY	-	expression tag	UNP Q1M7F4
B	-6	LEU	-	expression tag	UNP Q1M7F4
B	-5	VAL	-	expression tag	UNP Q1M7F4
B	-4	PRO	-	expression tag	UNP Q1M7F4
B	-3	ARG	-	expression tag	UNP Q1M7F4
B	-2	GLY	-	expression tag	UNP Q1M7F4
B	-1	SER	-	expression tag	UNP Q1M7F4
B	0	HIS	-	expression tag	UNP Q1M7F4
B	175	SER	CYS	conflict	UNP Q1M7F4
B	234	GLY	-	expression tag	UNP Q1M7F4
B	235	LEU	-	expression tag	UNP Q1M7F4
B	236	VAL	-	expression tag	UNP Q1M7F4
B	237	PRO	-	expression tag	UNP Q1M7F4
B	238	ARG	-	expression tag	UNP Q1M7F4
B	239	GLY	-	expression tag	UNP Q1M7F4
B	240	SER	-	expression tag	UNP Q1M7F4
B	241	ILE	-	expression tag	UNP Q1M7F4
B	242	GLU	-	expression tag	UNP Q1M7F4
C	-19	MET	-	initiating methionine	UNP Q1M7F4
C	-18	GLY	-	expression tag	UNP Q1M7F4
C	-17	SER	-	expression tag	UNP Q1M7F4
C	-16	SER	-	expression tag	UNP Q1M7F4
C	-15	HIS	-	expression tag	UNP Q1M7F4
C	-14	HIS	-	expression tag	UNP Q1M7F4
C	-13	HIS	-	expression tag	UNP Q1M7F4
C	-12	HIS	-	expression tag	UNP Q1M7F4
C	-11	HIS	-	expression tag	UNP Q1M7F4
C	-10	HIS	-	expression tag	UNP Q1M7F4
C	-9	SER	-	expression tag	UNP Q1M7F4
C	-8	SER	-	expression tag	UNP Q1M7F4
C	-7	GLY	-	expression tag	UNP Q1M7F4
C	-6	LEU	-	expression tag	UNP Q1M7F4
C	-5	VAL	-	expression tag	UNP Q1M7F4
C	-4	PRO	-	expression tag	UNP Q1M7F4
C	-3	ARG	-	expression tag	UNP Q1M7F4
C	-2	GLY	-	expression tag	UNP Q1M7F4
C	-1	SER	-	expression tag	UNP Q1M7F4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	HIS	-	expression tag	UNP Q1M7F4
C	175	SER	CYS	conflict	UNP Q1M7F4
C	234	GLY	-	expression tag	UNP Q1M7F4
C	235	LEU	-	expression tag	UNP Q1M7F4
C	236	VAL	-	expression tag	UNP Q1M7F4
C	237	PRO	-	expression tag	UNP Q1M7F4
C	238	ARG	-	expression tag	UNP Q1M7F4
C	239	GLY	-	expression tag	UNP Q1M7F4
C	240	SER	-	expression tag	UNP Q1M7F4
C	241	ILE	-	expression tag	UNP Q1M7F4
C	242	GLU	-	expression tag	UNP Q1M7F4
D	-19	MET	-	initiating methionine	UNP Q1M7F4
D	-18	GLY	-	expression tag	UNP Q1M7F4
D	-17	SER	-	expression tag	UNP Q1M7F4
D	-16	SER	-	expression tag	UNP Q1M7F4
D	-15	HIS	-	expression tag	UNP Q1M7F4
D	-14	HIS	-	expression tag	UNP Q1M7F4
D	-13	HIS	-	expression tag	UNP Q1M7F4
D	-12	HIS	-	expression tag	UNP Q1M7F4
D	-11	HIS	-	expression tag	UNP Q1M7F4
D	-10	HIS	-	expression tag	UNP Q1M7F4
D	-9	SER	-	expression tag	UNP Q1M7F4
D	-8	SER	-	expression tag	UNP Q1M7F4
D	-7	GLY	-	expression tag	UNP Q1M7F4
D	-6	LEU	-	expression tag	UNP Q1M7F4
D	-5	VAL	-	expression tag	UNP Q1M7F4
D	-4	PRO	-	expression tag	UNP Q1M7F4
D	-3	ARG	-	expression tag	UNP Q1M7F4
D	-2	GLY	-	expression tag	UNP Q1M7F4
D	-1	SER	-	expression tag	UNP Q1M7F4
D	0	HIS	-	expression tag	UNP Q1M7F4
D	175	SER	CYS	conflict	UNP Q1M7F4
D	234	GLY	-	expression tag	UNP Q1M7F4
D	235	LEU	-	expression tag	UNP Q1M7F4
D	236	VAL	-	expression tag	UNP Q1M7F4
D	237	PRO	-	expression tag	UNP Q1M7F4
D	238	ARG	-	expression tag	UNP Q1M7F4
D	239	GLY	-	expression tag	UNP Q1M7F4
D	240	SER	-	expression tag	UNP Q1M7F4
D	241	ILE	-	expression tag	UNP Q1M7F4
D	242	GLU	-	expression tag	UNP Q1M7F4
E	-19	MET	-	initiating methionine	UNP Q1M7F4

*Continued on next page...*



*Continued from previous page...*

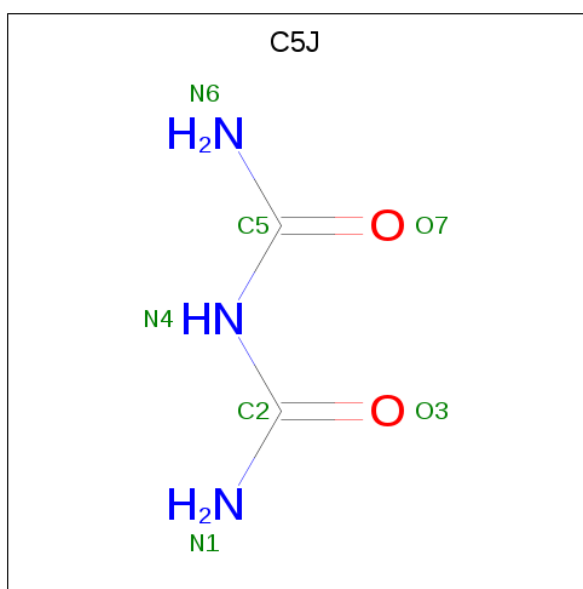
Chain	Residue	Modelled	Actual	Comment	Reference
E	-18	GLY	-	expression tag	UNP Q1M7F4
E	-17	SER	-	expression tag	UNP Q1M7F4
E	-16	SER	-	expression tag	UNP Q1M7F4
E	-15	HIS	-	expression tag	UNP Q1M7F4
E	-14	HIS	-	expression tag	UNP Q1M7F4
E	-13	HIS	-	expression tag	UNP Q1M7F4
E	-12	HIS	-	expression tag	UNP Q1M7F4
E	-11	HIS	-	expression tag	UNP Q1M7F4
E	-10	HIS	-	expression tag	UNP Q1M7F4
E	-9	SER	-	expression tag	UNP Q1M7F4
E	-8	SER	-	expression tag	UNP Q1M7F4
E	-7	GLY	-	expression tag	UNP Q1M7F4
E	-6	LEU	-	expression tag	UNP Q1M7F4
E	-5	VAL	-	expression tag	UNP Q1M7F4
E	-4	PRO	-	expression tag	UNP Q1M7F4
E	-3	ARG	-	expression tag	UNP Q1M7F4
E	-2	GLY	-	expression tag	UNP Q1M7F4
E	-1	SER	-	expression tag	UNP Q1M7F4
E	0	HIS	-	expression tag	UNP Q1M7F4
E	175	SER	CYS	conflict	UNP Q1M7F4
E	234	GLY	-	expression tag	UNP Q1M7F4
E	235	LEU	-	expression tag	UNP Q1M7F4
E	236	VAL	-	expression tag	UNP Q1M7F4
E	237	PRO	-	expression tag	UNP Q1M7F4
E	238	ARG	-	expression tag	UNP Q1M7F4
E	239	GLY	-	expression tag	UNP Q1M7F4
E	240	SER	-	expression tag	UNP Q1M7F4
E	241	ILE	-	expression tag	UNP Q1M7F4
E	242	GLU	-	expression tag	UNP Q1M7F4
F	-19	MET	-	initiating methionine	UNP Q1M7F4
F	-18	GLY	-	expression tag	UNP Q1M7F4
F	-17	SER	-	expression tag	UNP Q1M7F4
F	-16	SER	-	expression tag	UNP Q1M7F4
F	-15	HIS	-	expression tag	UNP Q1M7F4
F	-14	HIS	-	expression tag	UNP Q1M7F4
F	-13	HIS	-	expression tag	UNP Q1M7F4
F	-12	HIS	-	expression tag	UNP Q1M7F4
F	-11	HIS	-	expression tag	UNP Q1M7F4
F	-10	HIS	-	expression tag	UNP Q1M7F4
F	-9	SER	-	expression tag	UNP Q1M7F4
F	-8	SER	-	expression tag	UNP Q1M7F4
F	-7	GLY	-	expression tag	UNP Q1M7F4

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	-6	LEU	-	expression tag	UNP Q1M7F4
F	-5	VAL	-	expression tag	UNP Q1M7F4
F	-4	PRO	-	expression tag	UNP Q1M7F4
F	-3	ARG	-	expression tag	UNP Q1M7F4
F	-2	GLY	-	expression tag	UNP Q1M7F4
F	-1	SER	-	expression tag	UNP Q1M7F4
F	0	HIS	-	expression tag	UNP Q1M7F4
F	175	SER	CYS	conflict	UNP Q1M7F4
F	234	GLY	-	expression tag	UNP Q1M7F4
F	235	LEU	-	expression tag	UNP Q1M7F4
F	236	VAL	-	expression tag	UNP Q1M7F4
F	237	PRO	-	expression tag	UNP Q1M7F4
F	238	ARG	-	expression tag	UNP Q1M7F4
F	239	GLY	-	expression tag	UNP Q1M7F4
F	240	SER	-	expression tag	UNP Q1M7F4
F	241	ILE	-	expression tag	UNP Q1M7F4
F	242	GLU	-	expression tag	UNP Q1M7F4

- Molecule 2 is dicarbonimidic diamide (three-letter code: C5J) (formula:  $C_2H_5N_3O_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	G	1	Total	C	N	O	0	0
			7	2	3	2		
2	H	1	Total	C	N	O	0	0
			7	2	3	2		
2	A	1	Total	C	N	O	0	0
			7	2	3	2		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	N	O	0	0
			7	2	3	2		
2	C	1	Total	C	N	O	0	0
			7	2	3	2		
2	D	1	Total	C	N	O	0	0
			7	2	3	2		
2	E	1	Total	C	N	O	0	0
			7	2	3	2		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	1	Total	Ca	0	0
			1	1		
3	B	1	Total	Ca	0	0
			1	1		
3	C	2	Total	Ca	0	0
			2	2		
3	F	1	Total	Ca	0	0
			1	1		

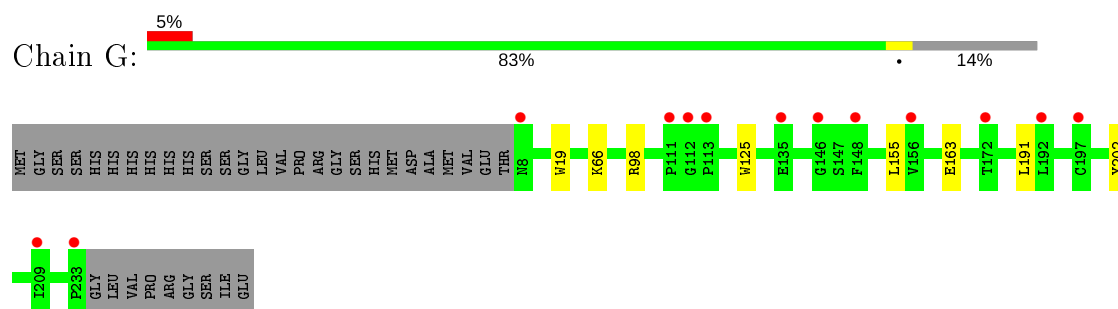
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	60	Total	O	0	0
			60	60		
4	H	64	Total	O	0	0
			64	64		
4	A	49	Total	O	0	0
			49	49		
4	B	61	Total	O	0	0
			61	61		
4	C	62	Total	O	0	0
			62	62		
4	D	54	Total	O	0	0
			54	54		
4	E	32	Total	O	0	0
			32	32		
4	F	27	Total	O	0	0
			27	27		

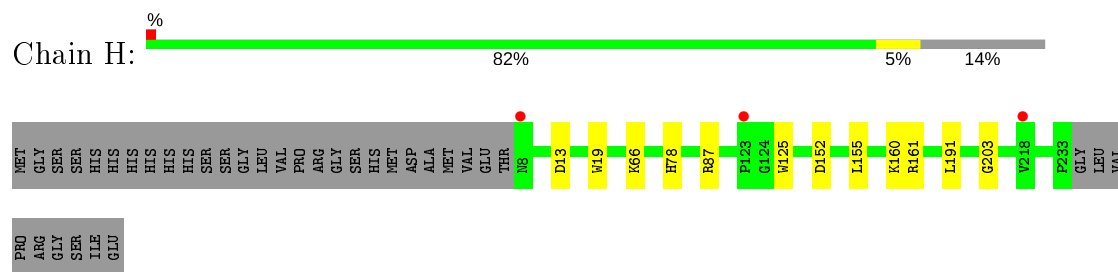
### 3 Residue-property plots [i](#)

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.

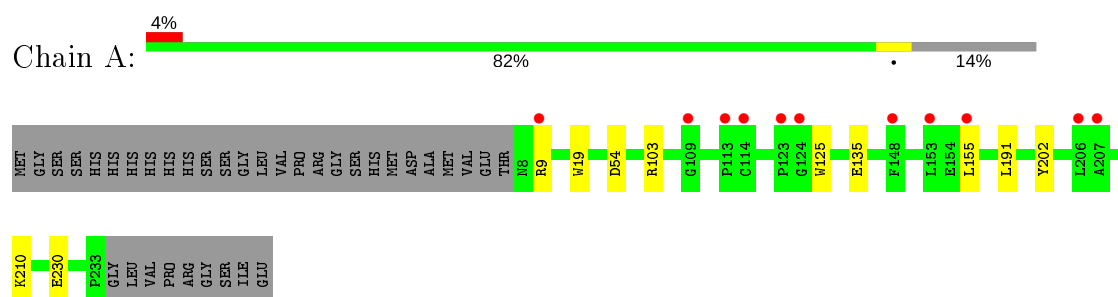
- Molecule 1: Putative amidase



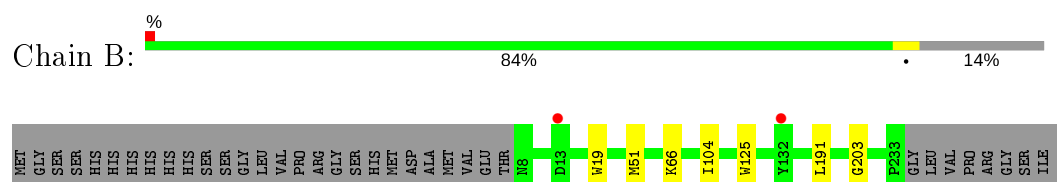
- Molecule 1: Putative amidase



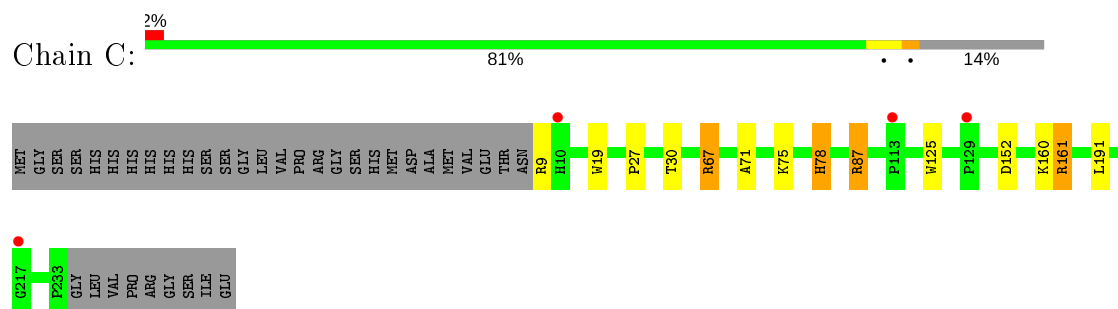
- Molecule 1: Putative amidase



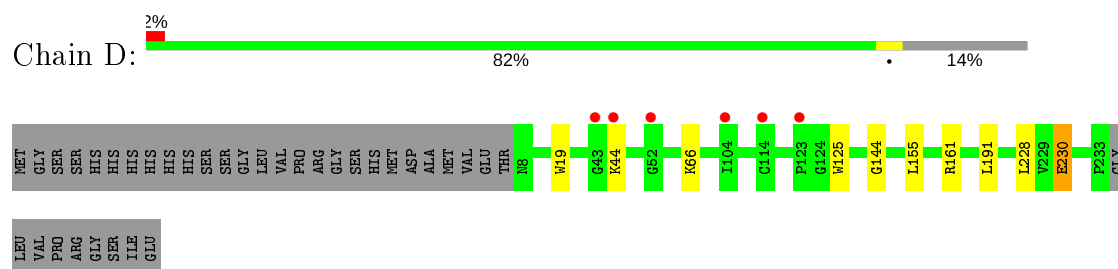
- Molecule 1: Putative amidase



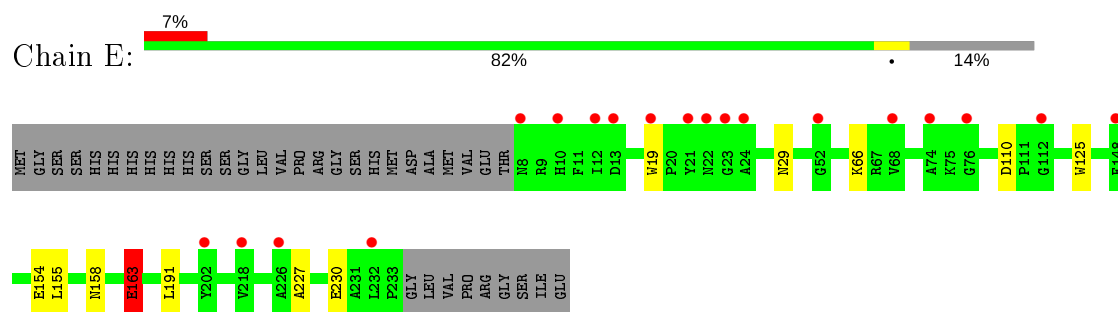
- Molecule 1: Putative amidase



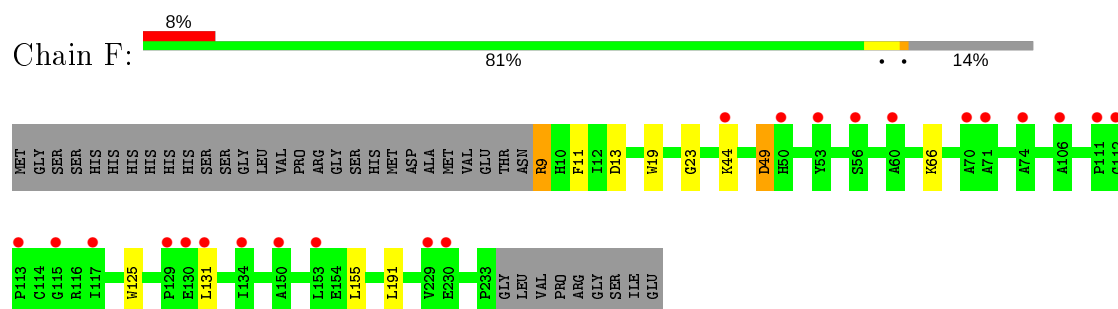
- Molecule 1: Putative amidase



- Molecule 1: Putative amidase



- Molecule 1: Putative amidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.70 Å 87.30 Å 341.69 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.85 – 2.22 42.61 – 2.22	Depositor EDS
% Data completeness (in resolution range)	99.9 (42.85-2.22) 100.0 (42.61-2.22)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.26 (at 2.22 Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.249 , 0.277 0.252 , 0.280	Depositor DCC
$R_{free}$ test set	5462 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.5	Xtriage
Anisotropy	0.030	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 33.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	14376	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.80% 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.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CA, C5J

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	0.52	0/1770	0.75	2/2403 (0.1%)
1	B	0.55	0/1776	0.76	0/2411
1	C	0.56	0/1782	0.76	3/2419 (0.1%)
1	D	0.52	0/1770	0.79	5/2403 (0.2%)
1	E	0.52	1/1776 (0.1%)	0.78	2/2411 (0.1%)
1	F	0.49	0/1771	0.77	3/2404 (0.1%)
1	G	0.54	0/1787	0.76	1/2426 (0.0%)
1	H	0.60	1/1787 (0.1%)	0.81	4/2426 (0.2%)
All	All	0.54	2/14219 (0.0%)	0.77	20/19303 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	13	ASP	CB-CG	-6.62	1.37	1.51
1	E	163	GLU	CD-OE2	-5.17	1.20	1.25

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	13	ASP	CB-CG-OD1	10.31	127.58	118.30
1	F	49	ASP	CB-CG-OD2	-10.18	109.14	118.30
1	E	110	ASP	CB-CG-OD2	9.72	127.05	118.30
1	G	98	ARG	NE-CZ-NH1	-8.47	116.06	120.30
1	H	13	ASP	CB-CG-OD2	-8.18	110.94	118.30

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	1732	0	1713	5	0
1	B	1738	0	1717	4	0
1	C	1743	0	1721	16	0
1	D	1732	0	1713	2	0
1	E	1738	0	1717	7	0
1	F	1733	0	1715	5	0
1	G	1749	0	1724	4	0
1	H	1748	0	1723	6	0
2	A	7	0	0	0	0
2	B	7	0	0	0	0
2	C	7	0	0	0	0
2	D	7	0	0	0	0
2	E	7	0	0	0	0
2	G	7	0	0	0	0
2	H	7	0	0	0	0
3	B	1	0	0	0	0
3	C	2	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
4	A	49	0	0	1	0
4	B	61	0	0	0	0
4	C	62	0	0	2	0
4	D	54	0	0	0	0
4	E	32	0	0	1	0
4	F	27	0	0	0	0
4	G	60	0	0	1	0
4	H	64	0	0	0	0
All	All	14376	0	13743	42	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 42 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:78[A]:HIS:CD2	1:C:78[A]:HIS:H	1.95	0.79
1:H:203:GLY:HA3	1:A:202:TYR:CD1	2.23	0.73
1:C:27:PRO:O	1:C:78[B]:HIS:CE1	2.45	0.70
1:F:44:LYS:HA	1:F:49:ASP:OD2	1.93	0.69
1:G:202:TYR:CD1	1:B:203:GLY:HA3	2.28	0.68

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	224/262 (86%)	220 (98%)	4 (2%)	0	100	100
1	B	225/262 (86%)	221 (98%)	4 (2%)	0	100	100
1	C	225/262 (86%)	221 (98%)	4 (2%)	0	100	100
1	D	224/262 (86%)	220 (98%)	4 (2%)	0	100	100
1	E	225/262 (86%)	221 (98%)	4 (2%)	0	100	100
1	F	224/262 (86%)	220 (98%)	4 (2%)	0	100	100
1	G	226/262 (86%)	222 (98%)	4 (2%)	0	100	100
1	H	226/262 (86%)	222 (98%)	4 (2%)	0	100	100
All	All	1799/2096 (86%)	1767 (98%)	32 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	183/214 (86%)	179 (98%)	4 (2%)	52	64
1	B	184/214 (86%)	182 (99%)	2 (1%)	73	84
1	C	185/214 (86%)	181 (98%)	4 (2%)	52	64
1	D	183/214 (86%)	180 (98%)	3 (2%)	62	75
1	E	184/214 (86%)	180 (98%)	4 (2%)	52	64
1	F	184/214 (86%)	181 (98%)	3 (2%)	62	75
1	G	186/214 (87%)	183 (98%)	3 (2%)	62	75
1	H	185/214 (86%)	183 (99%)	2 (1%)	73	84
All	All	1474/1712 (86%)	1449 (98%)	25 (2%)	62	73

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	9	ARG
1	C	78[B]	HIS
1	F	125	TRP
1	C	78[A]	HIS
1	C	125	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	102	GLN
1	E	158	ASN
1	E	215	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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 5 are monoatomic - leaving 7 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$
2	C5J	D	301	-	6,6,6	1.28	0	7,7,7	1.68	3 (42%)
2	C5J	H	301	-	6,6,6	1.71	2 (33%)	7,7,7	1.71	1 (14%)
2	C5J	A	301	-	6,6,6	1.50	1 (16%)	7,7,7	1.86	2 (28%)
2	C5J	C	301	-	6,6,6	2.13	2 (33%)	7,7,7	2.13	3 (42%)
2	C5J	E	301	-	6,6,6	1.62	1 (16%)	7,7,7	1.75	2 (28%)
2	C5J	G	301	-	6,6,6	1.65	1 (16%)	7,7,7	1.75	3 (42%)
2	C5J	B	301	-	6,6,6	1.30	0	7,7,7	1.68	3 (42%)

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	C5J	D	301	-	-	2/4/4/4	-
2	C5J	H	301	-	-	2/4/4/4	-
2	C5J	A	301	-	-	2/4/4/4	-
2	C5J	C	301	-	-	2/4/4/4	-
2	C5J	E	301	-	-	2/4/4/4	-
2	C5J	G	301	-	-	2/4/4/4	-
2	C5J	B	301	-	-	2/4/4/4	-

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	301	C5J	C5-N4	-4.29	1.34	1.40
2	A	301	C5J	C5-N4	-3.19	1.35	1.40
2	E	301	C5J	O7-C5	-2.65	1.20	1.24
2	G	301	C5J	C2-N4	-2.45	1.36	1.40
2	H	301	C5J	C5-N6	2.30	1.38	1.33

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	301	C5J	C5-N4-C2	-3.79	122.45	127.53
2	C	301	C5J	C5-N4-C2	-3.43	122.94	127.53
2	C	301	C5J	N1-C2-N4	-3.29	108.31	115.58
2	E	301	C5J	O3-C2-N1	-3.01	118.06	123.22
2	A	301	C5J	O3-C2-N1	-2.95	118.16	123.22

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	301	C5J	N1-C2-N4-C5
2	D	301	C5J	O3-C2-N4-C5
2	H	301	C5J	N6-C5-N4-C2
2	H	301	C5J	O7-C5-N4-C2
2	A	301	C5J	N6-C5-N4-C2

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<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	226/262 (86%)	0.44	11 (4%) 29 28	19, 31, 44, 57	0
1	B	226/262 (86%)	0.21	2 (0%) 84 83	19, 28, 39, 55	0
1	C	225/262 (85%)	0.39	4 (1%) 68 66	17, 32, 48, 60	0
1	D	226/262 (86%)	0.39	6 (2%) 54 52	16, 32, 50, 70	0
1	E	226/262 (86%)	0.60	19 (8%) 11 9	22, 37, 58, 68	0
1	F	225/262 (85%)	0.73	22 (9%) 7 6	22, 42, 64, 72	0
1	G	226/262 (86%)	0.50	13 (5%) 23 21	16, 29, 44, 60	0
1	H	226/262 (86%)	0.26	3 (1%) 77 75	16, 25, 39, 63	0
All	All	1806/2096 (86%)	0.44	80 (4%) 34 32	16, 31, 52, 72	0

The worst 5 of 80 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	21	TYR	4.2
1	F	60	ALA	4.0
1	F	117	ILE	3.7
1	D	123	PRO	3.7
1	E	10	HIS	3.6

### 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 [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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
3	CA	C	303	1/1	0.86	0.07	67,67,67,67	0
3	CA	F	301	1/1	0.90	0.27	52,52,52,52	0
2	C5J	A	301	7/7	0.94	0.15	20,22,23,25	0
2	C5J	H	301	7/7	0.95	0.12	13,15,17,19	0
2	C5J	E	301	7/7	0.96	0.13	19,21,28,28	0
3	CA	B	302	1/1	0.96	0.21	50,50,50,50	0
2	C5J	B	301	7/7	0.96	0.11	23,26,28,30	0
3	CA	C	302	1/1	0.96	0.22	57,57,57,57	0
3	CA	G	302	1/1	0.96	0.23	51,51,51,51	0
2	C5J	D	301	7/7	0.97	0.10	25,26,28,28	0
2	C5J	G	301	7/7	0.97	0.17	24,25,27,30	0
2	C5J	C	301	7/7	0.97	0.12	21,25,26,27	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.