



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 10, 2023 – 12:36 AM EDT

PDB ID : 4HXG  
Title : Pyrococcus horikoshii acylaminoacyl peptidase (orthorhombic crystal form)  
Authors : Kiss-Szeman, A.; Menyhard, D.K.; Tichy-Racs, E.; Hornung, B.; Radi, K.; Szeltner, Z.; Domokos, K.; Szamosi, I.; Naray-Szabo, G.; Polgar, L.; Harmat, V.  
Deposited on : 2012-11-09  
Resolution : 2.70 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35.1  
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) : 2.35.1

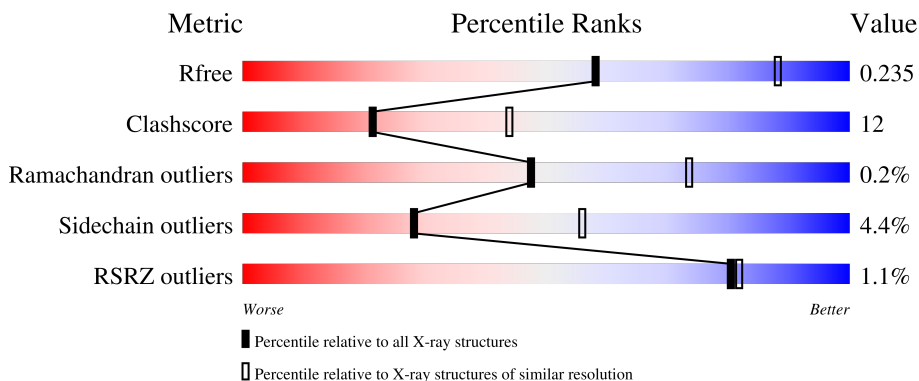
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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  $\geq 3$ , 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	622	
1	B	622	
1	C	622	
1	D	622	
1	E	622	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain	
1	F	622	73%	24% ..
1	G	622	72%	24% ..
1	H	622	66%	30% ..
1	I	622	69%	27% ..
1	J	622	4%	67% 29% ..
1	K	622	5%	70% 26% ..
1	L	622	%	68% 29% ..

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	HEZ	D	701	-	-	X	-
2	HEZ	E	702	-	-	-	X

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 59910 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 uncharacterized protein PH0594.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	608	4940	3225	803	900	12	0	3	0
1	B	612	4926	3211	800	903	12	0	2	0
1	C	612	4883	3183	791	897	12	0	1	0
1	D	612	4989	3254	812	911	12	0	3	0
1	E	613	5005	3259	815	919	12	0	2	0
1	F	614	4975	3237	814	912	12	0	1	0
1	G	608	4886	3190	792	892	12	0	2	0
1	H	609	4866	3168	795	891	12	0	0	0
1	I	611	4910	3201	794	903	12	0	2	0
1	J	609	4617	2999	753	853	12	0	1	0
1	K	610	4659	3017	762	868	12	0	1	0
1	L	611	4862	3165	784	901	12	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

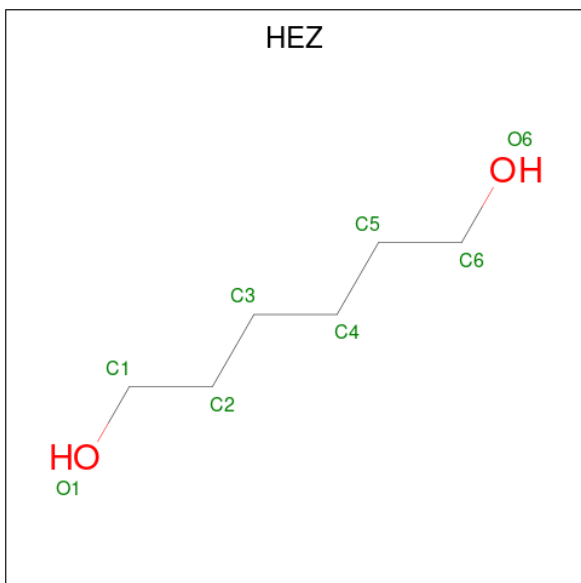
Chain	Residue	Modelled	Actual	Comment	Reference
A	466	ALA	SER	engineered mutation	UNP O58323
B	466	ALA	SER	engineered mutation	UNP O58323
C	466	ALA	SER	engineered mutation	UNP O58323
D	466	ALA	SER	engineered mutation	UNP O58323
E	466	ALA	SER	engineered mutation	UNP O58323

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	466	ALA	SER	engineered mutation	UNP O58323
G	466	ALA	SER	engineered mutation	UNP O58323
H	466	ALA	SER	engineered mutation	UNP O58323
I	466	ALA	SER	engineered mutation	UNP O58323
J	466	ALA	SER	engineered mutation	UNP O58323
K	466	ALA	SER	engineered mutation	UNP O58323
L	466	ALA	SER	engineered mutation	UNP O58323

- Molecule 2 is HEXANE-1,6-DIOL (three-letter code: HEZ) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 8 6 2	0	0
2	A	1	Total C O 8 6 2	0	0
2	A	1	Total C O 8 6 2	0	0
2	B	1	Total C O 8 6 2	0	0
2	B	1	Total C O 8 6 2	0	0
2	C	1	Total C O 8 6 2	0	0
2	D	1	Total C O 8 6 2	0	0
2	D	1	Total C O 8 6 2	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	1	Total C O 8 6 2	0	0
2	D	1	Total C O 8 6 2	0	0
2	D	1	Total C O 8 6 2	0	0
2	D	1	Total C O 8 6 2	0	0
2	D	1	Total C O 8 6 2	0	0
2	D	1	Total C O 8 6 2	0	0
2	E	1	Total C O 8 6 2	0	0
2	E	1	Total C O 8 6 2	0	0
2	E	1	Total C O 8 6 2	0	0
2	F	1	Total C O 8 6 2	0	0
2	F	1	Total C O 8 6 2	0	0
2	F	1	Total C O 8 6 2	0	0
2	G	1	Total C O 8 6 2	0	0
2	H	1	Total C O 8 6 2	0	0
2	I	1	Total C O 8 6 2	0	0
2	I	1	Total C O 8 6 2	0	0
2	K	1	Total C O 8 6 2	0	0
2	L	1	Total C O 8 6 2	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total Mg 2 2	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Mg 1 1	0	0
3	C	2	Total Mg 2 2	0	0
3	D	2	Total Mg 2 2	0	0
3	E	3	Total Mg 3 3	0	0
3	F	3	Total Mg 3 3	0	0
3	G	2	Total Mg 2 2	0	0
3	H	2	Total Mg 2 2	0	0
3	I	2	Total Mg 2 2	0	0
3	L	2	Total Mg 2 2	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	1	Total Cl 1 1	0	0
4	F	1	Total Cl 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	149	Total O 152 152	0	3
5	B	85	Total O 86 86	0	1
5	C	59	Total O 60 60	0	1
5	D	221	Total O 221 221	0	0
5	E	188	Total O 191 191	0	3
5	F	155	Total O 156 156	0	1

*Continued on next page...*

*Continued from previous page...*

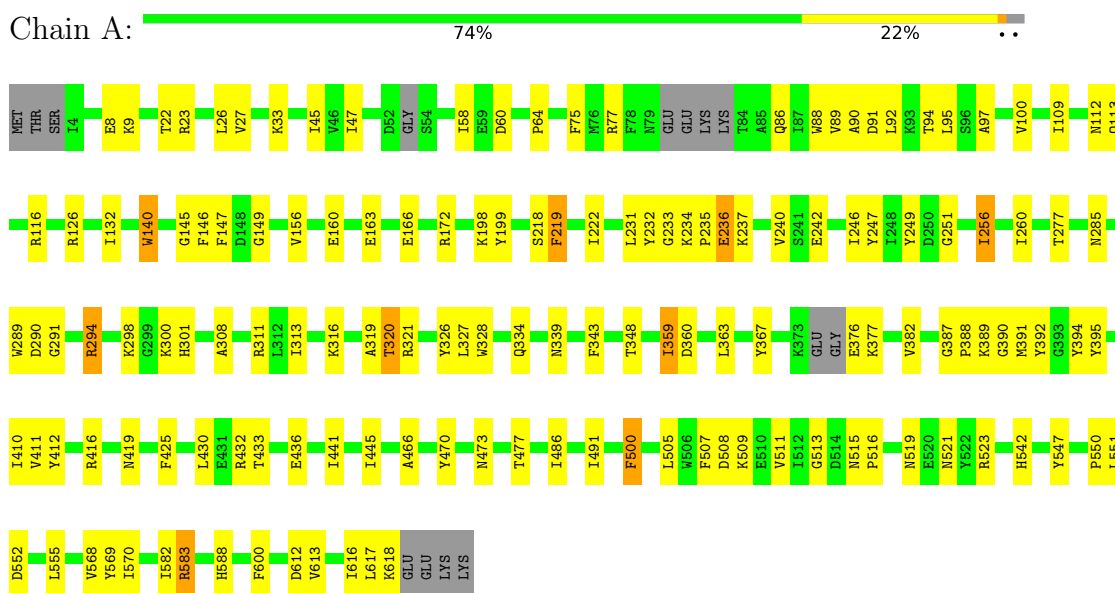
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	G	76	Total 79	O 79	0	3
5	H	58	Total 58	O 58	0	0
5	I	77	Total 78	O 78	0	1
5	J	17	Total 17	O 17	0	0
5	K	11	Total 12	O 12	0	1
5	L	51	Total 51	O 51	0	0



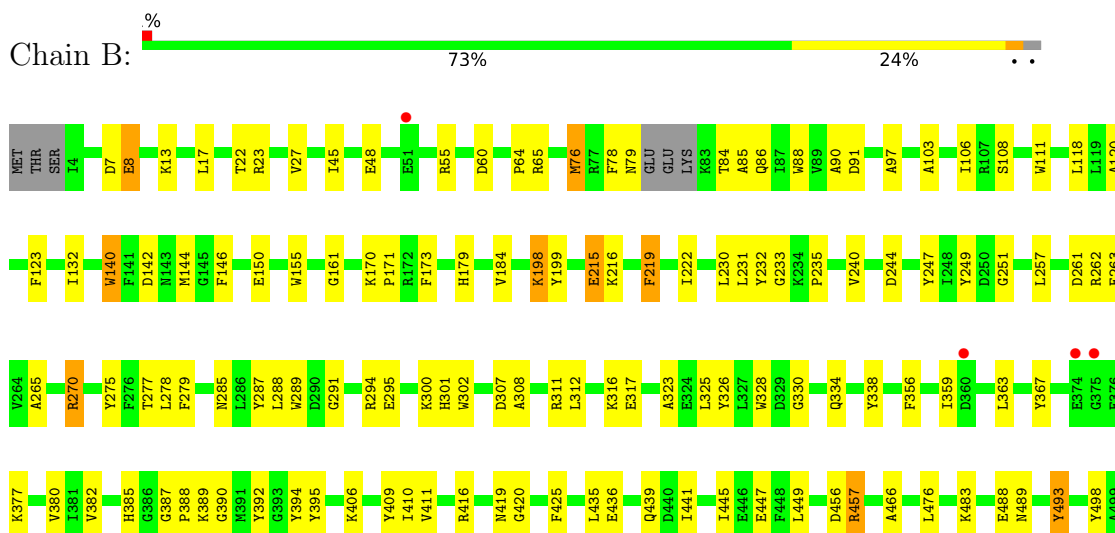
### 3 Residue-property plots

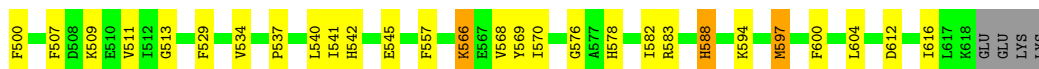
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Putative uncharacterized protein PH0594

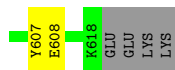
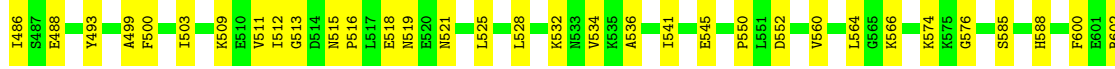
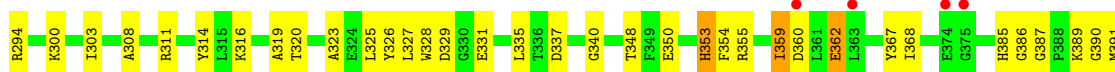
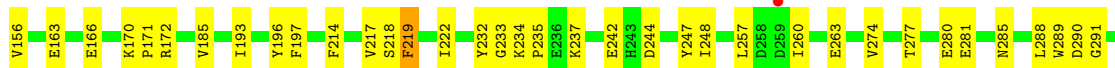
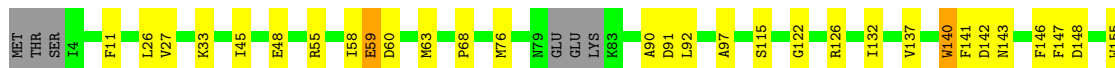


- Molecule 1: Putative uncharacterized protein PH0594

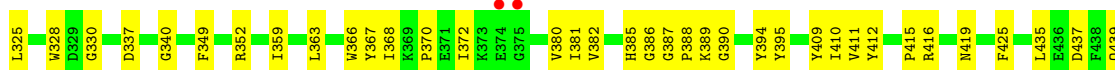
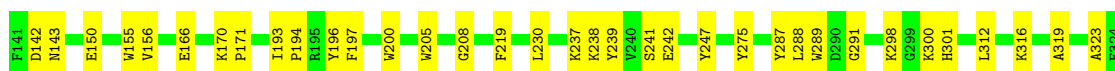




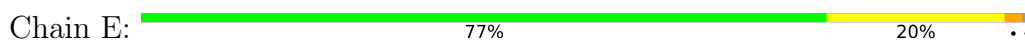
- Molecule 1: Putative uncharacterized protein PH0594

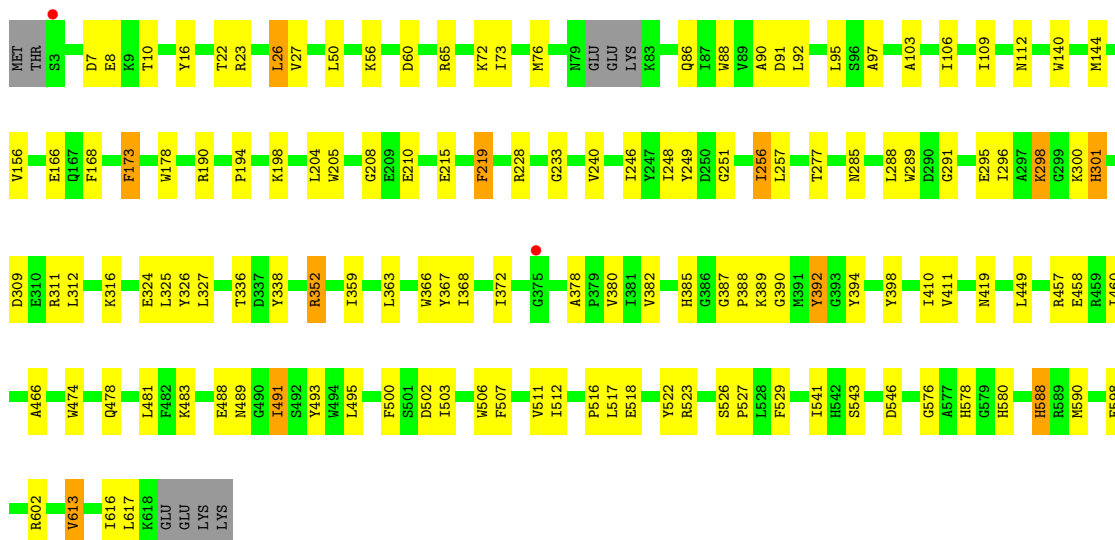


- Molecule 1: Putative uncharacterized protein PH0594

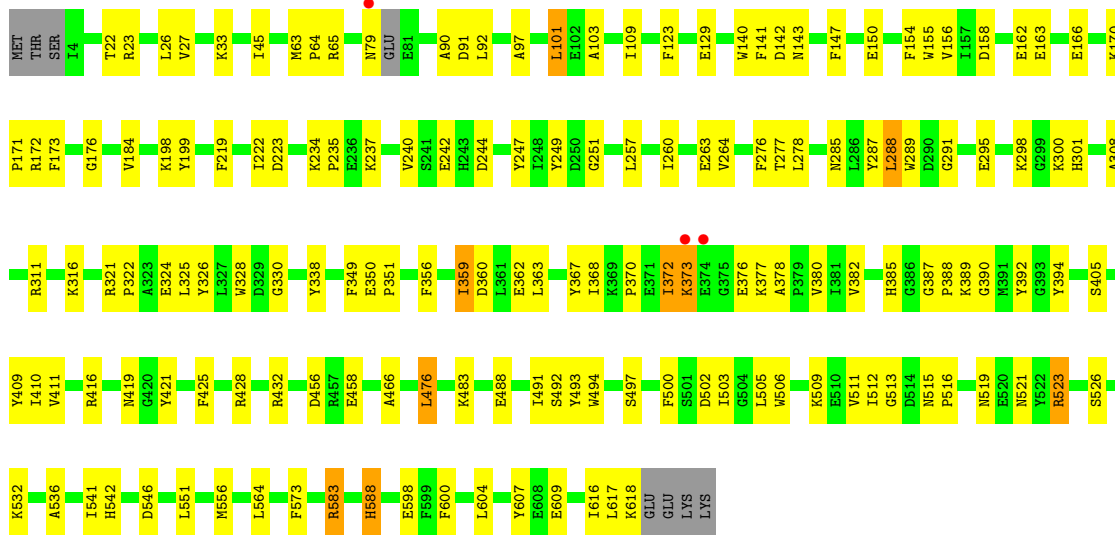


- Molecule 1: Putative uncharacterized protein PH0594

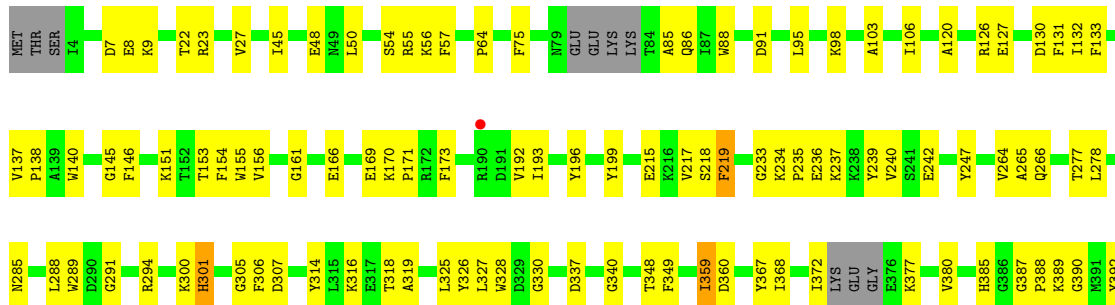


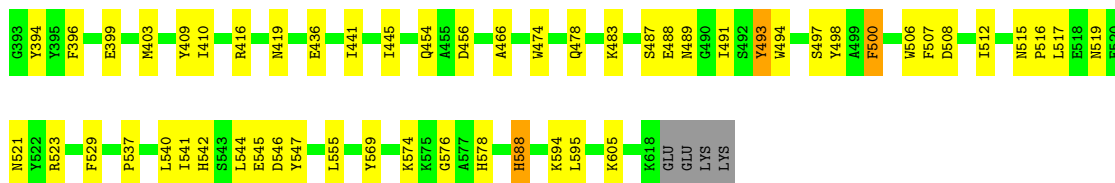


• Molecule 1: Putative uncharacterized protein PH0594

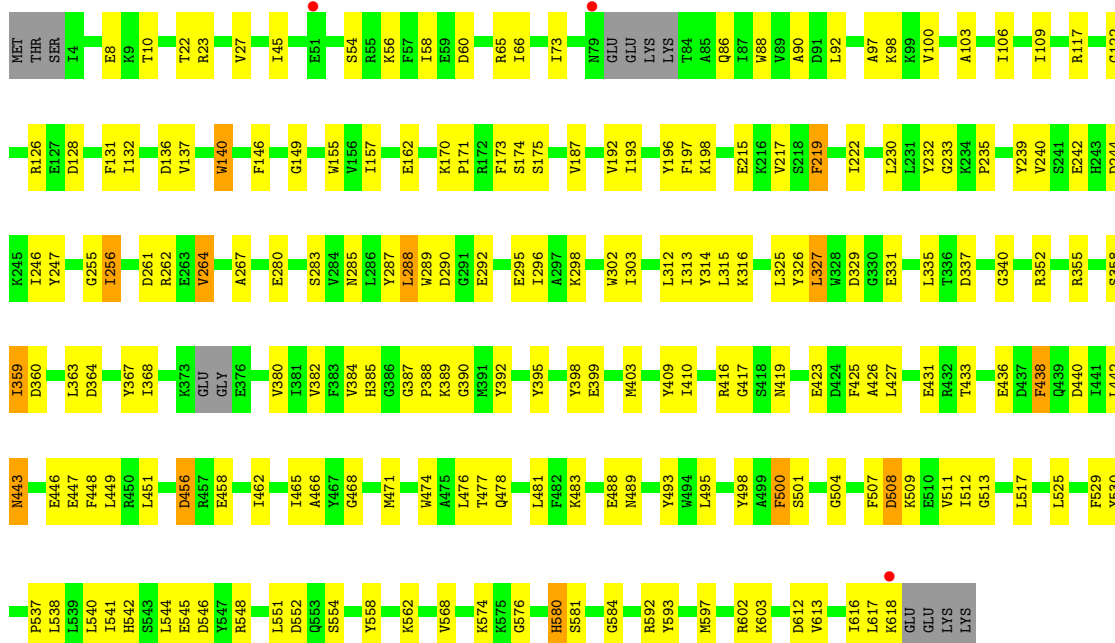


• Molecule 1: Putative uncharacterized protein PH0594

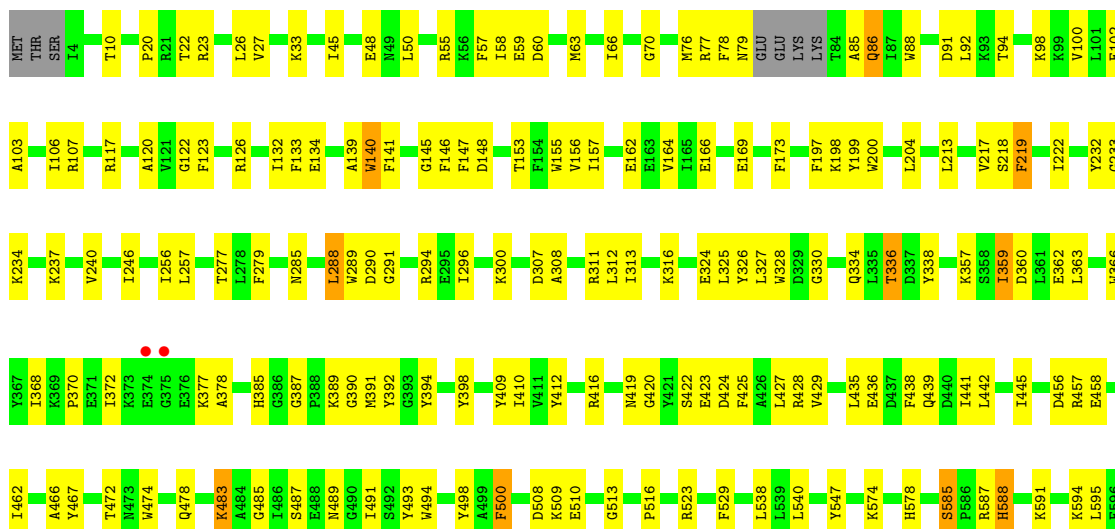




• Molecule 1: Putative uncharacterized protein PH0594

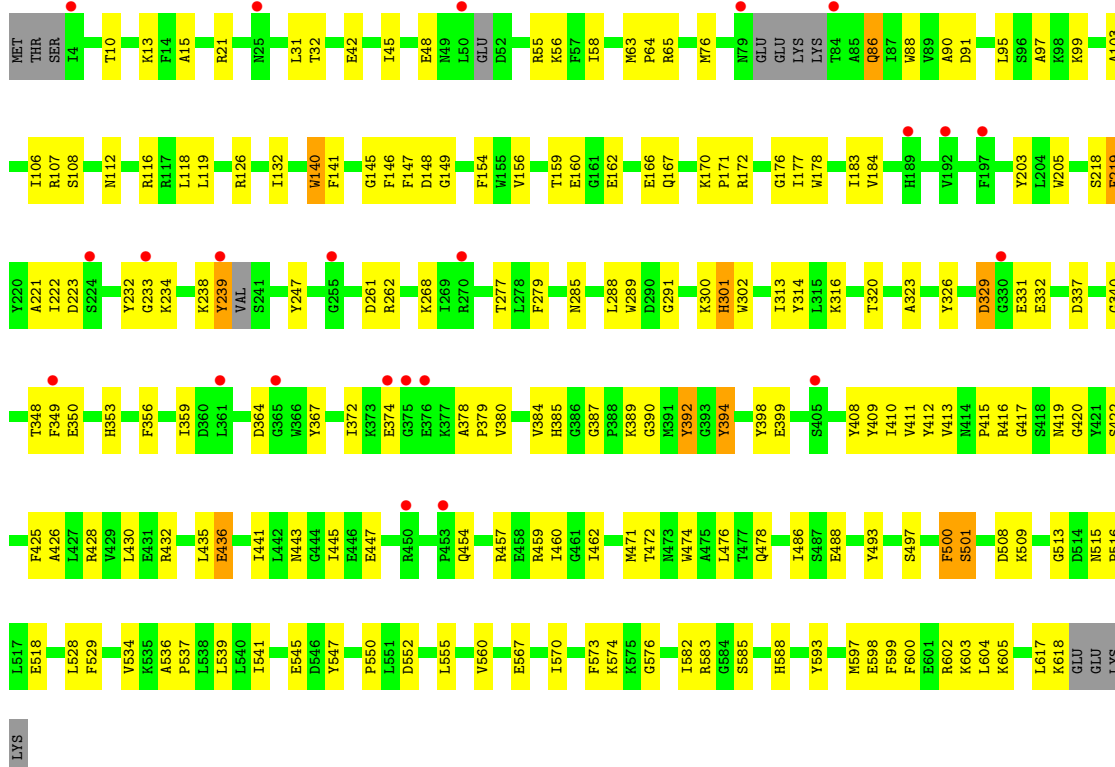


• Molecule 1: Putative uncharacterized protein PH0594

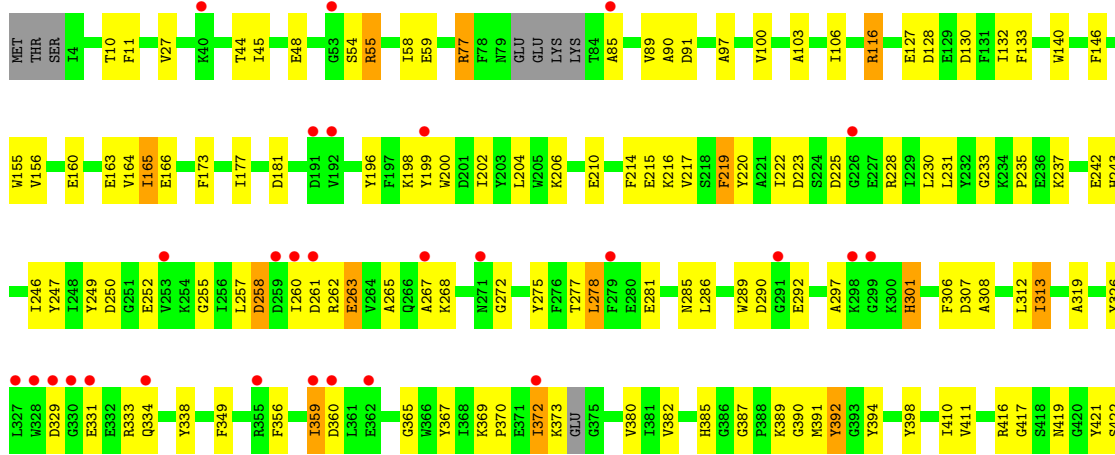


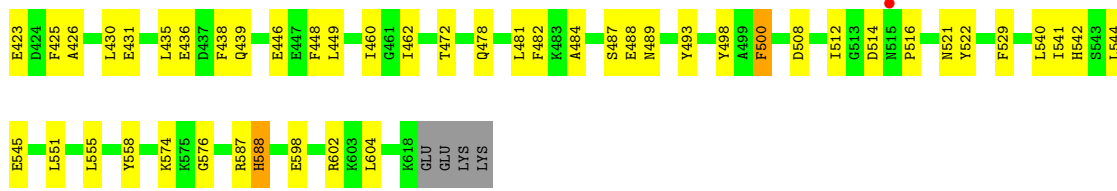


● Molecule 1: Putative uncharacterized protein PH0594

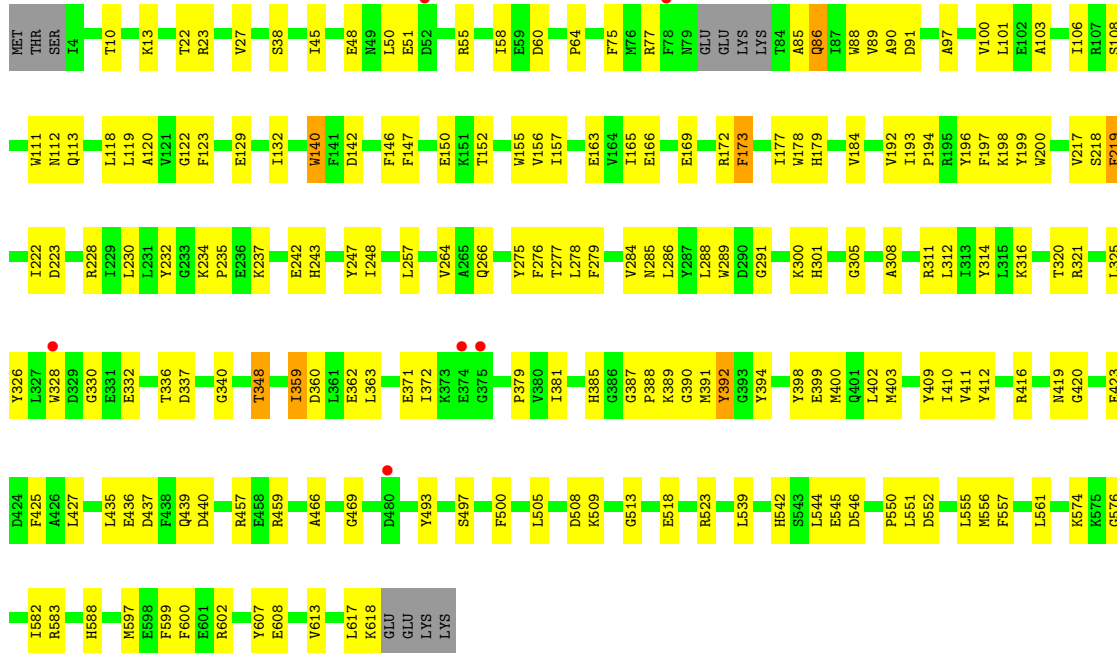


● Molecule 1: Putative uncharacterized protein PH0594





● Molecule 1: Putative uncharacterized protein PH0594



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	183.31Å 183.80Å 275.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.98 – 2.70 19.98 – 2.70	Depositor EDS
% Data completeness (in resolution range)	87.6 (19.98-2.70) 87.1 (19.98-2.70)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.48 (at 2.71Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7_650)	Depositor
R, $R_{free}$	0.200 , 0.241 0.194 , 0.235	Depositor DCC
$R_{free}$ test set	1906 reflections (0.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.7	Xtrriage
Anisotropy	0.384	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 67.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.073 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	59910	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	67.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.83% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: CL, HEZ, MG

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	0.33	0/5083	0.54	0/6881
1	B	0.30	0/5068	0.50	0/6869
1	C	0.28	0/5022	0.48	0/6815
1	D	0.35	0/5134	0.55	0/6945
1	E	0.35	0/5147	0.54	0/6964
1	F	0.32	0/5112	0.53	0/6917
1	G	0.29	0/5027	0.48	0/6814
1	H	0.28	0/4999	0.46	0/6779
1	I	0.29	0/5050	0.49	0/6844
1	J	0.25	0/4750	0.44	0/6470
1	K	0.26	0/4790	0.45	0/6513
1	L	0.28	0/4996	0.47	0/6781
All	All	0.30	0/60178	0.50	0/81592

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 [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	4940	0	4668	102	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	4926	0	4614	104	0
1	C	4883	0	4516	109	0
1	D	4989	0	4741	103	0
1	E	5005	0	4745	91	0
1	F	4975	0	4709	113	0
1	G	4886	0	4563	111	0
1	H	4866	0	4530	149	0
1	I	4910	0	4605	135	0
1	J	4617	0	4007	124	0
1	K	4659	0	4088	137	0
1	L	4862	0	4488	122	0
2	A	24	0	42	3	0
2	B	16	0	28	0	0
2	C	8	0	14	1	0
2	D	64	0	112	12	0
2	E	24	0	42	5	0
2	F	24	0	42	4	0
2	G	8	0	14	4	0
2	H	8	0	14	2	0
2	I	16	0	28	4	0
2	K	8	0	14	0	0
2	L	8	0	14	1	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
3	C	2	0	0	0	0
3	D	2	0	0	0	0
3	E	3	0	0	0	0
3	F	3	0	0	0	0
3	G	2	0	0	0	0
3	H	2	0	0	0	0
3	I	2	0	0	0	0
3	L	2	0	0	0	0
4	D	1	0	0	0	0
4	F	1	0	0	0	0
5	A	152	0	0	1	0
5	B	86	0	0	1	0
5	C	60	0	0	0	0
5	D	221	0	0	3	0
5	E	191	0	0	2	0
5	F	156	0	0	2	0
5	G	79	0	0	2	0
5	H	58	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	I	78	0	0	0	0
5	J	17	0	0	0	0
5	K	12	0	0	0	0
5	L	51	0	0	0	0
All	All	59910	0	54638	1362	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 12.

The worst 5 of 1362 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:K:77:ARG:HG2	1:K:77:ARG:HH11	1.08	1.09
1:I:103:ALA:HB1	1:I:106:ILE:HG13	1.41	1.02
1:I:86:GLN:HG3	1:I:88:TRP:HE1	1.25	1.00
1:B:8:GLU:HG2	1:B:594:LYS:HB2	1.45	0.97
1:J:156:VAL:HB	1:J:166:GLU:HB2	1.50	0.92

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	603/622 (97%)	579 (96%)	23 (4%)	1 (0%)	47 73
1	B	610/622 (98%)	591 (97%)	18 (3%)	1 (0%)	47 73
1	C	609/622 (98%)	580 (95%)	28 (5%)	1 (0%)	47 73
1	D	611/622 (98%)	589 (96%)	21 (3%)	1 (0%)	47 73
1	E	611/622 (98%)	591 (97%)	19 (3%)	1 (0%)	47 73
1	F	611/622 (98%)	588 (96%)	22 (4%)	1 (0%)	47 73

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	604/622 (97%)	580 (96%)	23 (4%)	1 (0%)	47	73
1	H	603/622 (97%)	566 (94%)	36 (6%)	1 (0%)	47	73
1	I	609/622 (98%)	586 (96%)	22 (4%)	1 (0%)	47	73
1	J	602/622 (97%)	566 (94%)	35 (6%)	1 (0%)	47	73
1	K	605/622 (97%)	568 (94%)	35 (6%)	2 (0%)	41	66
1	L	607/622 (98%)	579 (95%)	27 (4%)	1 (0%)	47	73
All	All	7285/7464 (98%)	6963 (96%)	309 (4%)	13 (0%)	47	73

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	K	372	ILE
1	A	359	ILE
1	B	359	ILE
1	C	359	ILE
1	D	359	ILE

### 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	490/548 (89%)	476 (97%)	14 (3%)	42	71
1	B	485/548 (88%)	460 (95%)	25 (5%)	23	49
1	C	473/548 (86%)	461 (98%)	12 (2%)	47	76
1	D	499/548 (91%)	476 (95%)	23 (5%)	27	54
1	E	502/548 (92%)	477 (95%)	25 (5%)	24	51
1	F	496/548 (90%)	479 (97%)	17 (3%)	37	66
1	G	477/548 (87%)	451 (94%)	26 (6%)	21	46
1	H	476/548 (87%)	456 (96%)	20 (4%)	30	58
1	I	485/548 (88%)	466 (96%)	19 (4%)	32	61
1	J	407/548 (74%)	384 (94%)	23 (6%)	20	44

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	K	418/548 (76%)	392 (94%)	26 (6%)	18	40
1	L	473/548 (86%)	452 (96%)	21 (4%)	28	56
All	All	5681/6576 (86%)	5430 (96%)	251 (4%)	28	56

5 of 251 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	609	GLU
1	K	313	ILE
1	H	60	ASP
1	K	278	LEU
1	L	198	LYS

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

Mol	Chain	Res	Type
1	G	478	GLN
1	K	243	HIS
1	H	454	GLN
1	L	266	GLN
1	J	385	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](#)

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

### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

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

Of 49 ligands modelled in this entry, 23 are monoatomic - leaving 26 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	HEZ	D	707	-	7,7,7	0.27	0	6,6,6	0.55	0
2	HEZ	I	702	-	7,7,7	0.28	0	6,6,6	0.61	0
2	HEZ	D	702	-	7,7,7	0.45	0	6,6,6	0.38	0
2	HEZ	C	701	-	7,7,7	0.25	0	6,6,6	0.66	0
2	HEZ	B	701	-	7,7,7	0.24	0	6,6,6	0.66	0
2	HEZ	D	701	-	7,7,7	0.29	0	6,6,6	0.53	0
2	HEZ	H	701	-	7,7,7	0.23	0	6,6,6	0.68	0
2	HEZ	D	704	-	7,7,7	0.35	0	6,6,6	0.46	0
2	HEZ	G	701	-	7,7,7	0.29	0	6,6,6	0.55	0
2	HEZ	A	701	-	7,7,7	0.30	0	6,6,6	0.46	0
2	HEZ	E	701	-	7,7,7	0.28	0	6,6,6	0.59	0
2	HEZ	D	705	-	7,7,7	0.34	0	6,6,6	0.43	0
2	HEZ	D	703	-	7,7,7	0.31	0	6,6,6	0.50	0
2	HEZ	B	702	-	7,7,7	0.40	0	6,6,6	0.40	0
2	HEZ	K	701	-	7,7,7	0.32	0	6,6,6	0.52	0
2	HEZ	F	703	-	7,7,7	0.30	0	6,6,6	0.51	0
2	HEZ	A	703	-	7,7,7	0.32	0	6,6,6	0.42	0
2	HEZ	E	703	-	7,7,7	0.43	0	6,6,6	0.29	0
2	HEZ	D	708	-	7,7,7	0.38	0	6,6,6	0.28	0
2	HEZ	I	701	-	7,7,7	0.31	0	6,6,6	0.43	0
2	HEZ	F	702	-	7,7,7	0.32	0	6,6,6	0.46	0
2	HEZ	A	702	-	7,7,7	0.35	0	6,6,6	0.42	0
2	HEZ	E	702	-	7,7,7	0.23	0	6,6,6	0.70	0
2	HEZ	D	706	-	7,7,7	0.39	0	6,6,6	0.29	0
2	HEZ	L	701	-	7,7,7	0.32	0	6,6,6	0.43	0
2	HEZ	F	701	-	7,7,7	0.33	0	6,6,6	0.55	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
2	HEZ	D	707	-	-	1/5/5/5	-
2	HEZ	I	702	-	-	2/5/5/5	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEZ	D	702	-	-	2/5/5/5	-
2	HEZ	C	701	-	-	1/5/5/5	-
2	HEZ	B	701	-	-	0/5/5/5	-
2	HEZ	D	701	-	-	0/5/5/5	-
2	HEZ	H	701	-	-	2/5/5/5	-
2	HEZ	D	704	-	-	0/5/5/5	-
2	HEZ	G	701	-	-	0/5/5/5	-
2	HEZ	A	701	-	-	3/5/5/5	-
2	HEZ	E	701	-	-	0/5/5/5	-
2	HEZ	D	705	-	-	1/5/5/5	-
2	HEZ	D	703	-	-	2/5/5/5	-
2	HEZ	B	702	-	-	3/5/5/5	-
2	HEZ	K	701	-	-	0/5/5/5	-
2	HEZ	F	703	-	-	3/5/5/5	-
2	HEZ	A	703	-	-	3/5/5/5	-
2	HEZ	E	703	-	-	0/5/5/5	-
2	HEZ	D	708	-	-	0/5/5/5	-
2	HEZ	I	701	-	-	4/5/5/5	-
2	HEZ	F	702	-	-	1/5/5/5	-
2	HEZ	A	702	-	-	1/5/5/5	-
2	HEZ	E	702	-	-	1/5/5/5	-
2	HEZ	D	706	-	-	1/5/5/5	-
2	HEZ	L	701	-	-	2/5/5/5	-
2	HEZ	F	701	-	-	2/5/5/5	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 35 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	HEZ	C2-C3-C4-C5
2	D	707	HEZ	C3-C4-C5-C6
2	D	706	HEZ	C3-C4-C5-C6
2	I	701	HEZ	C1-C2-C3-C4
2	A	701	HEZ	C1-C2-C3-C4

There are no ring outliers.

14 monomers are involved in 36 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	701	HEZ	1	0
2	D	701	HEZ	6	0
2	H	701	HEZ	2	0
2	G	701	HEZ	4	0
2	A	701	HEZ	3	0
2	E	701	HEZ	3	0
2	D	703	HEZ	1	0
2	D	708	HEZ	4	0
2	I	701	HEZ	4	0
2	F	702	HEZ	1	0
2	E	702	HEZ	2	0
2	D	706	HEZ	1	0
2	L	701	HEZ	1	0
2	F	701	HEZ	3	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	608/622 (97%)	-0.67	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	30, 50, 75, 93	0
1	B	612/622 (98%)	-0.43	4 (0%) <span style="border: 1px solid blue; padding: 2px;">87</span> <span style="border: 1px solid blue; padding: 2px;">89</span>	38, 62, 89, 126	0
1	C	612/622 (98%)	-0.32	5 (0%) <span style="border: 1px solid blue; padding: 2px;">86</span> <span style="border: 1px solid blue; padding: 2px;">87</span>	47, 70, 98, 126	0
1	D	612/622 (98%)	-0.78	3 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">92</span>	23, 39, 62, 109	0
1	E	613/622 (98%)	-0.71	2 (0%) <span style="border: 1px solid blue; padding: 2px;">94</span> <span style="border: 1px solid blue; padding: 2px;">95</span>	27, 44, 67, 108	0
1	F	614/622 (98%)	-0.64	3 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">92</span>	26, 49, 73, 103	1 (0%)
1	G	608/622 (97%)	-0.40	1 (0%) <span style="border: 1px solid blue; padding: 2px;">95</span> <span style="border: 1px solid blue; padding: 2px;">96</span>	44, 70, 93, 139	0
1	H	609/622 (97%)	-0.29	3 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">92</span>	44, 75, 99, 126	0
1	I	611/622 (98%)	-0.47	2 (0%) <span style="border: 1px solid blue; padding: 2px;">94</span> <span style="border: 1px solid blue; padding: 2px;">95</span>	44, 64, 87, 124	0
1	J	609/622 (97%)	0.24	23 (3%) <span style="border: 1px solid red; padding: 2px;">40</span> <span style="border: 1px solid red; padding: 2px;">39</span>	65, 103, 130, 143	4 (0%)
1	K	610/622 (98%)	0.35	29 (4%) <span style="border: 1px solid red; padding: 2px;">30</span> <span style="border: 1px solid red; padding: 2px;">28</span>	71, 101, 133, 161	3 (0%)
1	L	611/622 (98%)	-0.19	6 (0%) <span style="border: 1px solid blue; padding: 2px;">82</span> <span style="border: 1px solid blue; padding: 2px;">83</span>	49, 77, 106, 139	0
All	All	7329/7464 (98%)	-0.36	81 (1%) <span style="border: 1px solid blue; padding: 2px;">80</span> <span style="border: 1px solid blue; padding: 2px;">82</span>	23, 65, 113, 161	8 (0%)

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	291	GLY	7.2
1	K	372	ILE	5.9
1	K	330	GLY	4.8
1	K	279	PHE	4.3
1	K	271	ASN	4.2

### 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 monosaccharides 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	MG	C	703	1/1	0.56	0.08	93,93,93,93	0
2	HEZ	E	702	8/8	0.68	0.42	67,69,76,77	0
2	HEZ	K	701	8/8	0.76	0.28	74,76,78,78	0
3	MG	D	709	1/1	0.81	0.17	63,63,63,63	0
3	MG	I	704	1/1	0.81	0.10	75,75,75,75	0
2	HEZ	C	701	8/8	0.83	0.22	64,71,73,74	0
2	HEZ	A	703	8/8	0.84	0.32	69,71,75,76	0
3	MG	G	703	1/1	0.84	0.11	82,82,82,82	0
2	HEZ	G	701	8/8	0.84	0.27	74,77,79,79	0
2	HEZ	D	706	8/8	0.86	0.26	63,64,65,65	0
2	HEZ	D	705	8/8	0.88	0.23	60,64,75,77	0
3	MG	F	704	1/1	0.88	0.09	80,80,80,80	0
3	MG	A	705	1/1	0.88	0.13	60,60,60,60	0
2	HEZ	B	701	8/8	0.88	0.22	50,56,57,59	0
2	HEZ	I	702	8/8	0.89	0.25	73,74,75,76	0
3	MG	H	702	1/1	0.89	0.06	81,81,81,81	0
2	HEZ	D	707	8/8	0.89	0.26	61,63,64,68	0
2	HEZ	I	701	8/8	0.90	0.20	66,70,72,72	0
2	HEZ	E	703	8/8	0.90	0.21	54,54,55,56	0
3	MG	H	703	1/1	0.91	0.14	41,41,41,41	0
2	HEZ	A	702	8/8	0.91	0.21	47,50,65,67	0
2	HEZ	D	703	8/8	0.92	0.17	63,65,67,68	0
2	HEZ	D	704	8/8	0.92	0.22	67,71,72,73	0
3	MG	E	706	1/1	0.92	0.08	44,44,44,44	0
2	HEZ	B	702	8/8	0.92	0.37	11,13,17,18	8
2	HEZ	D	702	8/8	0.93	0.34	10,14,16,16	8
2	HEZ	H	701	8/8	0.93	0.17	58,61,65,67	0
2	HEZ	L	701	8/8	0.93	0.17	57,60,62,62	0
2	HEZ	F	703	8/8	0.93	0.25	69,70,74,75	0
3	MG	L	703	1/1	0.93	0.16	66,66,66,66	0
3	MG	D	710	1/1	0.94	0.23	42,42,42,42	0
3	MG	E	705	1/1	0.94	0.18	29,29,29,29	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MG	L	702	1/1	0.94	0.15	80,80,80,80	0
2	HEZ	D	708	8/8	0.94	0.16	50,58,68,69	0
2	HEZ	F	701	8/8	0.95	0.15	43,45,50,50	0
2	HEZ	D	701	8/8	0.95	0.15	43,43,47,52	0
3	MG	E	704	1/1	0.95	0.09	58,58,58,58	0
2	HEZ	F	702	8/8	0.96	0.30	17,21,25,27	8
2	HEZ	E	701	8/8	0.96	0.14	36,40,47,47	0
3	MG	I	703	1/1	0.96	0.17	45,45,45,45	0
2	HEZ	A	701	8/8	0.96	0.14	47,48,49,49	0
3	MG	F	705	1/1	0.96	0.13	32,32,32,32	0
3	MG	C	702	1/1	0.96	0.15	60,60,60,60	0
3	MG	B	703	1/1	0.97	0.14	37,37,37,37	0
3	MG	G	702	1/1	0.98	0.07	34,34,34,34	0
3	MG	F	706	1/1	0.98	0.11	67,67,67,67	0
4	CL	D	711	1/1	0.98	0.09	47,47,47,47	0
4	CL	F	707	1/1	0.98	0.06	44,44,44,44	0
3	MG	A	704	1/1	0.99	0.14	36,36,36,36	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.