



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 22, 2016 – 09:10 AM EDT

PDB ID : 5FPH  
Title : The GTPase domains of the immunity-related Irga6 dimerize in a parallel head-to-head fashion  
Authors : Schulte, K.; Pawlowski, N.; Faelber, K.; Froehlich, C.; Howard, J.; Daumke, O.  
Deposited on : 2015-11-30  
Resolution : 3.20 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027107  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0122  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027107

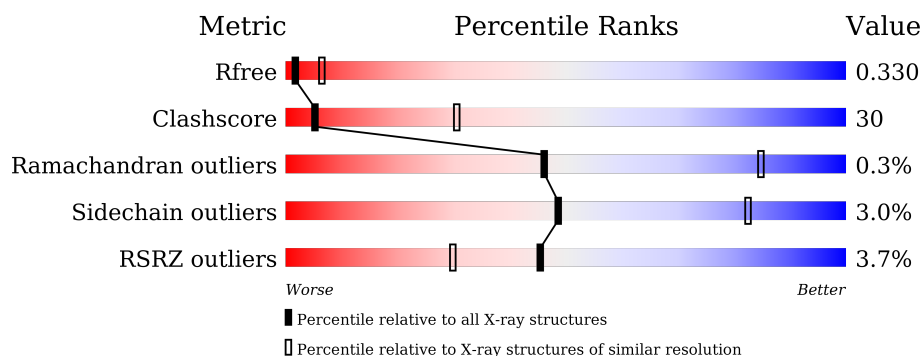
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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	423	
1	B	423	
1	C	423	
1	D	423	
1	E	423	
1	F	423	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	423	

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	GNP	A	1414	-	-	X	-
2	GNP	C	1414	-	-	X	-
2	GNP	E	1414	-	-	X	-
4	TRS	E	1417	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 22647 atoms, of which 24 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 INTERFERON-INDUCIBLE GTPASE 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	395	Total	C	N	O	S	Se	0	0	0
			3204	2062	522	606	4	10			
1	B	396	Total	C	N	O	S	Se	0	0	0
			3204	2066	518	606	4	10			
1	C	389	Total	C	N	O	S	Se	0	0	0
			3153	2033	514	592	4	10			
1	D	396	Total	C	N	O	S	Se	0	0	0
			3193	2051	521	607	4	10			
1	E	396	Total	C	N	O	S	Se	0	0	0
			3205	2065	521	605	4	10			
1	F	394	Total	C	N	O	S	Se	0	0	0
			3194	2055	521	604	4	10			
1	G	399	Total	C	N	O	S	Se	0	0	0
			3241	2086	527	614	4	10			

There are 98 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-9	GLY	-	EXPRESSION TAG	UNP Q9QZ85
A	-8	SER	-	EXPRESSION TAG	UNP Q9QZ85
A	-7	PRO	-	EXPRESSION TAG	UNP Q9QZ85
A	-6	GLY	-	EXPRESSION TAG	UNP Q9QZ85
A	-5	ILE	-	EXPRESSION TAG	UNP Q9QZ85
A	-4	PRO	-	EXPRESSION TAG	UNP Q9QZ85
A	-3	GLY	-	EXPRESSION TAG	UNP Q9QZ85
A	-2	SER	-	EXPRESSION TAG	UNP Q9QZ85
A	-1	THR	-	EXPRESSION TAG	UNP Q9QZ85
A	0	THR	-	EXPRESSION TAG	UNP Q9QZ85
A	31	GLU	ARG	ENGINEERED MUTATION	UNP Q9QZ85
A	32	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
A	176	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
A	246	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
B	-9	GLY	-	EXPRESSION TAG	UNP Q9QZ85

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	SER	-	EXPRESSION TAG	UNP Q9QZ85
B	-7	PRO	-	EXPRESSION TAG	UNP Q9QZ85
B	-6	GLY	-	EXPRESSION TAG	UNP Q9QZ85
B	-5	ILE	-	EXPRESSION TAG	UNP Q9QZ85
B	-4	PRO	-	EXPRESSION TAG	UNP Q9QZ85
B	-3	GLY	-	EXPRESSION TAG	UNP Q9QZ85
B	-2	SER	-	EXPRESSION TAG	UNP Q9QZ85
B	-1	THR	-	EXPRESSION TAG	UNP Q9QZ85
B	0	THR	-	EXPRESSION TAG	UNP Q9QZ85
B	31	GLU	ARG	ENGINEERED MUTATION	UNP Q9QZ85
B	32	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
B	176	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
B	246	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
C	-9	GLY	-	EXPRESSION TAG	UNP Q9QZ85
C	-8	SER	-	EXPRESSION TAG	UNP Q9QZ85
C	-7	PRO	-	EXPRESSION TAG	UNP Q9QZ85
C	-6	GLY	-	EXPRESSION TAG	UNP Q9QZ85
C	-5	ILE	-	EXPRESSION TAG	UNP Q9QZ85
C	-4	PRO	-	EXPRESSION TAG	UNP Q9QZ85
C	-3	GLY	-	EXPRESSION TAG	UNP Q9QZ85
C	-2	SER	-	EXPRESSION TAG	UNP Q9QZ85
C	-1	THR	-	EXPRESSION TAG	UNP Q9QZ85
C	0	THR	-	EXPRESSION TAG	UNP Q9QZ85
C	31	GLU	ARG	ENGINEERED MUTATION	UNP Q9QZ85
C	32	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
C	176	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
C	246	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
D	-9	GLY	-	EXPRESSION TAG	UNP Q9QZ85
D	-8	SER	-	EXPRESSION TAG	UNP Q9QZ85
D	-7	PRO	-	EXPRESSION TAG	UNP Q9QZ85
D	-6	GLY	-	EXPRESSION TAG	UNP Q9QZ85
D	-5	ILE	-	EXPRESSION TAG	UNP Q9QZ85
D	-4	PRO	-	EXPRESSION TAG	UNP Q9QZ85
D	-3	GLY	-	EXPRESSION TAG	UNP Q9QZ85
D	-2	SER	-	EXPRESSION TAG	UNP Q9QZ85
D	-1	THR	-	EXPRESSION TAG	UNP Q9QZ85
D	0	THR	-	EXPRESSION TAG	UNP Q9QZ85
D	31	GLU	ARG	ENGINEERED MUTATION	UNP Q9QZ85
D	32	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
D	176	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
D	246	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
E	-9	GLY	-	EXPRESSION TAG	UNP Q9QZ85

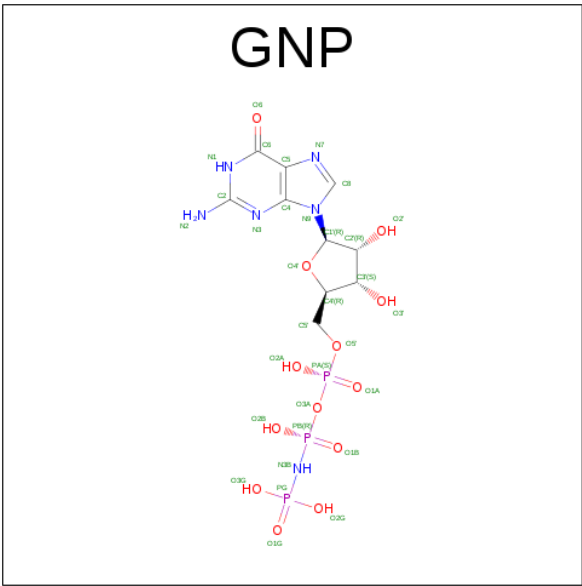
*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-8	SER	-	EXPRESSION TAG	UNP Q9QZ85
E	-7	PRO	-	EXPRESSION TAG	UNP Q9QZ85
E	-6	GLY	-	EXPRESSION TAG	UNP Q9QZ85
E	-5	ILE	-	EXPRESSION TAG	UNP Q9QZ85
E	-4	PRO	-	EXPRESSION TAG	UNP Q9QZ85
E	-3	GLY	-	EXPRESSION TAG	UNP Q9QZ85
E	-2	SER	-	EXPRESSION TAG	UNP Q9QZ85
E	-1	THR	-	EXPRESSION TAG	UNP Q9QZ85
E	0	THR	-	EXPRESSION TAG	UNP Q9QZ85
E	31	GLU	ARG	ENGINEERED MUTATION	UNP Q9QZ85
E	32	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
E	176	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
E	246	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
F	-9	GLY	-	EXPRESSION TAG	UNP Q9QZ85
F	-8	SER	-	EXPRESSION TAG	UNP Q9QZ85
F	-7	PRO	-	EXPRESSION TAG	UNP Q9QZ85
F	-6	GLY	-	EXPRESSION TAG	UNP Q9QZ85
F	-5	ILE	-	EXPRESSION TAG	UNP Q9QZ85
F	-4	PRO	-	EXPRESSION TAG	UNP Q9QZ85
F	-3	GLY	-	EXPRESSION TAG	UNP Q9QZ85
F	-2	SER	-	EXPRESSION TAG	UNP Q9QZ85
F	-1	THR	-	EXPRESSION TAG	UNP Q9QZ85
F	0	THR	-	EXPRESSION TAG	UNP Q9QZ85
F	31	GLU	ARG	ENGINEERED MUTATION	UNP Q9QZ85
F	32	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
F	176	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
F	246	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
G	-9	GLY	-	EXPRESSION TAG	UNP Q9QZ85
G	-8	SER	-	EXPRESSION TAG	UNP Q9QZ85
G	-7	PRO	-	EXPRESSION TAG	UNP Q9QZ85
G	-6	GLY	-	EXPRESSION TAG	UNP Q9QZ85
G	-5	ILE	-	EXPRESSION TAG	UNP Q9QZ85
G	-4	PRO	-	EXPRESSION TAG	UNP Q9QZ85
G	-3	GLY	-	EXPRESSION TAG	UNP Q9QZ85
G	-2	SER	-	EXPRESSION TAG	UNP Q9QZ85
G	-1	THR	-	EXPRESSION TAG	UNP Q9QZ85
G	0	THR	-	EXPRESSION TAG	UNP Q9QZ85
G	31	GLU	ARG	ENGINEERED MUTATION	UNP Q9QZ85
G	32	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
G	176	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85
G	246	GLU	LYS	ENGINEERED MUTATION	UNP Q9QZ85

- Molecule 2 is PHOSPHOAMINOPHOSPHONIC ACID-GUANYLATE ESTER (three-letter

code: GNP) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>6</sub>O<sub>13</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
2	B	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
2	C	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
2	E	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
2	F	1	Total	C	N	O	P	0	0
			32	10	6	13	3		
2	G	1	Total	C	N	O	P	0	0
			32	10	6	13	3		

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

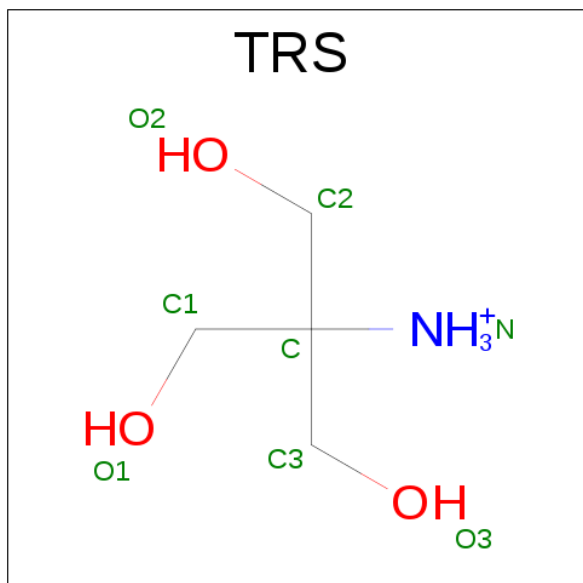
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	2	Total	Mg	0	0
			2	2		
3	D	1	Total	Mg	0	0
			1	1		
3	E	2	Total	Mg	0	0
			2	2		
3	B	2	Total	Mg	0	0
			2	2		
3	C	2	Total	Mg	0	0
			2	2		

Continued on next page...

Continued from previous page...

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

- Molecule 4 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula:  $C_4H_{12}NO_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	C	1	Total	C	H	N	O	0	0
			20	4	12	1	3		
4	E	1	Total	C	H	N	O	0	0
			20	4	12	1	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	O	0	0
			1	1		
5	B	2	Total	O	0	0
			2	2		
5	C	1	Total	O	0	0
			1	1		
5	D	1	Total	O	0	0
			1	1		
5	F	2	Total	O	0	0
			2	2		

Continued on next page...



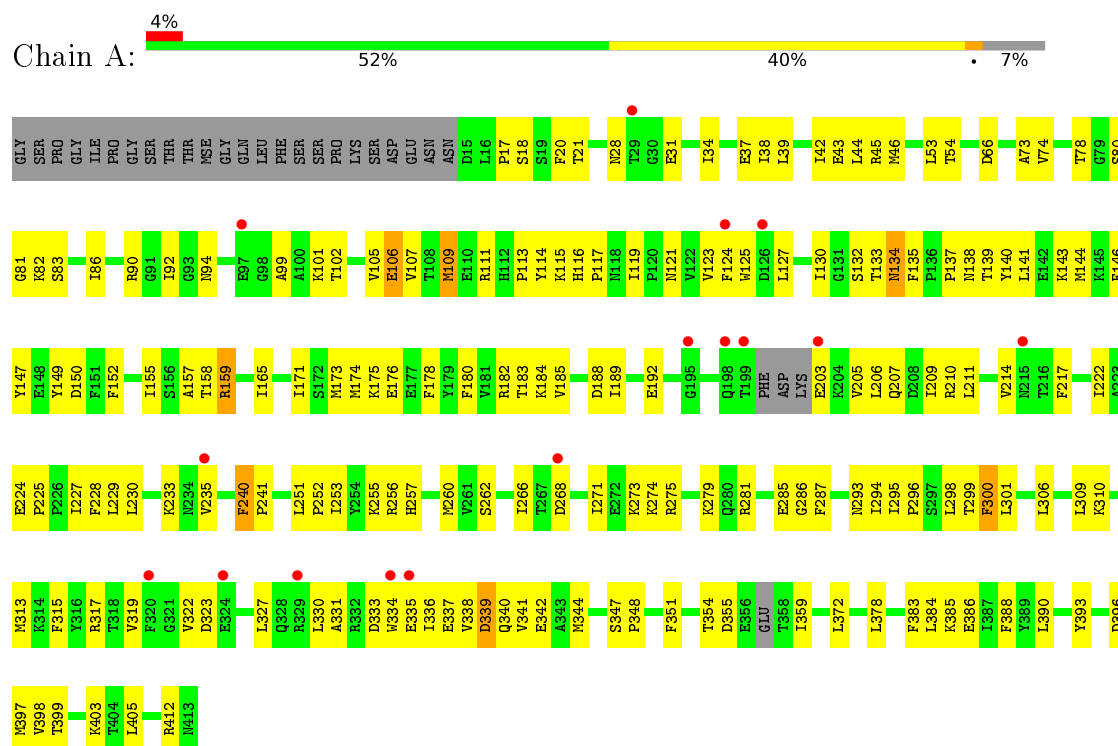
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	G	1	Total	O	0	0
			1	1		

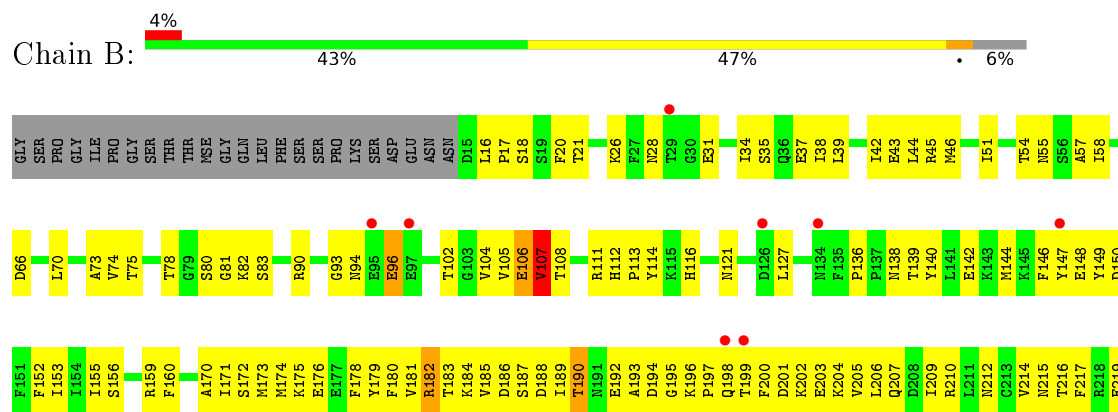
### 3 Residue-property plots

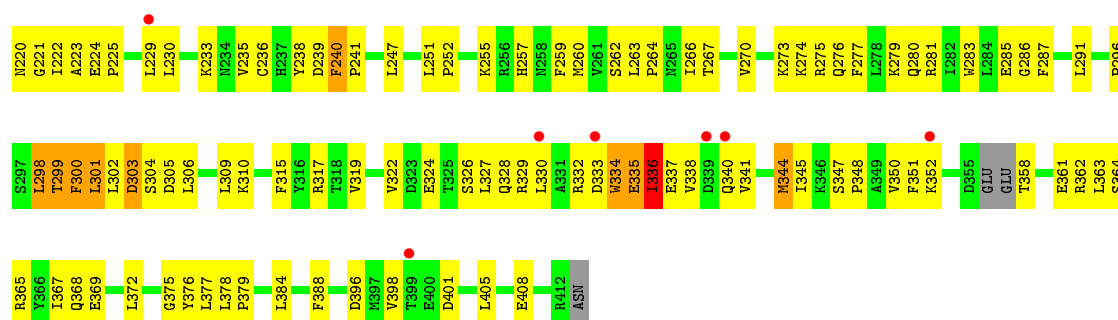
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: INTERFERON-INDUCIBLE GTPASE 1

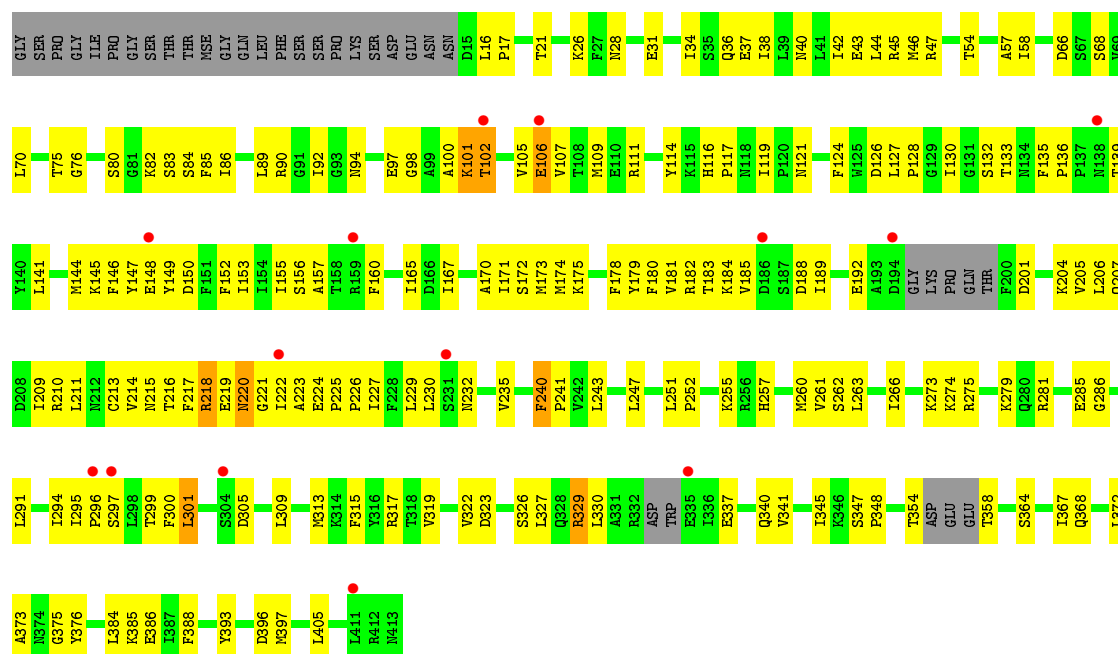


#### • Molecule 1: INTERFERON-INDUCIBLE GTPASE 1

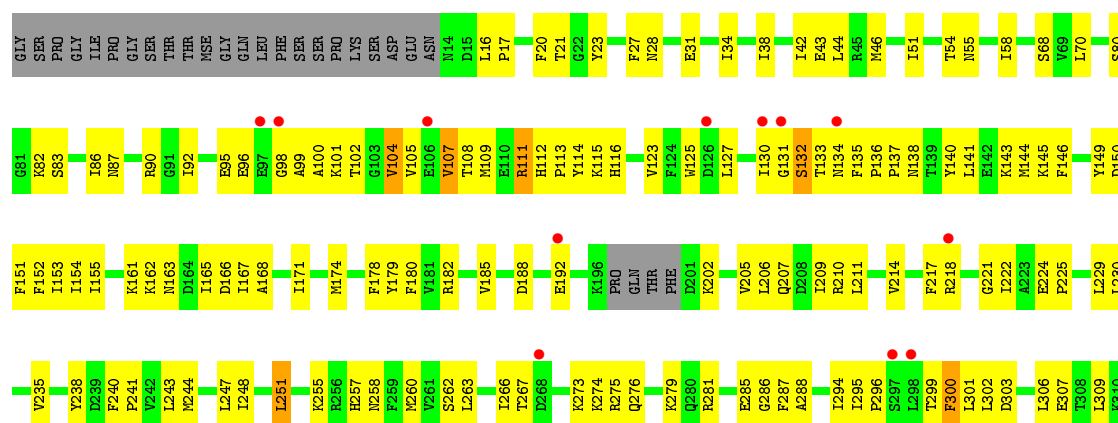


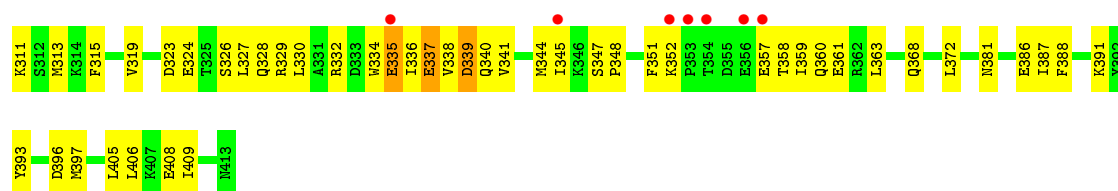


• Molecule 1: INTERFERON-INDUCIBLE GTPASE 1

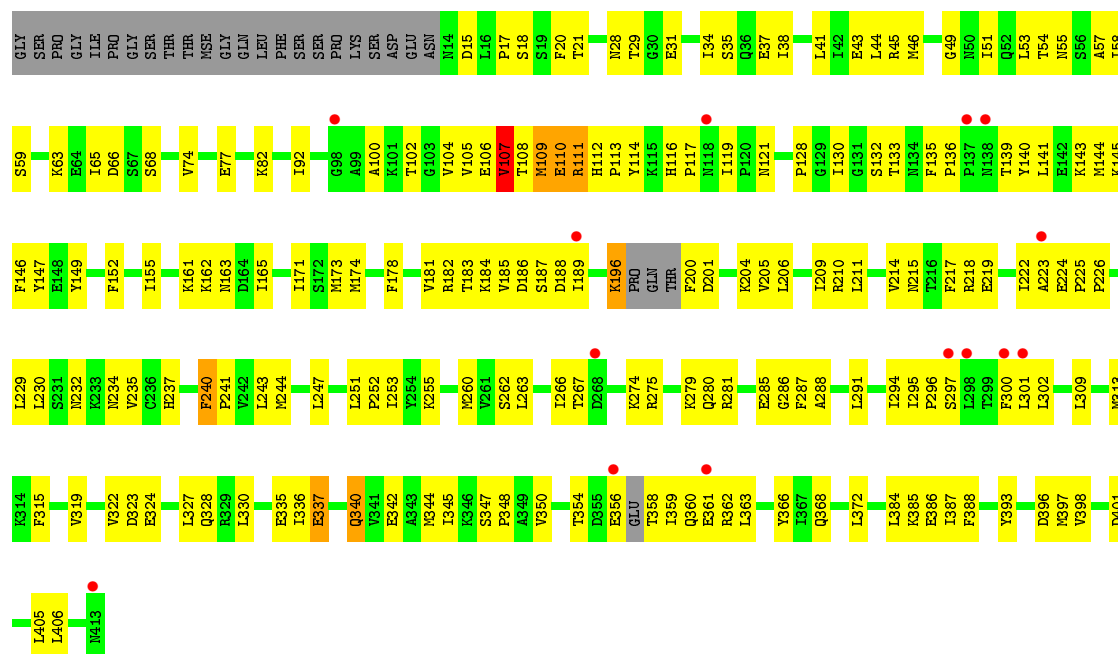


• Molecule 1: INTERFERON-INDUCIBLE GTPASE 1

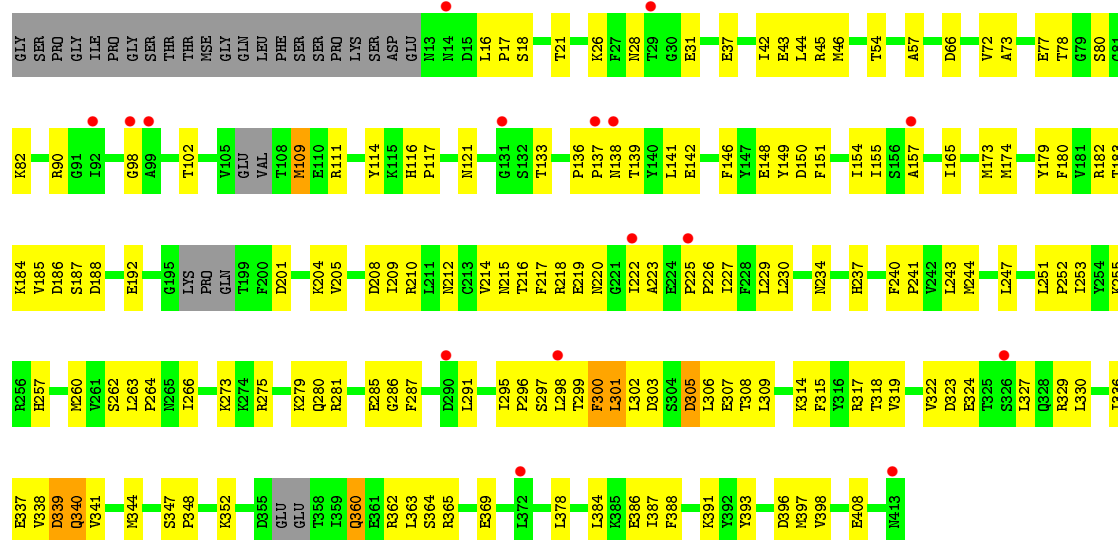




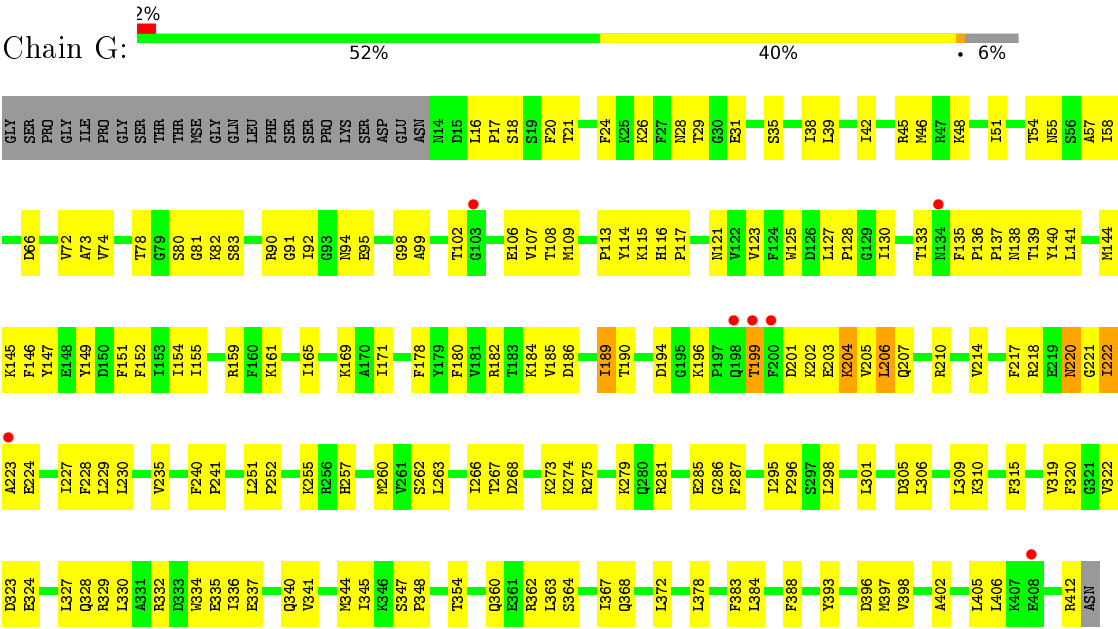
• Molecule 1: INTERFERON-INDUCIBLE GTPASE 1



• Molecule 1: INTERFERON-INDUCIBLE GTPASE 1



• Molecule 1: INTERFERON-INDUCIBLE GTPASE 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	98.51Å 98.51Å 1289.40Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	68.98 – 3.20 68.98 – 3.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (68.98-3.20) 100.0 (68.98-3.20)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.00 (at 3.19Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.297 , 0.317 0.310 , 0.330	Depositor DCC
$R_{free}$ test set	3215 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	75.4	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 9.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.36$ , $\langle L^2 \rangle = 0.19$	Xtriage
Outliers	0 of 64300 reflections	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	22647	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.71% 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: TRS, MG, GNP

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	0.24	0/3258	0.43	0/4383
1	B	0.24	0/3259	0.45	0/4386
1	C	0.24	0/3204	0.42	0/4306
1	D	0.24	0/3245	0.43	0/4366
1	E	0.23	0/3259	0.41	0/4383
1	F	0.22	0/3247	0.41	0/4367
1	G	0.23	0/3298	0.41	0/4439
All	All	0.23	0/22770	0.42	0/30630

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	3204	0	3214	218	0
1	B	3204	0	3213	267	0
1	C	3153	0	3172	208	0
1	D	3193	0	3195	238	0
1	E	3205	0	3204	176	0
1	F	3194	0	3191	176	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	3241	0	3247	158	0
2	A	32	0	13	18	0
2	B	32	0	13	8	0
2	C	32	0	13	11	0
2	E	32	0	13	9	0
2	F	32	0	13	5	0
2	G	32	0	13	8	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	1	0	0	0	0
3	E	2	0	0	0	0
3	F	2	0	0	0	0
3	G	2	0	0	0	0
4	C	8	12	12	1	0
4	E	8	12	12	1	0
5	A	1	0	0	1	0
5	B	2	0	0	1	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	F	2	0	0	0	0
5	G	1	0	0	0	0
All	All	22623	24	22538	1372	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 30.

The worst 5 of 1372 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:73:ALA:HB1	1:A:127:LEU:HD11	1.23	1.11
1:C:173:MSE:HE2	1:D:141:LEU:HB3	1.26	1.10
1:B:78:THR:HG21	1:B:106:GLU:HG2	1.29	1.08
1:G:186:ASP:HA	1:G:189:ILE:HD11	1.36	1.06
1:C:181:VAL:HG13	1:C:230:LEU:HD21	1.40	1.04

There are no symmetry-related clashes.



## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	389/423 (92%)	373 (96%)	16 (4%)	0	100	100
1	B	392/423 (93%)	378 (96%)	12 (3%)	2 (0%)	34	78
1	C	381/423 (90%)	369 (97%)	12 (3%)	0	100	100
1	D	392/423 (93%)	371 (95%)	18 (5%)	3 (1%)	24	69
1	E	390/423 (92%)	377 (97%)	12 (3%)	1 (0%)	46	85
1	F	386/423 (91%)	369 (96%)	17 (4%)	0	100	100
1	G	397/423 (94%)	376 (95%)	19 (5%)	2 (0%)	34	78
All	All	2727/2961 (92%)	2613 (96%)	106 (4%)	8 (0%)	46	85

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	336	ILE
1	D	335	GLU
1	G	222	ILE
1	D	104	VAL
1	D	107	VAL

### 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	361/374 (96%)	350 (97%)	11 (3%)	48	82
1	B	360/374 (96%)	343 (95%)	17 (5%)	32	73

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	355/374 (95%)	345 (97%)	10 (3%)	51	84
1	D	358/374 (96%)	351 (98%)	7 (2%)	63	88
1	E	359/374 (96%)	350 (98%)	9 (2%)	55	86
1	F	359/374 (96%)	348 (97%)	11 (3%)	47	82
1	G	365/374 (98%)	355 (97%)	10 (3%)	52	85
All	All	2517/2618 (96%)	2442 (97%)	75 (3%)	48	82

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	240	PHE
1	D	300	PHE
1	G	202	LYS
1	C	300	PHE
1	D	111	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	94	ASN
1	C	40	ASN
1	C	87	ASN
1	D	360	GLN
1	F	212	ASN

### 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

Of 21 ligands modelled in this entry, 13 are monoatomic - leaving 8 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	GNP	A	1414	3	29,34,34	2.65	11 (37%)	28,54,54	2.17	8 (28%)
2	GNP	B	1413	3	29,34,34	2.70	10 (34%)	28,54,54	1.95	7 (25%)
2	GNP	C	1414	3	29,34,34	2.75	12 (41%)	28,54,54	2.15	7 (25%)
4	TRS	C	1417	-	7,7,7	1.05	1 (14%)	9,9,9	0.94	0
2	GNP	E	1414	3	29,34,34	2.67	10 (34%)	28,54,54	1.93	7 (25%)
4	TRS	E	1417	-	7,7,7	1.19	1 (14%)	9,9,9	0.48	0
2	GNP	F	1414	3	29,34,34	2.71	10 (34%)	28,54,54	1.94	7 (25%)
2	GNP	G	1413	3	29,34,34	2.72	12 (41%)	28,54,54	1.91	7 (25%)

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	GNP	A	1414	3	-	0/16/38/38	0/3/3/3
2	GNP	B	1413	3	-	1/16/38/38	0/3/3/3
2	GNP	C	1414	3	-	0/16/38/38	0/3/3/3
4	TRS	C	1417	-	-	0/9/9/9	0/0/0/0
2	GNP	E	1414	3	-	0/16/38/38	0/3/3/3
4	TRS	E	1417	-	-	0/9/9/9	0/0/0/0
2	GNP	F	1414	3	-	0/16/38/38	0/3/3/3
2	GNP	G	1413	3	-	0/16/38/38	0/3/3/3

The worst 5 of 67 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1414	GNP	C4-N9	-7.45	1.37	1.47
2	G	1413	GNP	C4-N9	-7.35	1.37	1.47

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1413	GNP	C4-N9	-7.32	1.38	1.47
2	A	1414	GNP	C4-N9	-7.26	1.38	1.47
2	F	1414	GNP	C4-N9	-7.26	1.38	1.47

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1414	GNP	PA-O3A-PB	-4.45	116.58	132.71
2	A	1414	GNP	PA-O3A-PB	-4.02	118.14	132.71
2	B	1413	GNP	O3G-PG-O1G	-3.57	104.19	113.58
2	G	1413	GNP	O3G-PG-O1G	-3.33	104.80	113.58
2	E	1414	GNP	O3G-PG-O1G	-3.28	104.95	113.58

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1413	GNP	O1B-PB-N3B-PG

There are no ring outliers.

8 monomers are involved in 59 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1414	GNP	18	0
2	B	1413	GNP	8	0
2	C	1414	GNP	11	0
4	C	1417	TRS	1	0
2	E	1414	GNP	9	0
4	E	1417	TRS	1	0
2	F	1414	GNP	5	0
2	G	1413	GNP	8	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	385/423 (91%)	0.22	16 (4%)	40	26	39, 68, 108, 144	0
1	B	386/423 (91%)	0.30	15 (3%)	43	28	38, 67, 104, 126	0
1	C	379/423 (89%)	0.28	14 (3%)	45	30	38, 64, 103, 126	0
1	D	386/423 (91%)	0.36	19 (4%)	33	20	38, 67, 112, 137	0
1	E	386/423 (91%)	0.19	14 (3%)	46	31	36, 63, 99, 135	0
1	F	384/423 (90%)	0.31	16 (4%)	40	26	37, 65, 98, 111	0
1	G	389/423 (91%)	0.17	7 (1%)	71	58	33, 59, 95, 147	0
All	All	2695/2961 (91%)	0.26	101 (3%)	45	30	33, 65, 103, 147	0

The worst 5 of 101 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	98	GLY	6.1
1	D	131	GLY	6.0
1	C	297	SER	5.1
1	D	134	ASN	4.9
1	C	194	ASP	4.7

### 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(Å <sup>2</sup> )	Q<0.9
4	TRS	E	1417	8/8	0.55	0.28	2.01	79,109,138,138	0
2	GNP	E	1414	32/32	0.87	0.25	0.30	34,67,79,88	32
2	GNP	A	1414	32/32	0.69	0.24	-0.15	60,80,100,116	32
2	GNP	G	1413	32/32	0.90	0.21	-0.28	44,59,74,79	32
2	GNP	B	1413	32/32	0.88	0.23	-0.64	56,69,88,95	32
2	GNP	C	1414	32/32	0.89	0.21	-0.82	46,63,84,91	32
2	GNP	F	1414	32/32	0.90	0.18	-1.10	49,66,77,80	32
4	TRS	C	1417	8/8	0.81	0.21	-1.57	81,99,142,142	0
3	MG	G	1414	1/1	0.94	0.08	-2.50	43,43,43,43	0
3	MG	F	1416	1/1	0.89	0.20	-	36,36,36,36	0
3	MG	E	1416	1/1	0.96	0.20	-	20,20,20,20	0
3	MG	F	1415	1/1	0.95	0.18	-	71,71,71,71	0
3	MG	A	1416	1/1	0.73	0.23	-	38,38,38,38	0
3	MG	B	1414	1/1	0.81	0.51	-	110,110,110,110	0
3	MG	C	1416	1/1	0.97	0.20	-	30,30,30,30	0
3	MG	E	1415	1/1	0.90	0.15	-	44,44,44,44	0
3	MG	C	1415	1/1	0.99	0.11	-	68,68,68,68	0
3	MG	A	1415	1/1	0.94	0.49	-	91,91,91,91	0
3	MG	G	1415	1/1	0.96	0.18	-	43,43,43,43	0
3	MG	D	1414	1/1	0.96	0.28	-	32,32,32,32	0
3	MG	B	1415	1/1	0.90	0.23	-	44,44,44,44	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.