



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 06:03 pm GMT

PDB ID : 1N4P  
Title : Protein Geranylgeranyltransferase type-I Complexed with Geranylgeranyl Diphosphate  
Authors : Taylor, J.S.; Reid, T.S.; Casey, P.J.; Beese, L.S.  
Deposited on : 2002-11-01  
Resolution : 2.65 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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

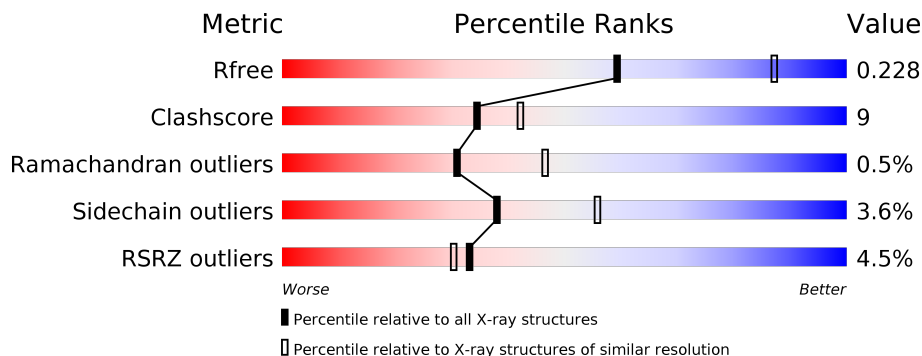
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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))
$R_{free}$	111664	1112 (2.68-2.64)
Clashscore	122126	1151 (2.68-2.64)
Ramachandran outliers	120053	1133 (2.68-2.64)
Sidechain outliers	120020	1133 (2.68-2.64)
RSRZ outliers	108989	1098 (2.68-2.64)

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	377	<div> <div>4%</div> <div> <div></div> <div>67%</div> <div>15%</div> <div>•</div> <div>17%</div> </div> </div>
1	C	377	<div> <div>3%</div> <div> <div></div> <div>66%</div> <div>17%</div> <div>•</div> <div>17%</div> </div> </div>
1	E	377	<div> <div>4%</div> <div> <div></div> <div>64%</div> <div>17%</div> <div>•</div> <div>17%</div> </div> </div>
1	G	377	<div> <div>3%</div> <div> <div></div> <div>63%</div> <div>20%</div> <div>•</div> <div>17%</div> </div> </div>
1	I	377	<div> <div>3%</div> <div> <div></div> <div>64%</div> <div>18%</div> <div>•</div> <div>17%</div> </div> </div>
1	K	377	<div> <div>2%</div> <div> <div></div> <div>69%</div> <div>14%</div> <div>•</div> <div>17%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	377	<div><div><div>5%</div><div>72%</div><div>19%</div><div>8%</div></div></div>
2	D	377	<div><div><div>4%</div><div>72%</div><div>17%</div><div>8%</div></div></div>
2	F	377	<div><div><div>4%</div><div>76%</div><div>15%</div><div>8%</div></div></div>
2	H	377	<div><div><div>8%</div><div>69%</div><div>21%</div><div>8%</div></div></div>
2	J	377	<div><div><div>4%</div><div>69%</div><div>21%</div><div>8%</div></div></div>
2	L	377	<div><div><div>2%</div><div>73%</div><div>16%</div><div>8%</div></div></div>
3	M	11	<div><div><div>27%</div><div>9%</div><div>27%</div><div>64%</div></div></div>
3	N	11	<div><div><div>18%</div><div>9%</div><div>27%</div><div>64%</div></div></div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 33443 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 farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	314	Total	C	N	O	S	0	0	0
			2629	1679	463	482	5			
1	C	314	Total	C	N	O	S	0	0	0
			2643	1689	461	488	5			
1	E	314	Total	C	N	O	S	0	0	0
			2642	1686	461	490	5			
1	G	314	Total	C	N	O	S	0	0	0
			2633	1683	459	486	5			
1	I	314	Total	C	N	O	S	0	0	0
			2656	1694	465	492	5			
1	K	314	Total	C	N	O	S	0	0	0
			2671	1703	467	496	5			

- Molecule 2 is a protein called Geranylgeranyl transferase type-1 subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	346	Total	C	N	O	S	0	0	0
			2697	1707	467	499	24			
2	D	346	Total	C	N	O	S	0	0	0
			2713	1715	472	502	24			
2	F	346	Total	C	N	O	S	0	0	0
			2718	1717	474	503	24			
2	H	346	Total	C	N	O	S	0	0	0
			2694	1706	464	500	24			
2	J	346	Total	C	N	O	S	0	0	0
			2711	1713	471	503	24			
2	L	346	Total	C	N	O	S	0	0	0
			2723	1720	473	506	24			

- Molecule 3 is a protein called Fusion protein consisting of transforming protein p21b and Ras related protein Rap-2b.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	M	4	Total	C	N	O	S	0	0	0
			30	20	4	5	1			
3	N	4	Total	C	N	O	S	0	0	0
			30	20	4	5	1			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

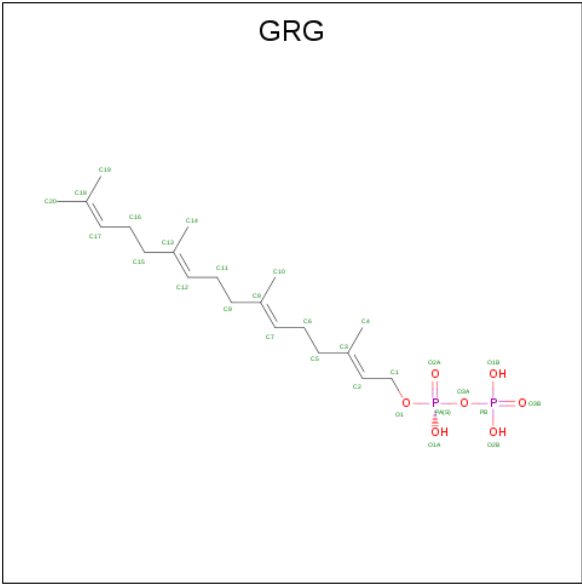
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	J	1	Total	Zn	0	0
			1	1		
4	D	1	Total	Zn	0	0
			1	1		
4	H	1	Total	Zn	0	0
			1	1		
4	B	1	Total	Zn	0	0
			1	1		
4	L	1	Total	Zn	0	0
			1	1		
4	F	1	Total	Zn	0	0
			1	1		

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	G	1	Total	Cl	0	0
			1	1		
5	J	1	Total	Cl	0	0
			1	1		
5	D	1	Total	Cl	0	0
			1	1		
5	K	1	Total	Cl	0	0
			1	1		
5	H	1	Total	Cl	0	0
			1	1		
5	B	1	Total	Cl	0	0
			1	1		
5	C	1	Total	Cl	0	0
			1	1		
5	L	1	Total	Cl	0	0
			1	1		
5	F	1	Total	Cl	0	0
			1	1		

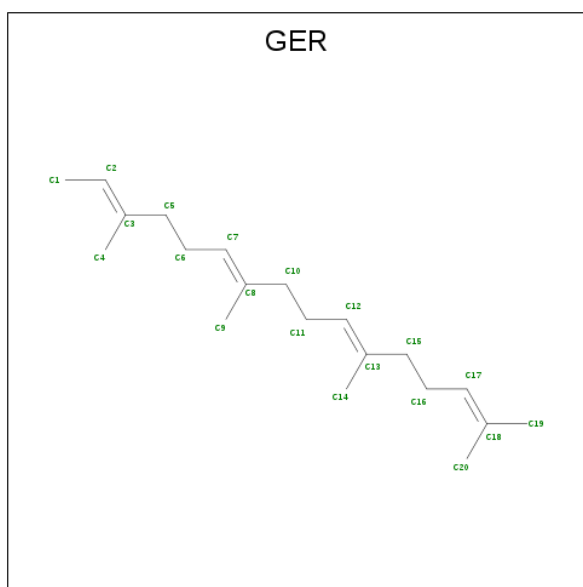
- Molecule 6 is GERANYLGERANYL DIPHOSPHATE (three-letter code: GRG) (formula:

C<sub>20</sub>H<sub>36</sub>O<sub>7</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total	C	O	P	0	0
			29	20	7	2		
6	D	1	Total	C	O	P	0	0
			29	20	7	2		
6	F	1	Total	C	O	P	0	0
			29	20	7	2		
6	H	1	Total	C	O	P	0	0
			29	20	7	2		
6	J	1	Total	C	O	P	0	0
			29	20	7	2		
6	L	1	Total	C	O	P	0	0
			29	20	7	2		

- Molecule 7 is GERAN-8-YL GERAN (three-letter code: GER) (formula: C<sub>20</sub>H<sub>34</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	M	1	Total C 20 20	0	0
7	N	1	Total C 20 20	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	71	Total O 71 71	0	0
8	B	63	Total O 63 63	0	0
8	C	80	Total O 80 80	0	0
8	D	94	Total O 94 94	0	0
8	E	64	Total O 64 64	0	0
8	F	92	Total O 92 92	0	0
8	G	58	Total O 58 58	0	0
8	H	48	Total O 48 48	0	0
8	I	87	Total O 87 87	0	0
8	J	73	Total O 73 73	0	0

*Continued on next page...*

*Continued from previous page...*

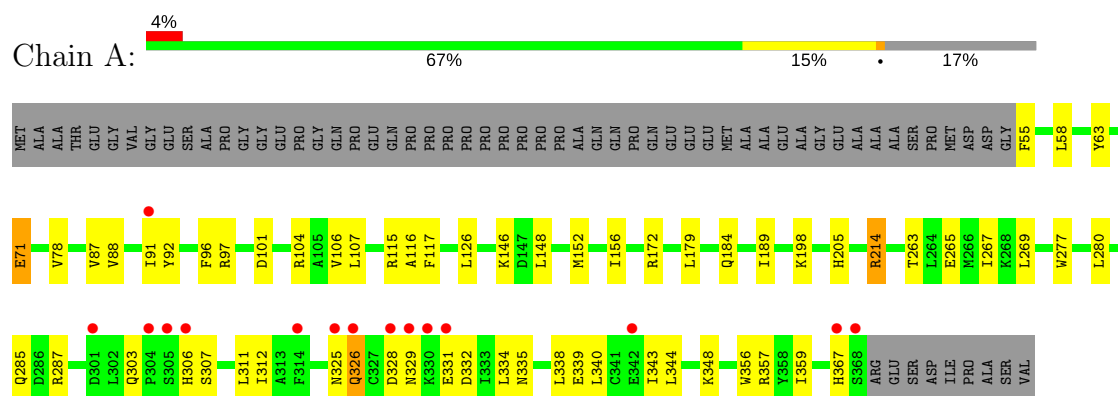
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	K	159	Total 159	O 159	0	0
8	L	134	Total 134	O 134	0	0
8	N	1	Total 1	O 1	0	0



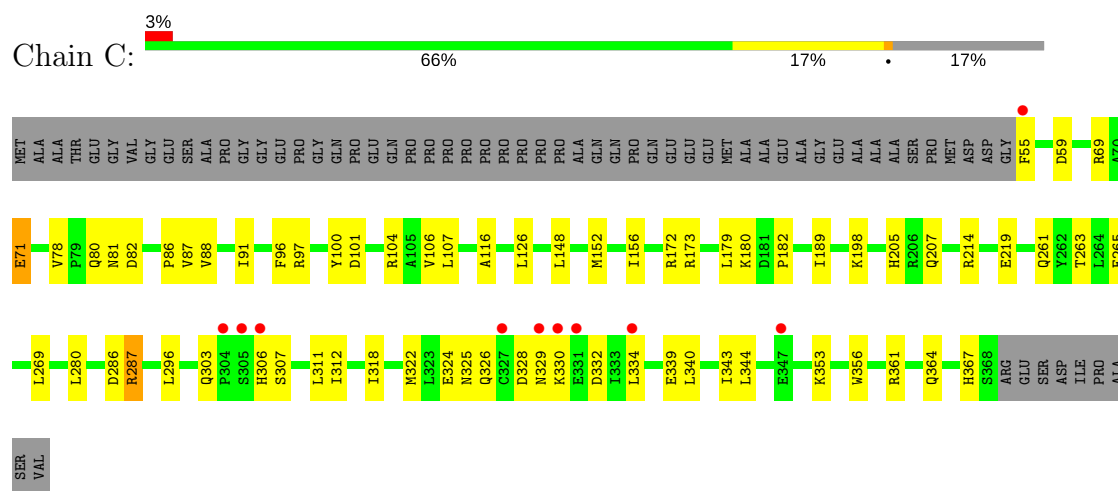
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

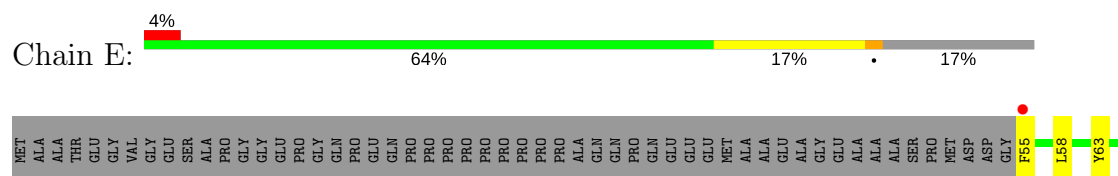
- Molecule 1: Protein farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha

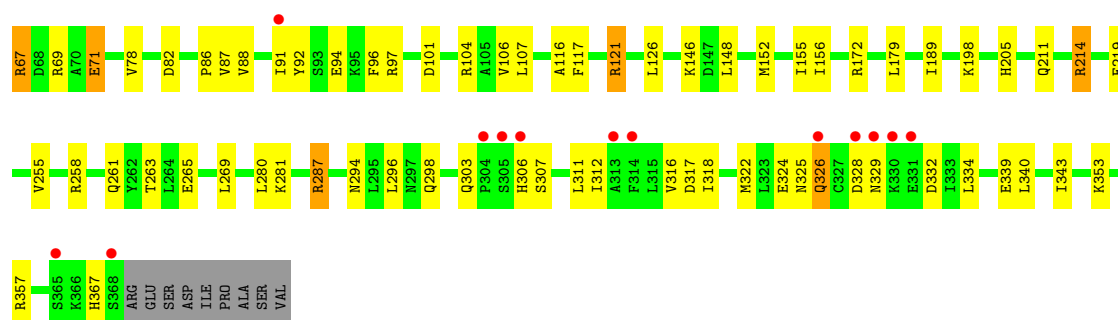


- Molecule 1: Protein farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha

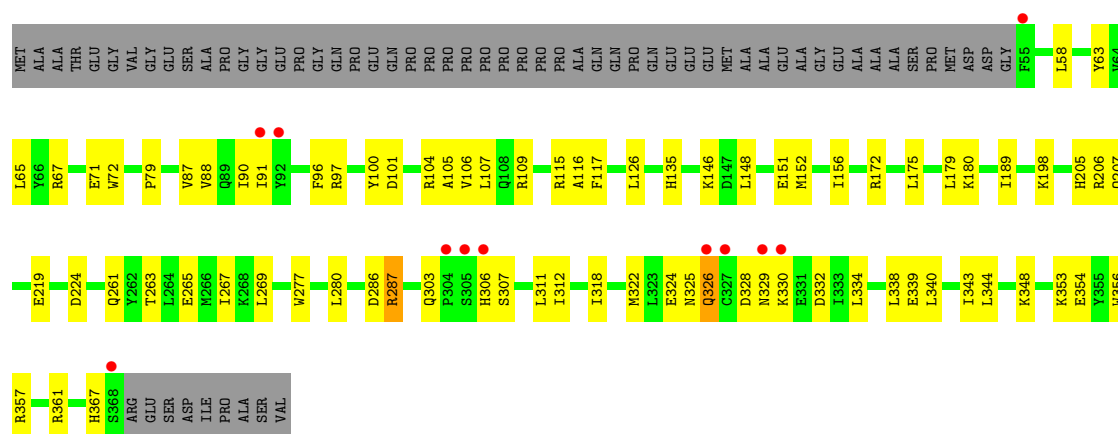


- Molecule 1: Protein farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha

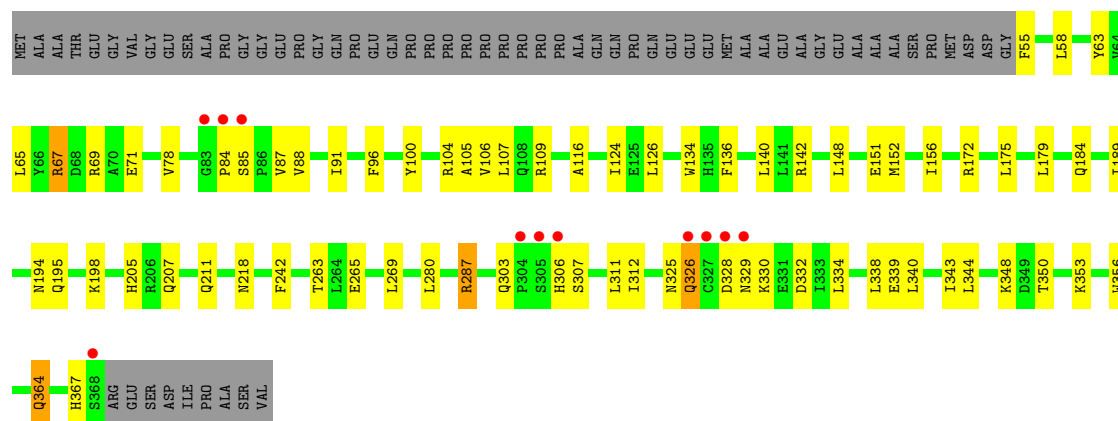




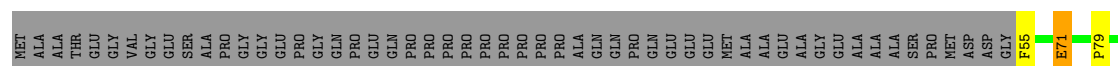
- Molecule 1: Protein farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha

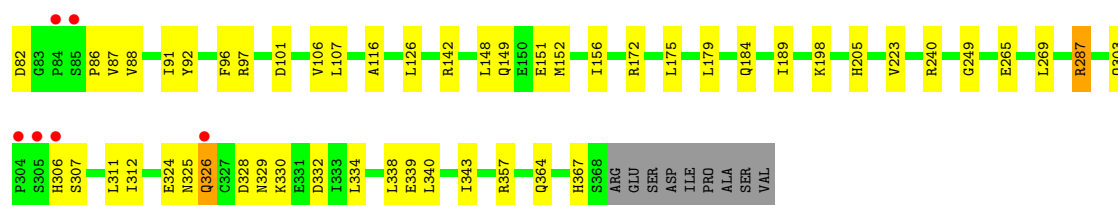


- Molecule 1: Protein farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha

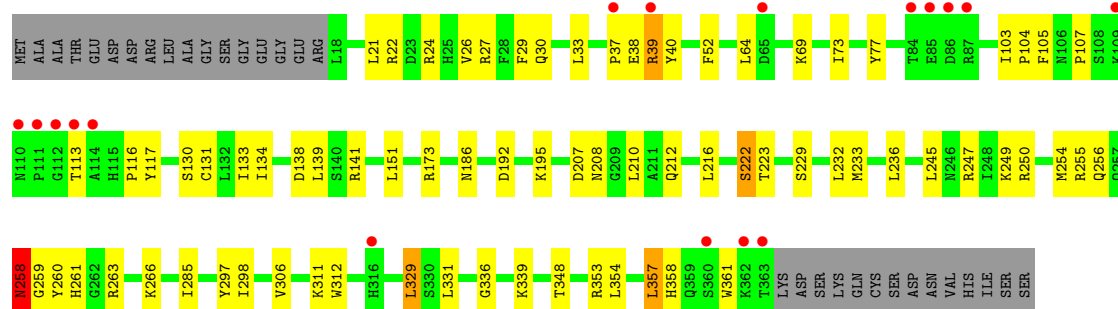
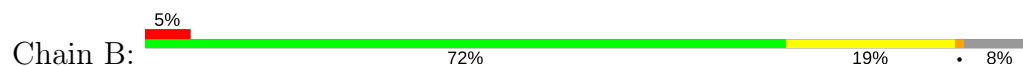


- Molecule 1: Protein farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha

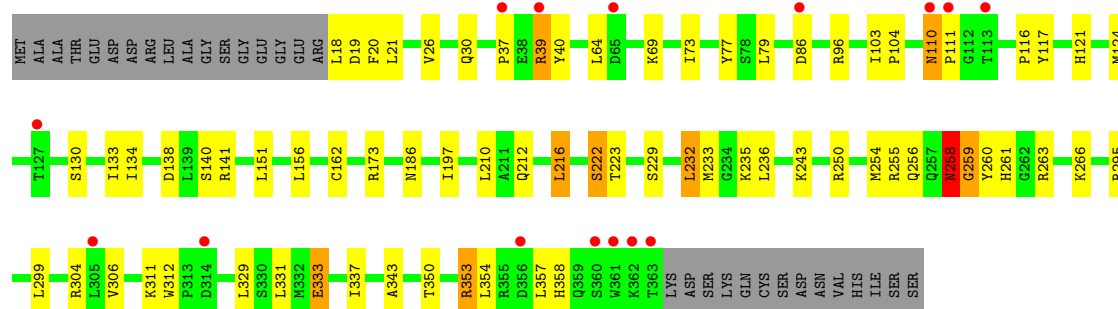




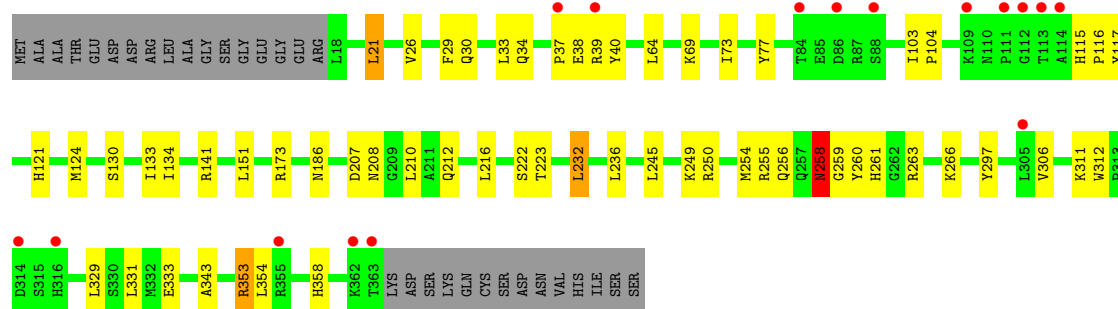
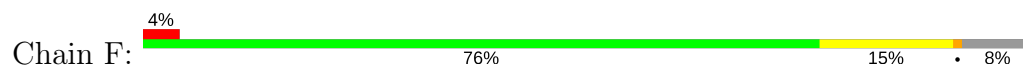
• Molecule 2: Geranylgeranyl transferase type-1 subunit beta



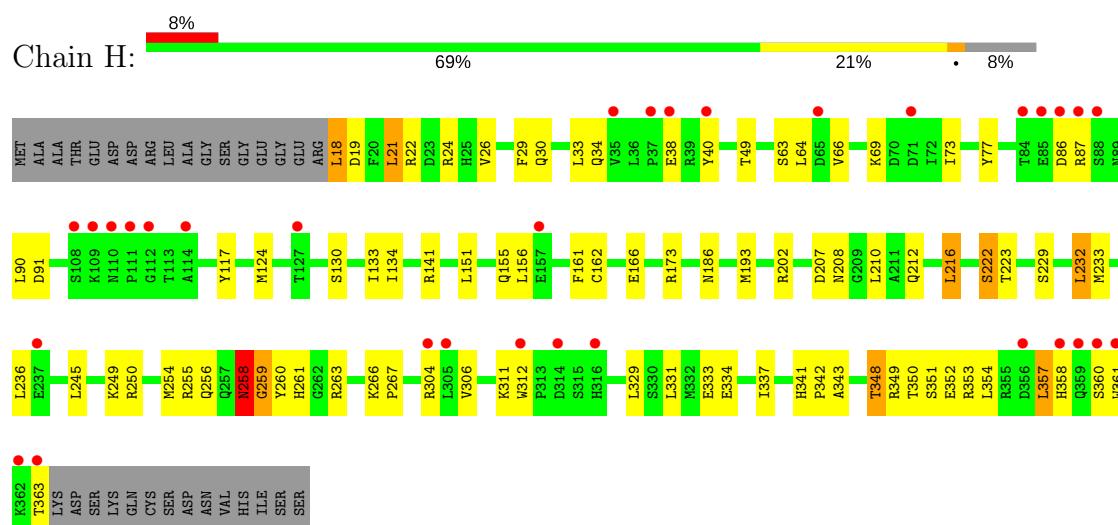
• Molecule 2: Geranylgeranyl transferase type-1 subunit beta



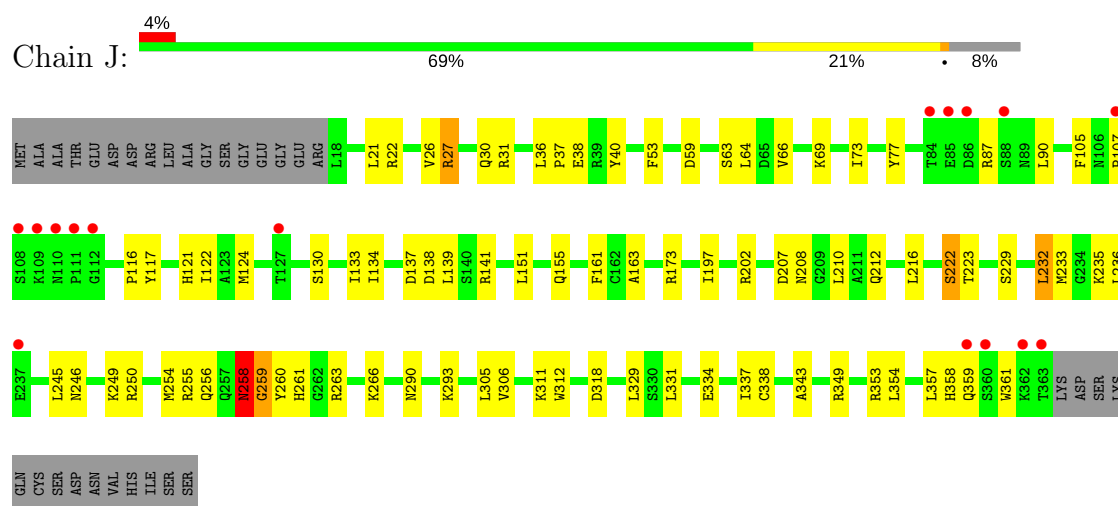
• Molecule 2: Geranylgeranyl transferase type-1 subunit beta



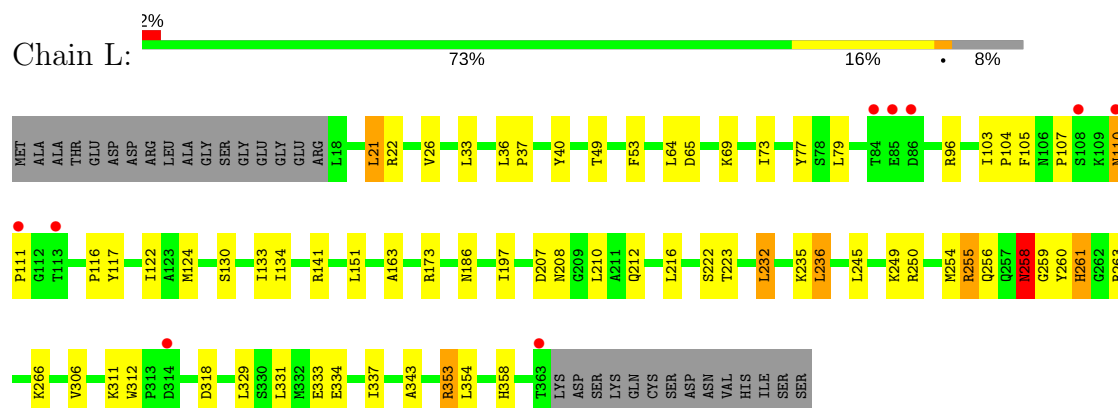
• Molecule 2: Geranylgeranyl transferase type-1 subunit beta



- Molecule 2: Geranylgeranyl transferase type-1 subunit beta

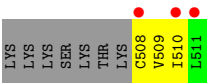


- Molecule 2: Geranylgeranyl transferase type-1 subunit beta



- Molecule 3: Fusion protein consisting of transforming protein p21b and Ras related protein Rap-2b





● Molecule 3: Fusion protein consisting of transforming protein p21b and Ras related protein Rap-2b



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	272.34Å 271.57Å 185.43Å 90.00° 131.56° 90.00°	Depositor
Resolution (Å)	29.92 – 2.65 29.92 – 2.65	Depositor EDS
% Data completeness (in resolution range)	98.0 (29.92-2.65) 98.7 (29.92-2.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.02 (at 2.64Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.204 , 0.228 0.205 , 0.228	Depositor DCC
$R_{free}$ test set	14292 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.7	Xtriage
Anisotropy	0.031	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 54.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.080 for -h-2*k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	33443	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.37% 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: GRG, ZN, CL, GER

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	0.34	0/2695	0.52	0/3668
1	C	0.36	0/2709	0.53	0/3684
1	E	0.35	0/2708	0.53	0/3684
1	G	0.35	0/2699	0.53	0/3672
1	I	0.36	0/2722	0.53	0/3700
1	K	0.40	0/2737	0.55	0/3717
2	B	0.36	0/2759	0.60	2/3733 (0.1%)
2	D	0.37	0/2775	0.59	2/3752 (0.1%)
2	F	0.38	0/2780	0.60	2/3758 (0.1%)
2	H	0.35	0/2756	0.58	2/3729 (0.1%)
2	J	0.36	0/2773	0.59	2/3750 (0.1%)
2	L	0.40	0/2785	0.61	2/3764 (0.1%)
3	M	0.67	0/29	0.82	0/37
3	N	0.66	0/29	0.81	0/37
All	All	0.37	0/32956	0.56	12/44685 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1
2	F	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	259	GLY	N-CA-C	-6.01	98.07	113.10
2	F	259	GLY	N-CA-C	-5.96	98.20	113.10
2	D	259	GLY	N-CA-C	-5.84	98.49	113.10
2	J	259	GLY	N-CA-C	-5.83	98.51	113.10
2	L	259	GLY	N-CA-C	-5.77	98.68	113.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	297	TYR	Sidechain
2	F	297	TYR	Sidechain

## 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	2629	0	2520	48	0
1	C	2643	0	2540	44	0
1	E	2642	0	2534	57	0
1	G	2633	0	2524	55	0
1	I	2656	0	2560	44	0
1	K	2671	0	2588	39	0
2	B	2697	0	2600	50	0
2	D	2713	0	2628	52	0
2	F	2718	0	2635	37	0
2	H	2694	0	2590	67	0
2	J	2711	0	2616	51	0
2	L	2723	0	2643	45	0
3	M	30	0	34	3	0
3	N	30	0	34	3	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	F	1	0	0	0	0
4	H	1	0	0	0	0
4	J	1	0	0	0	0
4	L	1	0	0	0	0
5	B	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	F	1	0	0	0	0
5	G	1	0	0	1	0
5	H	1	0	0	0	0
5	J	1	0	0	0	0
5	K	1	0	0	1	0
5	L	1	0	0	0	0
6	B	29	0	33	1	0
6	D	29	0	33	1	0
6	F	29	0	33	2	0
6	H	29	0	33	1	0
6	J	29	0	33	1	0
6	L	29	0	33	1	0
7	M	20	0	33	5	0
7	N	20	0	33	5	0
8	A	71	0	0	1	0
8	B	63	0	0	0	0
8	C	80	0	0	3	0
8	D	94	0	0	1	0
8	E	64	0	0	0	0
8	F	92	0	0	2	0
8	G	58	0	0	2	0
8	H	48	0	0	2	0
8	I	87	0	0	4	0
8	J	73	0	0	2	0
8	K	159	0	0	4	0
8	L	134	0	0	2	0
8	N	1	0	0	0	0
All	All	33443	0	31310	573	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 9.

The worst 5 of 573 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:I:156:ILE:HG12	1:I:172:ARG:HH12	1.10	1.15
1:E:156:ILE:HG12	1:E:172:ARG:HH12	1.09	1.11
1:A:156:ILE:HG12	1:A:172:ARG:HH12	1.03	1.08
1:K:156:ILE:HG12	1:K:172:ARG:HH12	0.99	1.08

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:156:ILE:HG12	1:G:172:ARG:HH12	1.17	1.04

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	312/377 (83%)	287 (92%)	23 (7%)	2 (1%)	27	41
1	C	312/377 (83%)	289 (93%)	22 (7%)	1 (0%)	43	59
1	E	312/377 (83%)	289 (93%)	21 (7%)	2 (1%)	27	41
1	G	312/377 (83%)	291 (93%)	19 (6%)	2 (1%)	27	41
1	I	312/377 (83%)	289 (93%)	21 (7%)	2 (1%)	27	41
1	K	312/377 (83%)	293 (94%)	17 (5%)	2 (1%)	27	41
2	B	344/377 (91%)	329 (96%)	14 (4%)	1 (0%)	43	59
2	D	344/377 (91%)	329 (96%)	13 (4%)	2 (1%)	27	41
2	F	344/377 (91%)	328 (95%)	15 (4%)	1 (0%)	43	59
2	H	344/377 (91%)	323 (94%)	19 (6%)	2 (1%)	27	41
2	J	344/377 (91%)	324 (94%)	19 (6%)	1 (0%)	43	59
2	L	344/377 (91%)	330 (96%)	13 (4%)	1 (0%)	43	59
3	M	2/11 (18%)	2 (100%)	0	0	100	100
3	N	2/11 (18%)	2 (100%)	0	0	100	100
All	All	3940/4546 (87%)	3705 (94%)	216 (6%)	19 (0%)	31	46

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	258	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	306	HIS
1	G	326	GLN
2	H	258	ASN
1	I	306	HIS

### 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	280/338 (83%)	276 (99%)	4 (1%)	69	84
1	C	283/338 (84%)	275 (97%)	8 (3%)	47	66
1	E	284/338 (84%)	276 (97%)	8 (3%)	47	66
1	G	281/338 (83%)	277 (99%)	4 (1%)	69	84
1	I	287/338 (85%)	279 (97%)	8 (3%)	47	66
1	K	291/338 (86%)	286 (98%)	5 (2%)	63	80
2	B	289/326 (89%)	275 (95%)	14 (5%)	28	43
2	D	293/326 (90%)	278 (95%)	15 (5%)	26	41
2	F	294/326 (90%)	281 (96%)	13 (4%)	31	48
2	H	288/326 (88%)	272 (94%)	16 (6%)	23	36
2	J	292/326 (90%)	276 (94%)	16 (6%)	24	37
2	L	296/326 (91%)	281 (95%)	15 (5%)	26	41
3	M	4/11 (36%)	4 (100%)	0	100	100
3	N	4/11 (36%)	4 (100%)	0	100	100
All	All	3466/4006 (86%)	3340 (96%)	126 (4%)	38	56

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	258	ASN
2	H	216	LEU
2	L	216	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	F	306	VAL
1	G	224	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	212	GLN
1	G	162	GLN
1	K	218	ASN
2	F	246	ASN
1	G	80	GLN

### 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 carbohydrates in this entry.

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

Of 23 ligands modelled in this entry, 15 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	$\# Z  > 2$
6	GRG	B	1721	-	27,28,28	0.87	1 (3%)	31,37,37	0.70	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	GRG	D	1722	-	27,28,28	0.86	1 (3%)	31,37,37	0.69	0
6	GRG	F	1723	-	27,28,28	0.87	1 (3%)	31,37,37	0.68	0
6	GRG	H	1724	-	27,28,28	0.81	1 (3%)	31,37,37	0.70	0
6	GRG	J	1725	-	27,28,28	0.84	1 (3%)	31,37,37	0.67	0
6	GRG	L	1726	-	27,28,28	0.89	1 (3%)	31,37,37	0.68	0
7	GER	M	1727	-	19,19,19	0.83	0	22,22,22	0.66	0
7	GER	N	1727	-	19,19,19	0.86	0	22,22,22	0.68	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
6	GRG	B	1721	-	-	0/31/31/31	0/0/0/0
6	GRG	D	1722	-	-	0/31/31/31	0/0/0/0
6	GRG	F	1723	-	-	0/31/31/31	0/0/0/0
6	GRG	H	1724	-	-	0/31/31/31	0/0/0/0
6	GRG	J	1725	-	-	0/31/31/31	0/0/0/0
6	GRG	L	1726	-	-	0/31/31/31	0/0/0/0
7	GER	M	1727	-	-	0/20/20/20	0/0/0/0
7	GER	N	1727	-	-	0/20/20/20	0/0/0/0

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	L	1726	GRG	PB-O3A	-2.93	1.55	1.60
6	F	1723	GRG	PB-O3A	-2.63	1.56	1.60
6	B	1721	GRG	PB-O3A	-2.60	1.56	1.60
6	J	1725	GRG	PB-O3A	-2.38	1.56	1.60
6	H	1724	GRG	PB-O3A	-2.27	1.56	1.60

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	1721	GRG	1	0
6	D	1722	GRG	1	0
6	F	1723	GRG	2	0
6	H	1724	GRG	1	0
6	J	1725	GRG	1	0
6	L	1726	GRG	1	0
7	M	1727	GER	5	0
7	N	1727	GER	5	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

### 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	314/377 (83%)	0.02	15 (4%)	30	27	35, 57, 92, 110	0
1	C	314/377 (83%)	-0.05	10 (3%)	47	44	33, 54, 80, 97	0
1	E	314/377 (83%)	0.16	14 (4%)	33	30	35, 59, 86, 103	0
1	G	314/377 (83%)	0.13	11 (3%)	44	40	35, 59, 87, 102	0
1	I	314/377 (83%)	-0.03	11 (3%)	44	40	29, 52, 84, 95	0
1	K	314/377 (83%)	-0.23	6 (1%)	66	63	23, 41, 67, 81	0
2	B	346/377 (91%)	-0.01	17 (4%)	29	26	35, 52, 75, 100	0
2	D	346/377 (91%)	0.01	15 (4%)	35	32	32, 46, 71, 95	0
2	F	346/377 (91%)	0.00	16 (4%)	32	29	34, 47, 74, 101	0
2	H	346/377 (91%)	0.41	32 (9%)	9	7	36, 65, 95, 112	0
2	J	346/377 (91%)	0.09	16 (4%)	32	29	30, 50, 80, 104	0
2	L	346/377 (91%)	-0.05	9 (2%)	56	51	25, 40, 65, 93	0
3	M	4/11 (36%)	2.72	3 (75%)	0	0	60, 65, 73, 79	4 (100%)
3	N	4/11 (36%)	2.34	2 (50%)	0	0	59, 65, 74, 80	4 (100%)
All	All	3968/4546 (87%)	0.04	177 (4%)	33	30	23, 52, 84, 112	8 (0%)

The worst 5 of 177 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	306	HIS	6.0
1	G	306	HIS	5.4
1	G	305	SER	5.3
2	D	363	THR	5.1
2	H	363	THR	5.1

## 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. 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( $\text{\AA}^2$ )	Q<0.9
7	GER	N	1727	20/20	0.79	0.37	63,70,82,82	20
7	GER	M	1727	20/20	0.82	0.33	67,74,81,82	20
6	GRG	L	1726	29/29	0.95	0.23	37,42,53,58	0
6	GRG	H	1724	29/29	0.95	0.18	52,57,72,73	0
6	GRG	D	1722	29/29	0.95	0.17	42,48,55,57	0
6	GRG	B	1721	29/29	0.96	0.18	48,53,59,60	0
6	GRG	J	1725	29/29	0.96	0.21	38,44,55,58	0
6	GRG	F	1723	29/29	0.96	0.17	47,52,59,59	0
5	CL	G	1711	1/1	0.97	0.19	55,55,55,55	0
5	CL	B	1702	1/1	0.97	0.07	64,64,64,64	0
5	CL	J	1715	1/1	0.97	0.07	58,58,58,58	0
5	CL	C	1705	1/1	0.97	0.25	55,55,55,55	0
5	CL	H	1712	1/1	0.97	0.06	58,58,58,58	0
4	ZN	H	378	1/1	0.98	0.10	54,54,54,54	0
5	CL	D	1706	1/1	0.98	0.07	41,41,41,41	0
5	CL	K	1717	1/1	0.98	0.09	55,55,55,55	0
5	CL	L	1718	1/1	0.99	0.06	43,43,43,43	0
4	ZN	J	378	1/1	0.99	0.10	35,35,35,35	0
5	CL	F	1709	1/1	0.99	0.11	46,46,46,46	0
4	ZN	F	378	1/1	1.00	0.08	40,40,40,40	0
4	ZN	L	378	1/1	1.00	0.09	27,27,27,27	0
4	ZN	D	378	1/1	1.00	0.09	37,37,37,37	0
4	ZN	B	378	1/1	1.00	0.10	38,38,38,38	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.