



# Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 01:10 AM GMT

PDB ID : 3GDV  
Title : Crystal structure of DegS H198P/D320A mutant modified by DFP and in complex with YQF peptide  
Authors : Sohn, J.; Grant, R.A.; Sauer, R.T.  
Deposited on : 2009-02-24  
Resolution : 2.49 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

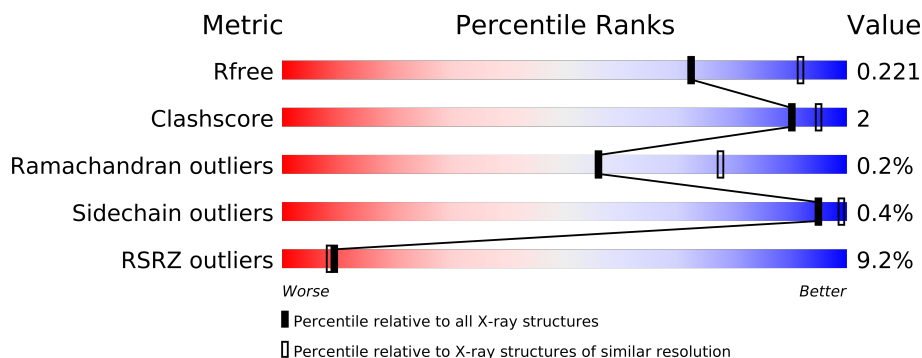
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.49 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))
$R_{free}$	66092	3277 (2.50-2.46)
Clashscore	79885	4136 (2.50-2.46)
Ramachandran outliers	78287	4052 (2.50-2.46)
Sidechain outliers	78261	4054 (2.50-2.46)
RSRZ outliers	66119	3279 (2.50-2.46)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	340	
1	B	340	
1	C	340	
2	D	3	
2	E	3	
2	F	3	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6662 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DegS protease.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	296	Total	C	N	O	P	S	0	0	0
			2193	1371	388	428	1	5			
1	B	284	Total	C	N	O	P	S	0	0	0
			2103	1321	367	410	1	4			
1	C	294	Total	C	N	O	P	S	0	0	0
			2189	1372	388	423	1	5			

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	16	MET	-	EXPRESSION TAG	UNP P0AEE3
A	17	ARG	-	EXPRESSION TAG	UNP P0AEE3
A	18	GLY	-	EXPRESSION TAG	UNP P0AEE3
A	19	SER	-	EXPRESSION TAG	UNP P0AEE3
A	20	HIS	-	EXPRESSION TAG	UNP P0AEE3
A	21	HIS	-	EXPRESSION TAG	UNP P0AEE3
A	22	HIS	-	EXPRESSION TAG	UNP P0AEE3
A	23	HIS	-	EXPRESSION TAG	UNP P0AEE3
A	24	HIS	-	EXPRESSION TAG	UNP P0AEE3
A	25	HIS	-	EXPRESSION TAG	UNP P0AEE3
A	26	GLY	-	EXPRESSION TAG	UNP P0AEE3
A	198	PRO	HIS	ENGINEERED	UNP P0AEE3
A	320	ALA	ASP	ENGINEERED	UNP P0AEE3
B	16	MET	-	EXPRESSION TAG	UNP P0AEE3
B	17	ARG	-	EXPRESSION TAG	UNP P0AEE3
B	18	GLY	-	EXPRESSION TAG	UNP P0AEE3
B	19	SER	-	EXPRESSION TAG	UNP P0AEE3
B	20	HIS	-	EXPRESSION TAG	UNP P0AEE3
B	21	HIS	-	EXPRESSION TAG	UNP P0AEE3
B	22	HIS	-	EXPRESSION TAG	UNP P0AEE3
B	23	HIS	-	EXPRESSION TAG	UNP P0AEE3
B	24	HIS	-	EXPRESSION TAG	UNP P0AEE3
B	25	HIS	-	EXPRESSION TAG	UNP P0AEE3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	26	GLY	-	EXPRESSION TAG	UNP P0AEE3
B	198	PRO	HIS	ENGINEERED	UNP P0AEE3
B	320	ALA	ASP	ENGINEERED	UNP P0AEE3
C	16	MET	-	EXPRESSION TAG	UNP P0AEE3
C	17	ARG	-	EXPRESSION TAG	UNP P0AEE3
C	18	GLY	-	EXPRESSION TAG	UNP P0AEE3
C	19	SER	-	EXPRESSION TAG	UNP P0AEE3
C	20	HIS	-	EXPRESSION TAG	UNP P0AEE3
C	21	HIS	-	EXPRESSION TAG	UNP P0AEE3
C	22	HIS	-	EXPRESSION TAG	UNP P0AEE3
C	23	HIS	-	EXPRESSION TAG	UNP P0AEE3
C	24	HIS	-	EXPRESSION TAG	UNP P0AEE3
C	25	HIS	-	EXPRESSION TAG	UNP P0AEE3
C	26	GLY	-	EXPRESSION TAG	UNP P0AEE3
C	198	PRO	HIS	ENGINEERED	UNP P0AEE3
C	320	ALA	ASP	ENGINEERED	UNP P0AEE3

- Molecule 2 is a protein called YQF peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	3	Total	C	N	O	0	0	0
			33	23	4	6			
2	E	3	Total	C	N	O	0	0	0
			33	23	4	6			
2	F	3	Total	C	N	O	0	0	0
			33	23	4	6			

- Molecule 3 is water.

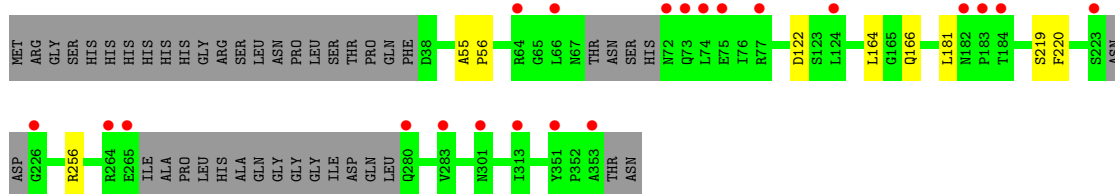
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	12	Total	O	0	0
			12	12		
3	B	34	Total	O	0	0
			34	34		
3	C	32	Total	O	0	0
			32	32		

### 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.

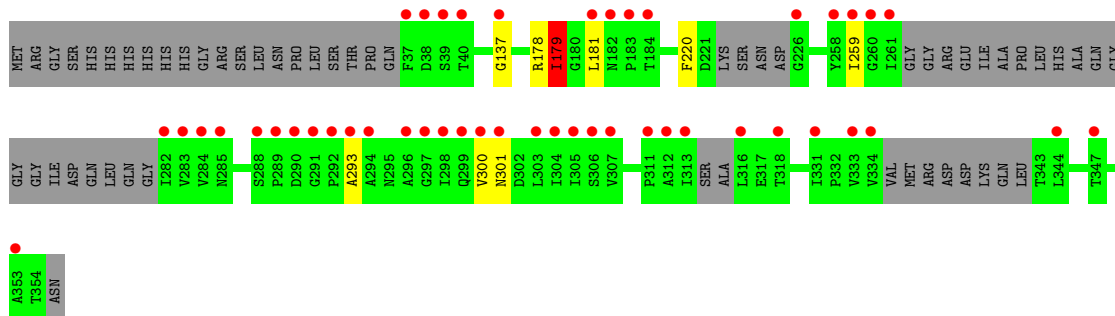
- Molecule 1: DegS protease

Chain A: 



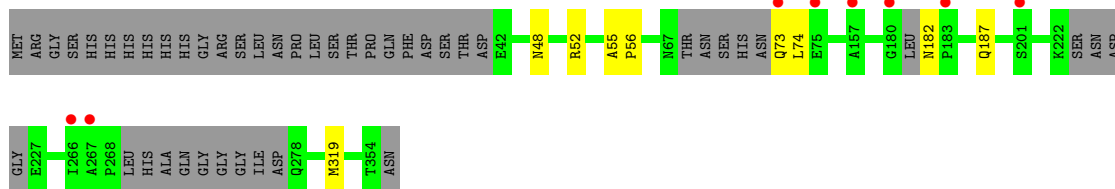
- Molecule 1: DegS protease

Chain B: 



- Molecule 1: DegS protease

Chain C: 



- Molecule 2: YQF peptide

Chain D: 



- Molecule 2: YQF peptide

Chain E: 

  
Y408  
Q409  
F410

- Molecule 2: YQF peptide

Chain F: 

There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.88Å 172.28Å 114.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.04 – 2.49 32.04 – 2.49	Depositor EDS
% Data completeness (in resolution range)	95.9 (32.04-2.49) 95.9 (32.04-2.49)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.43 (at 2.48Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.191 , 0.224 0.187 , 0.221	Depositor DCC
$R_{free}$ test set	2001 reflections (5.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	48.8	Xtriage
Anisotropy	0.454	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 49.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 39894 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6662	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

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

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	0/2204	0.45	0/2996
1	B	0.29	0/2115	0.50	0/2881
1	C	0.27	0/2200	0.49	0/2989
2	D	0.29	0/34	0.27	0/43
2	E	0.32	0/34	0.28	0/43
2	F	0.33	0/34	0.27	0/43
All	All	0.27	0/6621	0.48	0/8995

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2193	0	0	6	0
1	B	2103	0	8	7	0
1	C	2189	0	0	4	0
2	D	33	0	0	0	0
2	E	33	0	0	0	0
2	F	33	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	12	0	0	0	0
3	B	34	0	0	0	0
3	C	32	0	0	0	0
All	All	6662	0	8	14	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (14) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:181:LEU:CD2	1:B:220:PHE:CD1	2.74	0.71
1:A:164:LEU:O	1:B:178:ARG:NH1	2.33	0.62
1:A:181:LEU:CD2	1:A:220:PHE:CD1	2.88	0.56
1:B:178:ARG:O	1:B:179:ILE:CG1	2.56	0.54
1:A:164:LEU:C	1:B:178:ARG:NH1	2.61	0.53
1:A:166:GLN:O	1:B:178:ARG:NH2	2.47	0.48
1:C:48:ASN:OD1	1:C:52:ARG:NH1	2.47	0.47
1:C:182:ASN:ND2	1:C:187:GLN:OE1	2.50	0.44
1:A:122:ASP:OD2	1:A:256:ARG:NH2	2.51	0.44
1:C:55:ALA:N	1:C:56:PRO:CD	2.82	0.42
1:B:259:ILE:O	1:B:293:ALA:N	2.54	0.41
1:B:300:VAL:O	1:B:301:ASN:CB	2.69	0.41
1:A:55:ALA:N	1:A:56:PRO:CD	2.83	0.41
1:C:73:GLN:CG	1:C:74:LEU:N	2.84	0.40

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	287/340 (84%)	279 (97%)	8 (3%)	0	100	100
1	B	273/340 (80%)	262 (96%)	9 (3%)	2 (1%)	30	48
1	C	283/340 (83%)	282 (100%)	1 (0%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	1/3 (33%)	1 (100%)	0	0	100	100
2	E	1/3 (33%)	1 (100%)	0	0	100	100
2	F	1/3 (33%)	1 (100%)	0	0	100	100
All	All	846/1029 (82%)	826 (98%)	18 (2%)	2 (0%)	56	77

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	179	ILE
1	B	137	GLY

### 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	235/275 (86%)	234 (100%)	1 (0%)	95	99
1	B	226/275 (82%)	225 (100%)	1 (0%)	95	99
1	C	234/275 (85%)	233 (100%)	1 (0%)	95	99
2	D	3/3 (100%)	3 (100%)	0	100	100
2	E	3/3 (100%)	3 (100%)	0	100	100
2	F	3/3 (100%)	3 (100%)	0	100	100
All	All	704/834 (84%)	701 (100%)	3 (0%)	95	99

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	219	SER
1	B	179	ILE
1	C	319	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

3 non-standard protein/DNA/RNA residues are modelled in this entry.

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	$\# Z  > 2$
1	MIS	A	201	1	12,12,13	4.95	1 (8%)	14,16,18	0.79	0
1	MIS	B	201	1	12,12,13	4.81	1 (8%)	14,16,18	1.03	1 (7%)
1	MIS	C	201	1	12,12,13	5.01	1 (8%)	14,16,18	0.79	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
1	MIS	A	201	1	-	0/11/13/15	0/0/0/0
1	MIS	B	201	1	-	0/11/13/15	0/0/0/0
1	MIS	C	201	1	-	0/11/13/15	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	201	MIS	O-C	17.21	1.23	1.11
1	A	201	MIS	O-C	17.00	1.23	1.11
1	B	201	MIS	O-C	16.46	1.22	1.11

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	201	MIS	OG-P-O3P	-2.62	97.20	104.68

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	296/340 (87%)	0.30	21 (7%) 16 15	39, 82, 143, 168	0
1	B	284/340 (83%)	0.62	47 (16%) 2 2	28, 50, 176, 198	0
1	C	294/340 (86%)	0.03	8 (2%) 52 54	34, 54, 107, 150	0
2	D	3/3 (100%)	3.06	2 (66%) 0 0	128, 128, 139, 140	0
2	E	3/3 (100%)	4.01	3 (100%) 0 0	145, 145, 162, 164	0
2	F	3/3 (100%)	0.78	0 100 100	66, 66, 81, 99	0
All	All	883/1029 (85%)	0.34	81 (9%) 9 8	28, 60, 151, 198	0

All (81) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	304	ILE	11.9
1	B	313	ILE	10.1
1	B	316	LEU	6.6
1	A	66	LEU	6.6
1	A	74	LEU	6.4
1	B	181	LEU	6.3
1	B	301	ASN	6.2
2	E	408	TYR	6.2
1	B	282	ILE	6.1
1	B	37	PHE	5.5
1	B	291	GLY	5.5
2	D	410	PHE	5.4
1	B	289	PRO	5.4
1	B	182	ASN	5.4
1	B	305	ILE	5.1
1	B	261	ILE	4.8
1	B	303	LEU	4.7
1	B	300	VAL	4.7
1	A	264	ARG	4.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	297	GLY	4.5
1	B	283	VAL	4.5
1	A	73	GLN	4.2
1	A	313	ILE	4.0
1	B	258	TYR	4.0
1	B	312	ALA	3.9
1	B	344	LEU	3.9
1	B	183	PRO	3.9
1	A	72	ASN	3.8
1	C	183	PRO	3.7
1	B	311	PRO	3.7
1	B	294	ALA	3.6
1	B	293	ALA	3.6
1	B	296	ALA	3.6
1	B	298	ILE	3.5
1	A	280	GLN	3.5
1	B	38	ASP	3.5
1	B	299	GLN	3.4
2	D	408	TYR	3.4
1	C	75	GLU	3.4
1	C	267	ALA	3.3
1	B	285	ASN	3.3
1	B	184	THR	3.2
1	A	283	VAL	3.2
1	C	73	GLN	3.2
1	A	301	ASN	3.1
1	B	226	GLY	3.1
1	A	184	THR	3.0
2	E	409	GLN	3.0
1	B	39	SER	2.9
1	C	266	ILE	2.9
1	A	182	ASN	2.9
1	A	64	ARG	2.9
1	B	307	VAL	2.9
2	E	410	PHE	2.9
1	A	75	GLU	2.8
1	B	347	THR	2.8
1	B	259	ILE	2.8
1	A	183	PRO	2.8
1	B	331	ILE	2.7
1	C	180	GLY	2.7
1	B	284	VAL	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	290	ASP	2.7
1	A	226	GLY	2.6
1	B	333	VAL	2.5
1	A	265	GLU	2.4
1	A	223	SER	2.4
1	B	292	PRO	2.4
1	B	334	VAL	2.4
1	A	124	LEU	2.3
1	B	306	SER	2.3
1	C	157	ALA	2.2
1	A	353	ALA	2.2
1	B	353	ALA	2.2
1	A	77	ARG	2.1
1	B	40	THR	2.1
1	B	260	GLY	2.1
1	B	318	THR	2.1
1	B	288	SER	2.1
1	C	201	MIS	2.1
1	B	137	GLY	2.0
1	A	351	TYR	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	MIS	C	201	13/14	0.19	0.39	30,34,46,54	0
1	MIS	B	201	13/14	0.15	0.23	26,35,60,78	0
1	MIS	A	201	13/14	0.12	-0.25	40,51,60,67	0

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.