



# wwPDB X-ray Structure Validation Summary Report

Jun 11, 2014 – 11:41 AM EDT

PDB ID : 3WC9  
Title : The complex structure of HsSQS with ligand, FSPP  
Authors : Shang, N.; Li, Q.; Ko, T.P.; Chan, H.C.; Huang, C.H.; Ren, F.; Zheng, Y.;  
Zhu, Z.; Chen, C.C.; Guo, R.T.  
Deposited on : 2013-05-26  
Resolution : 2.82 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

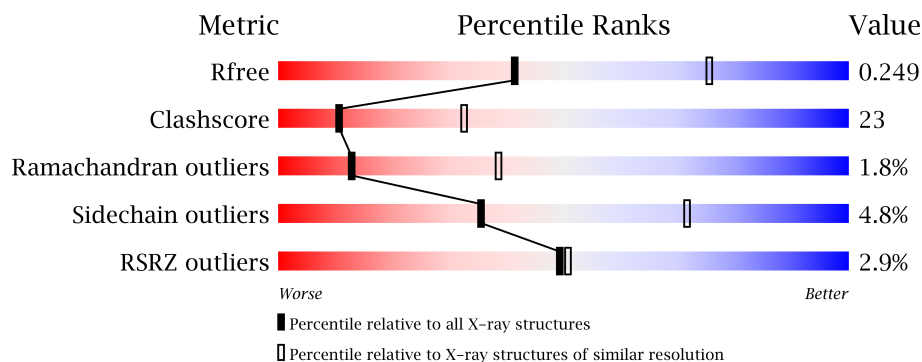
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	<b>FAILED</b>
Xtriage (Phenix)	:	dev-1439
EDS	:	stable23161
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)	:	stable23161

# 1 Overall quality at a glance

The reported resolution of this entry is 2.82 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	1963 (2.84-2.80)
Clashscore	79885	2478 (2.84-2.80)
Ramachandran outliers	78287	2429 (2.84-2.80)
Sidechain outliers	78261	2431 (2.84-2.80)
RSRZ outliers	66119	1966 (2.84-2.80)

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	360	
1	B	360	
1	C	360	
1	D	360	
1	E	360	
1	F	360	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	FPS	A	401	-	X
2	FPS	D	402	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 17168 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 Squalene synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	333	Total	C	N	O	S	0	0	0
			2687	1712	456	501	18			
1	B	334	Total	C	N	O	S	0	0	0
			2696	1717	458	503	18			
1	C	334	Total	C	N	O	S	0	0	0
			2696	1717	458	503	18			
1	D	334	Total	C	N	O	S	0	0	0
			2696	1717	458	503	18			
1	E	334	Total	C	N	O	S	0	0	0
			2696	1717	458	503	18			
1	F	334	Total	C	N	O	S	0	0	0
			2696	1717	458	503	18			

There are 24 discrepancies between the modelled and reference sequences:

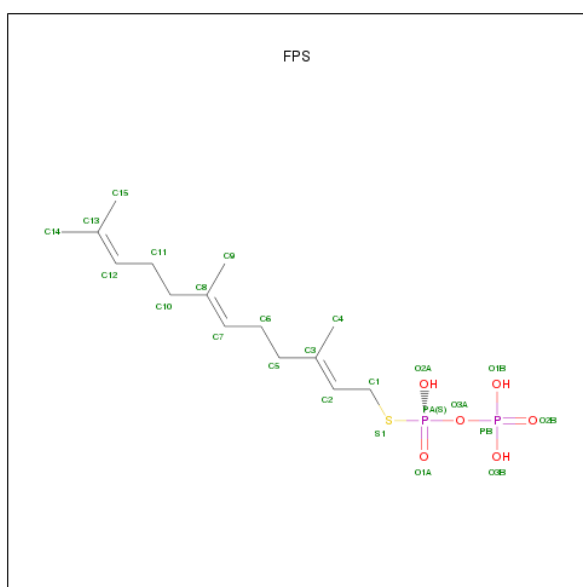
Chain	Residue	Modelled	Actual	Comment	Reference
A	248	LEU	LYS	ENGINEERED MUTATION	UNP P37268
A	315	LEU	LYS	ENGINEERED MUTATION	UNP P37268
A	318	LEU	LYS	ENGINEERED MUTATION	UNP P37268
A	353	ASN	ASP	SEE REMARK 999	UNP P37268
B	248	LEU	LYS	ENGINEERED MUTATION	UNP P37268
B	315	LEU	LYS	ENGINEERED MUTATION	UNP P37268
B	318	LEU	LYS	ENGINEERED MUTATION	UNP P37268
B	353	ASN	ASP	SEE REMARK 999	UNP P37268
C	248	LEU	LYS	ENGINEERED MUTATION	UNP P37268
C	315	LEU	LYS	ENGINEERED MUTATION	UNP P37268
C	318	LEU	LYS	ENGINEERED MUTATION	UNP P37268
C	353	ASN	ASP	SEE REMARK 999	UNP P37268
D	248	LEU	LYS	ENGINEERED MUTATION	UNP P37268
D	315	LEU	LYS	ENGINEERED MUTATION	UNP P37268
D	318	LEU	LYS	ENGINEERED MUTATION	UNP P37268
D	353	ASN	ASP	SEE REMARK 999	UNP P37268
E	248	LEU	LYS	ENGINEERED MUTATION	UNP P37268

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	315	LEU	LYS	ENGINEERED MUTATION	UNP P37268
E	318	LEU	LYS	ENGINEERED MUTATION	UNP P37268
E	353	ASN	ASP	SEE REMARK 999	UNP P37268
F	248	LEU	LYS	ENGINEERED MUTATION	UNP P37268
F	315	LEU	LYS	ENGINEERED MUTATION	UNP P37268
F	318	LEU	LYS	ENGINEERED MUTATION	UNP P37268
F	353	ASN	ASP	SEE REMARK 999	UNP P37268

- Molecule 2 is S-[(2E,6E)-3,7,11-TRIMETHYLDODECA-2,6,10-TRIENYL]TRIHYDROGEN THIODIPHOSPHATE (three-letter code: FPS) (formula: C<sub>15</sub>H<sub>28</sub>O<sub>6</sub>P<sub>2</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	O	P	S	0	0
			24	15	6	2	1		
2	B	1	Total	C	O	P	S	0	0
			24	15	6	2	1		
2	C	1	Total	C	O	P	S	0	0
			24	15	6	2	1		
2	D	1	Total	C	O	P	S	0	0
			24	15	6	2	1		
2	D	1	Total	C	O	P	S	0	0
			24	15	6	2	1		
2	E	1	Total	C	O	P	S	0	0
			24	15	6	2	1		
2	F	1	Total	C	O	P	S	0	0
			24	15	6	2	1		

- Molecule 3 is water.

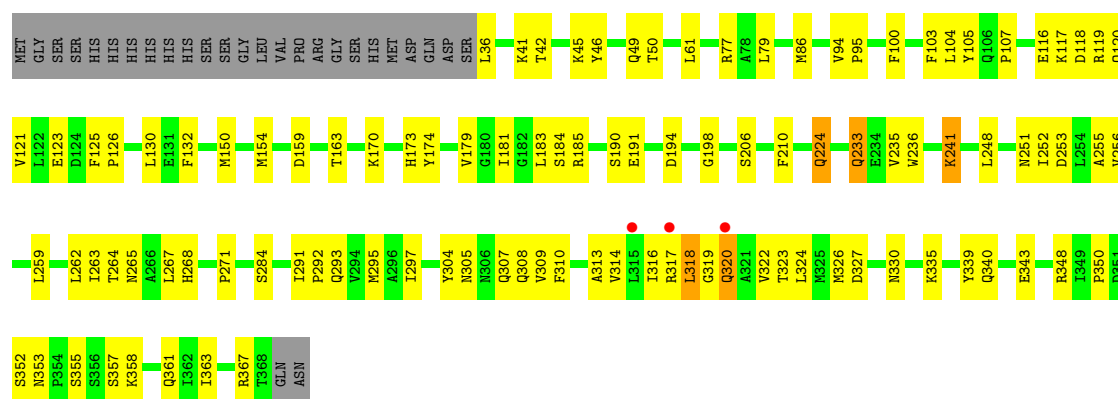
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	179	Total 179	O 179	0	0
3	B	179	Total 179	O 179	0	0
3	C	153	Total 153	O 153	0	0
3	D	114	Total 114	O 114	0	0
3	E	111	Total 111	O 111	0	0
3	F	97	Total 97	O 97	0	0

### 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 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.

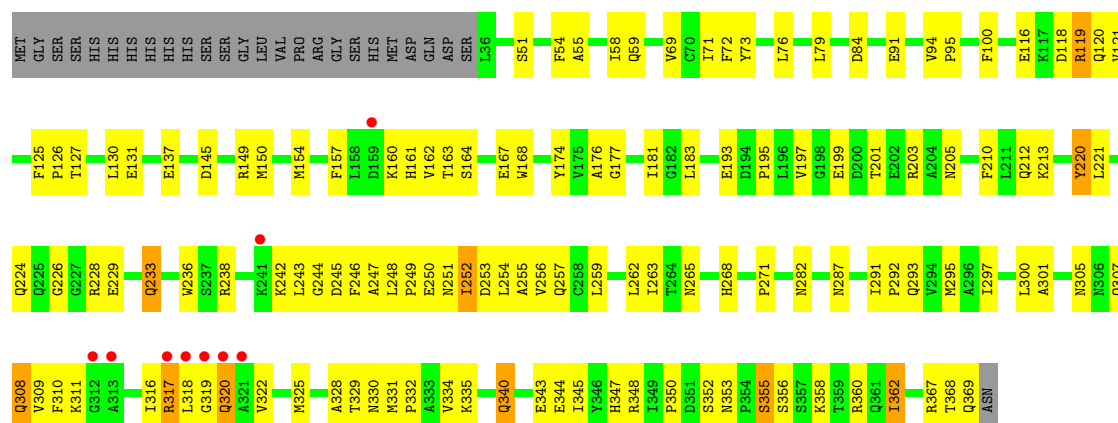
#### • Molecule 1: Squalene synthase

Chain A: 



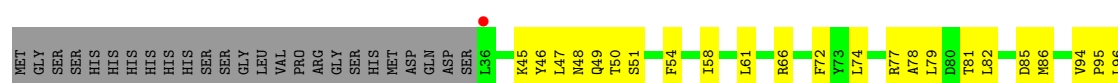
#### • Molecule 1: Squalene synthase

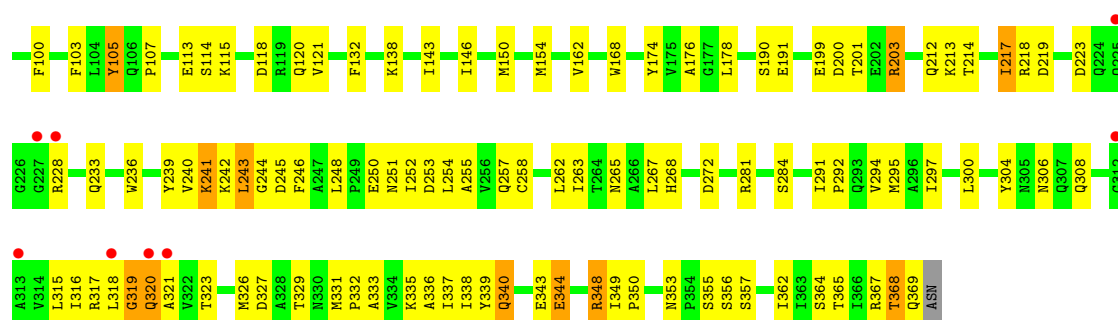
Chain B: 



#### • Molecule 1: Squalene synthase

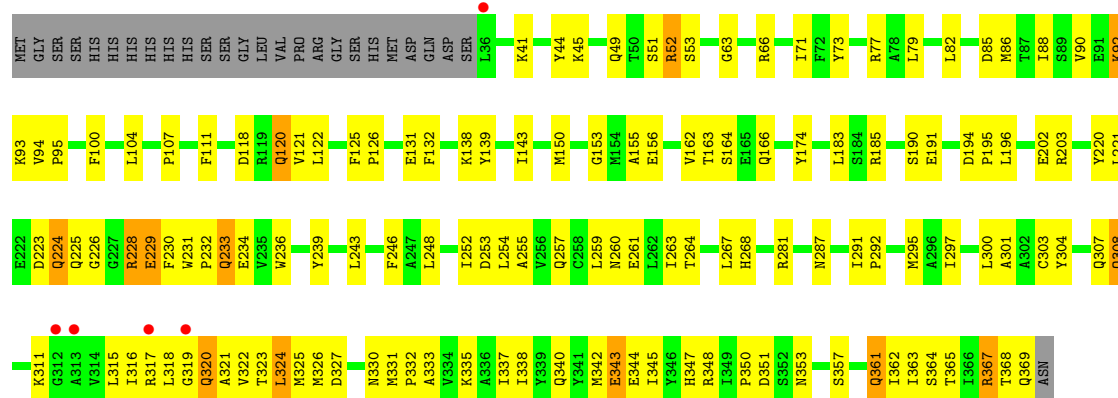
Chain C: 





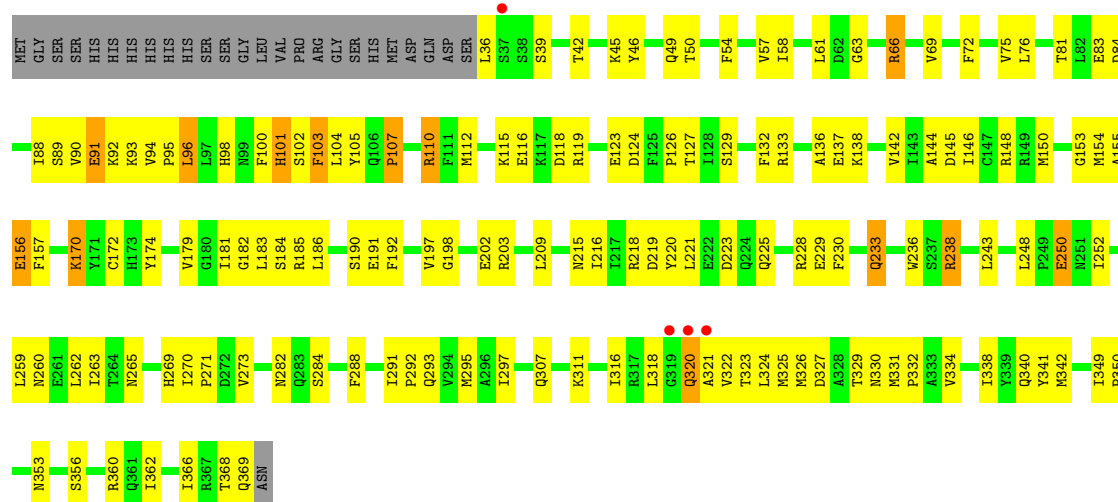
• Molecule 1: Squalene synthase

Chain D: 



• Molecule 1: Squalene synthase

Chain E: 



• Molecule 1: Squalene synthase

Chain F: 







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.19Å 153.33Å 90.99Å 90.00° 91.84° 90.00°	Depositor
Resolution (Å)	25.00 – 2.82 25.08 – 2.82	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.00-2.82) 95.0 (25.08-2.82)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.44 (at 2.84Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.184 , 0.248 0.186 , 0.249	Depositor DCC
$R_{free}$ test set	2688 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.3	Xtriage
Anisotropy	1.015	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 29.2	EDS
Estimated twinning fraction	0.036 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 53245 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	17168	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.14% 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: FPS

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	0.39	0/2742	0.60	0/3712
1	B	0.38	0/2751	0.60	0/3724
1	C	0.38	0/2751	0.59	0/3724
1	D	0.36	0/2751	0.56	0/3724
1	E	0.34	0/2751	0.57	1/3724 (0.0%)
1	F	0.33	0/2751	0.57	0/3724
All	All	0.36	0/16497	0.58	1/22332 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	103	PHE	N-CA-C	-5.34	96.59	111.00

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	2687	0	2668	84	0
1	B	2696	0	2676	120	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2696	0	2676	109	0
1	D	2696	0	2676	131	0
1	E	2696	0	2676	153	0
1	F	2696	0	2676	192	0
2	A	24	0	25	0	0
2	B	24	0	25	1	0
2	C	24	0	25	0	0
2	D	48	0	50	4	0
2	E	24	0	25	0	0
2	F	24	0	25	0	0
3	A	179	0	0	4	0
3	B	179	0	0	2	0
3	C	153	0	0	2	0
3	D	114	0	0	2	0
3	E	111	0	0	1	0
3	F	97	0	0	1	0
All	All	17168	0	16223	749	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 23.

The worst 5 of 749 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:F:291:ILE:HB	1:F:292:PRO:HD3	1.50	0.91
1:A:116:GLU:O	1:A:119:ARG:HG2	1.73	0.89
1:F:93:LYS:HD2	1:F:158:LEU:HD11	1.52	0.89
1:E:356:SER:HB2	1:E:360:ARG:HH12	1.38	0.87
1:B:51:SER:HB2	1:B:73:TYR:CZ	2.10	0.86

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	331/360 (92%)	308 (93%)	22 (7%)	1 (0%)	50	83
1	B	332/360 (92%)	301 (91%)	27 (8%)	4 (1%)	19	53
1	C	332/360 (92%)	297 (90%)	29 (9%)	6 (2%)	13	38
1	D	332/360 (92%)	297 (90%)	31 (9%)	4 (1%)	19	53
1	E	332/360 (92%)	293 (88%)	33 (10%)	6 (2%)	13	38
1	F	332/360 (92%)	278 (84%)	40 (12%)	14 (4%)	4	13
All	All	1991/2160 (92%)	1774 (89%)	182 (9%)	35 (2%)	13	38

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	320	GLN
1	B	320	GLN
1	C	320	GLN
1	D	317	ARG
1	E	320	GLN

### 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	296/320 (92%)	284 (96%)	12 (4%)	41	77
1	B	297/320 (93%)	283 (95%)	14 (5%)	36	72
1	C	297/320 (93%)	286 (96%)	11 (4%)	45	81
1	D	297/320 (93%)	282 (95%)	15 (5%)	33	69
1	E	297/320 (93%)	284 (96%)	13 (4%)	39	75
1	F	297/320 (93%)	277 (93%)	20 (7%)	23	55
All	All	1781/1920 (93%)	1696 (95%)	85 (5%)	35	71

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	86	MET
1	D	324	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	245	ASP
1	D	120	GLN
1	D	228	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 49 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	215	ASN
1	D	340	GLN
1	F	308	GLN
1	D	233	GLN
1	D	361	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

Mogul failed to run properly - this section will therefore be empty.

## 5.5 Carbohydrates ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.6 Ligand geometry ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.7 Other polymers ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 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	333/360 (92%)	-0.52	3 (0%) 81 81	5, 26, 54, 93	0
1	B	334/360 (92%)	-0.38	9 (2%) 52 53	11, 29, 65, 99	0
1	C	334/360 (92%)	-0.36	9 (2%) 52 53	11, 31, 61, 97	0
1	D	334/360 (92%)	-0.24	5 (1%) 70 71	13, 39, 76, 101	0
1	E	334/360 (92%)	-0.05	4 (1%) 75 76	23, 47, 83, 94	0
1	F	334/360 (92%)	0.30	27 (8%) 12 10	23, 61, 101, 108	0
All	All	2003/2160 (92%)	-0.21	57 (2%) 49 52	5, 38, 87, 108	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	321	ALA	6.2
1	F	315	LEU	5.0
1	F	312	GLY	4.9
1	E	321	ALA	4.9
1	F	319	GLY	4.7

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

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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
2	FPS	D	402	24/24	0.47	3.94	66,102,128,128	0
2	FPS	A	401	24/24	0.23	2.32	44,63,97,98	0
2	FPS	E	401	24/24	0.25	1.58	48,66,94,94	0
2	FPS	D	401	24/24	0.23	1.51	38,72,99,100	0
2	FPS	F	401	24/24	0.20	0.56	54,73,86,86	0
2	FPS	B	401	24/24	0.17	0.39	17,44,79,80	0
2	FPS	C	401	24/24	0.17	0.34	25,53,80,80	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.