



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 20, 2019 – 01:48 PM EST

PDB ID : 6E3B  
Title : STRUCTURE OF Siw14 CATALYTIC CORE  
Authors : Florio, T.; Lokareddy, R.; Cingolani, G.  
Deposited on : 2018-07-13  
Resolution : 2.50 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 : rb-20031633  
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) : rb-20031633

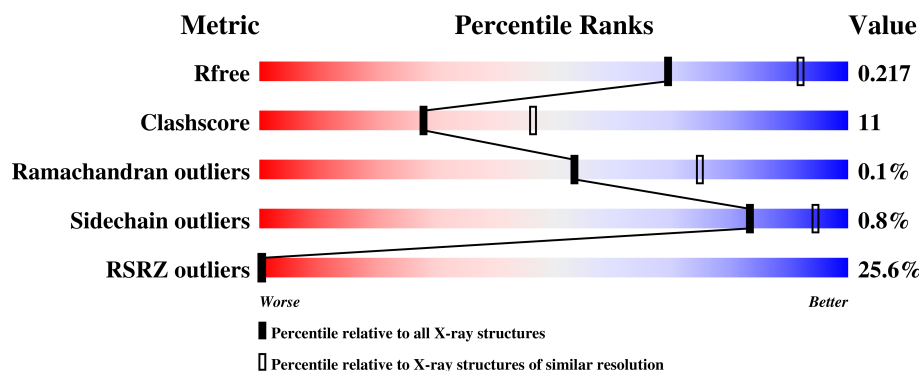
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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.



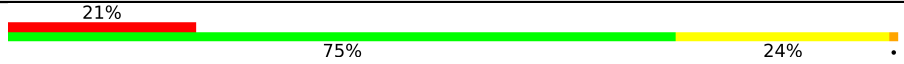
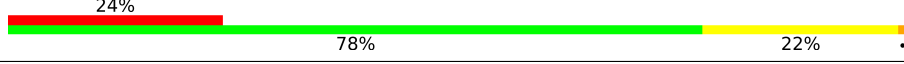




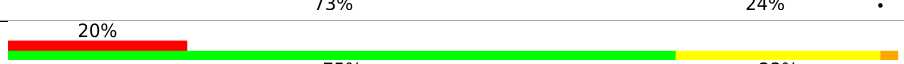
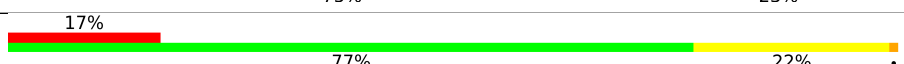
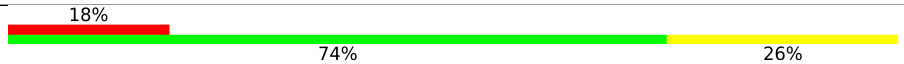
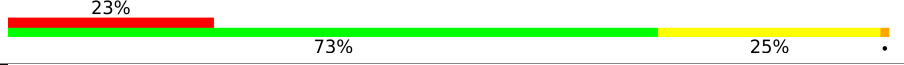
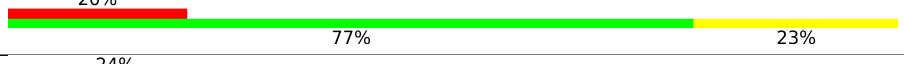



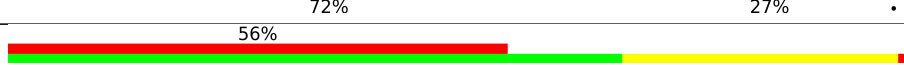
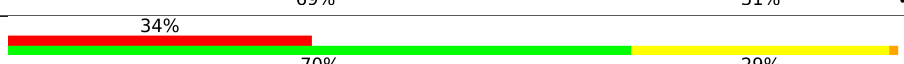
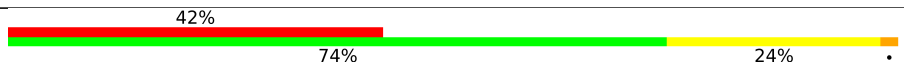

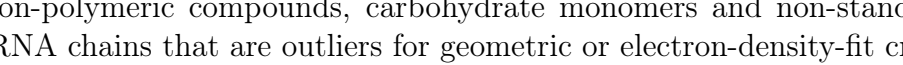
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	111664	4155 (2.50-2.50)
Clashscore	122126	4827 (2.50-2.50)
Ramachandran outliers	120053	4735 (2.50-2.50)
Sidechain outliers	120020	4737 (2.50-2.50)
RSRZ outliers	108989	4058 (2.50-2.50)

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	169	<div> <div>20%</div> <div>73%</div> <div>27%</div> <div>.</div> </div>
1	B	169	<div> <div>19%</div> <div>75%</div> <div>25%</div> <div>.</div> </div>
1	C	169	<div> <div>21%</div> <div>76%</div> <div>22%</div> <div>.</div> </div>
1	D	169	<div> <div>18%</div> <div>87%</div> <div>12%</div> <div>.</div> </div>
1	E	169	<div> <div>18%</div> <div>69%</div> <div>28%</div> <div>.</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	169	
1	G	169	
1	H	169	
1	I	169	
1	J	169	
1	K	169	
1	L	169	
1	M	169	
1	N	169	
1	O	169	
1	P	169	
1	Q	169	
1	R	169	
1	S	169	
1	T	169	
1	V	169	
1	W	169	
1	X	169	
1	Y	169	

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	SO4	P	301	-	-	X	-
2	SO4	W	301	-	-	X	-
2	SO4	Y	301	-	-	X	-

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 34047 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 Tyrosine-protein phosphatase SIW14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	B	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	C	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	D	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	E	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	F	169	Total	C	N	O	S	0	0	0
			1395	903	242	245	5			
1	G	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	H	169	Total	C	N	O	S	0	0	0
			1395	903	242	245	5			
1	I	169	Total	C	N	O	S	0	0	0
			1395	903	242	245	5			
1	J	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	K	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	L	169	Total	C	N	O	S	0	0	0
			1395	903	242	245	5			
1	M	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	N	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	O	169	Total	C	N	O	S	0	0	0
			1395	903	242	245	5			
1	P	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	169	Total	C	N	O	S	0	0	0
			1395	903	242	245	5			
1	R	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	S	169	Total	C	N	O	S	0	0	0
			1377	903	242	227	5			
1	T	168	Total	C	N	O	S	0	0	0
			1393	900	244	244	5			
1	V	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	W	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			
1	X	169	Total	C	N	O	S	0	0	0
			1387	898	241	243	5			
1	Y	169	Total	C	N	O	S	0	0	0
			1401	906	245	245	5			

There are 24 discrepancies between the modelled and reference sequences:

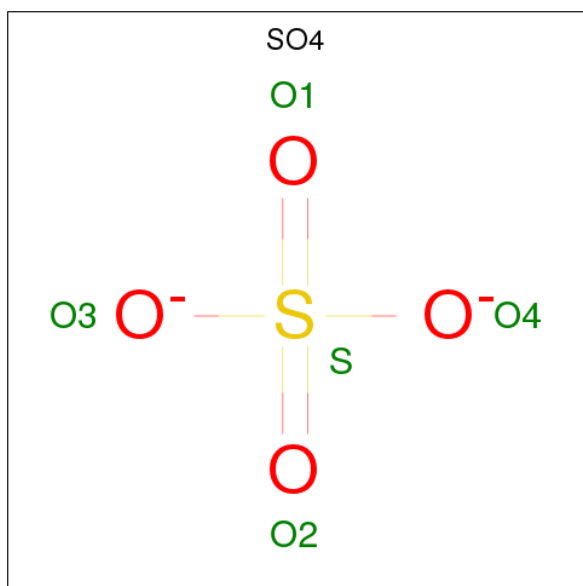
Chain	Residue	Modelled	Actual	Comment	Reference
A	214	SER	CYS	conflict	UNP P53965
B	214	SER	CYS	conflict	UNP P53965
C	214	SER	CYS	conflict	UNP P53965
D	214	SER	CYS	conflict	UNP P53965
E	214	SER	CYS	conflict	UNP P53965
F	214	SER	CYS	conflict	UNP P53965
G	214	SER	CYS	conflict	UNP P53965
H	214	SER	CYS	conflict	UNP P53965
I	214	SER	CYS	conflict	UNP P53965
J	214	SER	CYS	conflict	UNP P53965
K	214	SER	CYS	conflict	UNP P53965
L	214	SER	CYS	conflict	UNP P53965
M	214	SER	CYS	conflict	UNP P53965
N	214	SER	CYS	conflict	UNP P53965
O	214	SER	CYS	conflict	UNP P53965
P	214	SER	CYS	conflict	UNP P53965
Q	214	SER	CYS	conflict	UNP P53965
R	214	SER	CYS	conflict	UNP P53965
S	214	SER	CYS	conflict	UNP P53965
T	214	SER	CYS	conflict	UNP P53965
V	214	SER	CYS	conflict	UNP P53965
W	214	SER	CYS	conflict	UNP P53965
X	214	SER	CYS	conflict	UNP P53965

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
Y	214	SER	CYS	conflict	UNP P53965

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	G	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		
2	I	1	Total	O	S	0	0
			5	4	1		
2	J	1	Total	O	S	0	0
			5	4	1		
2	K	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	L	1	Total	O	S	0	0
			5	4	1		
2	M	1	Total	O	S	0	0
			5	4	1		
2	N	1	Total	O	S	0	0
			5	4	1		
2	O	1	Total	O	S	0	0
			5	4	1		
2	P	1	Total	O	S	0	0
			5	4	1		
2	Q	1	Total	O	S	0	0
			5	4	1		
2	R	1	Total	O	S	0	0
			5	4	1		
2	S	1	Total	O	S	0	0
			5	4	1		
2	T	1	Total	O	S	0	0
			5	4	1		
2	V	1	Total	O	S	0	0
			5	4	1		
2	W	1	Total	O	S	0	0
			5	4	1		
2	X	1	Total	O	S	0	0
			5	4	1		
2	Y	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	14	Total	O	0	0
			14	14		
3	B	17	Total	O	0	0
			17	17		
3	C	17	Total	O	0	0
			17	17		
3	D	17	Total	O	0	0
			17	17		
3	E	13	Total	O	0	0
			13	13		
3	F	23	Total	O	0	0
			23	23		

*Continued on next page...*

*Continued from previous page...*

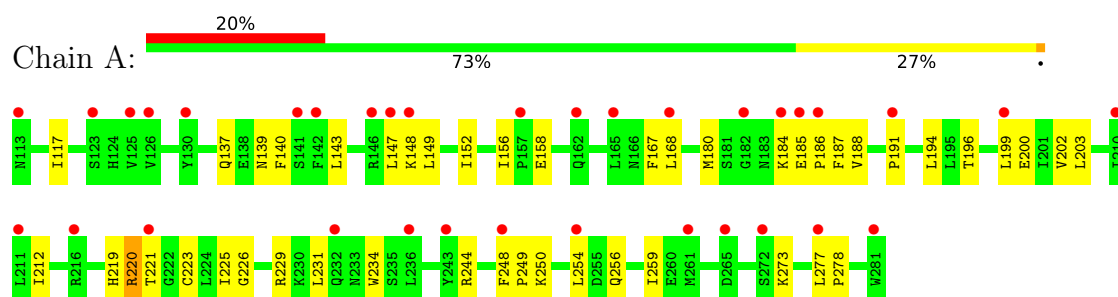
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	15	Total 15	O 15	0	0
3	H	6	Total 6	O 6	0	0
3	I	12	Total 12	O 12	0	0
3	J	18	Total 18	O 18	0	0
3	K	17	Total 17	O 17	0	0
3	L	21	Total 21	O 21	0	0
3	M	18	Total 18	O 18	0	0
3	N	15	Total 15	O 15	0	0
3	O	18	Total 18	O 18	0	0
3	P	16	Total 16	O 16	0	0
3	Q	23	Total 23	O 23	0	0
3	R	13	Total 13	O 13	0	0
3	S	12	Total 12	O 12	0	0
3	T	17	Total 17	O 17	0	0
3	V	9	Total 9	O 9	0	0
3	W	15	Total 15	O 15	0	0
3	X	25	Total 25	O 25	0	0
3	Y	14	Total 14	O 14	0	0



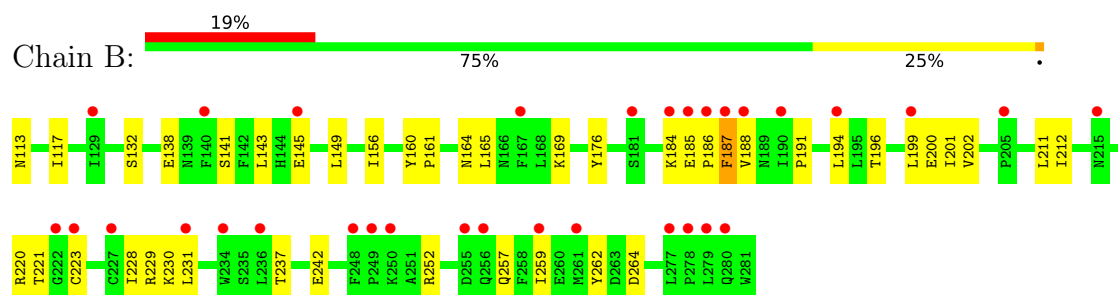
### 3 Residue-property plots

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.

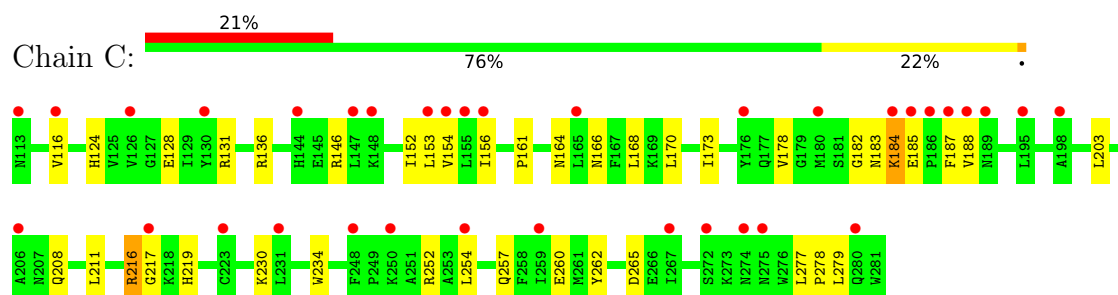
#### • Molecule 1: Tyrosine-protein phosphatase SIW14



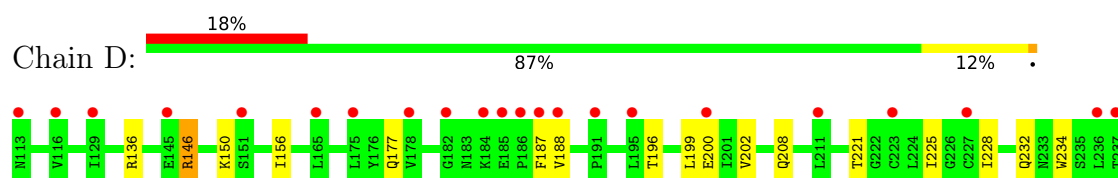
#### • Molecule 1: Tyrosine-protein phosphatase SIW14

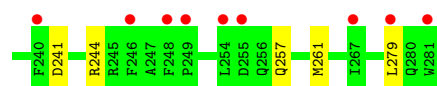


#### • Molecule 1: Tyrosine-protein phosphatase SIW14

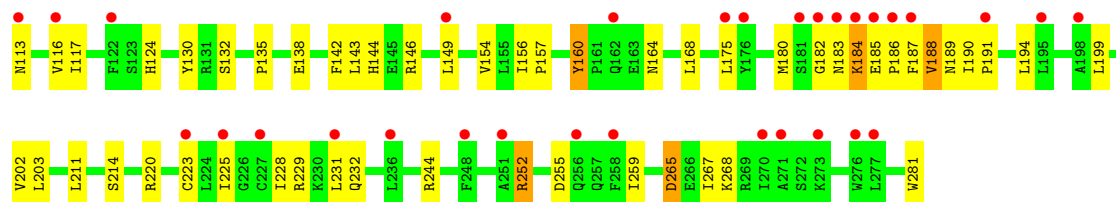


#### • Molecule 1: Tyrosine-protein phosphatase SIW14

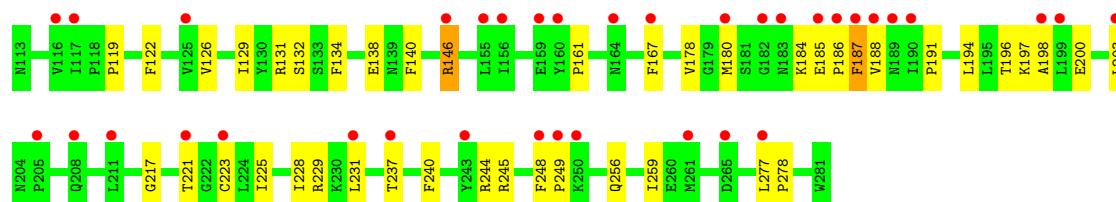
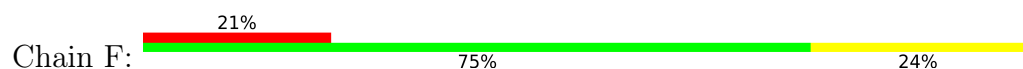




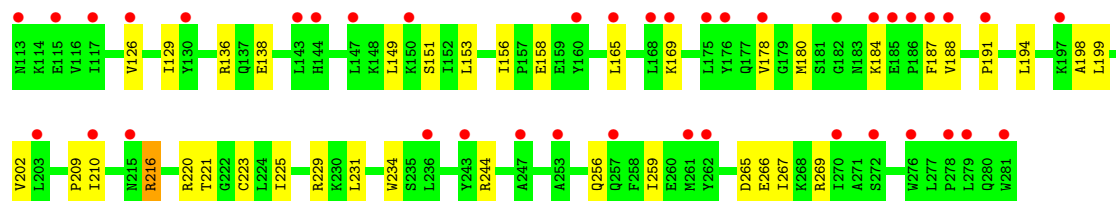
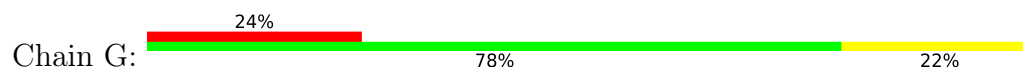
- Molecule 1: Tyrosine-protein phosphatase SIW14



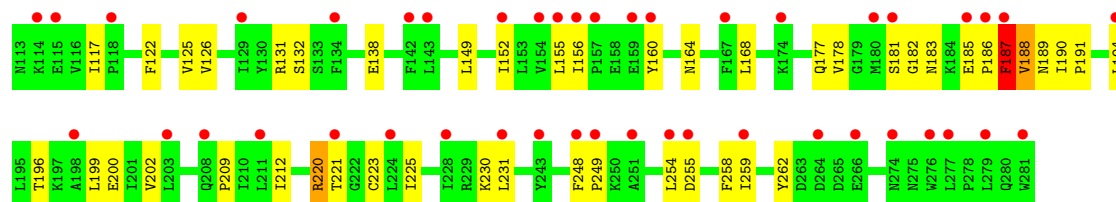
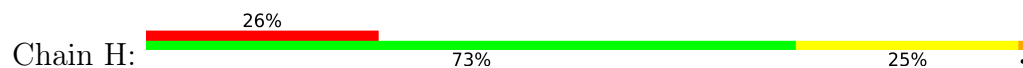
- Molecule 1: Tyrosine-protein phosphatase SIW14



- Molecule 1: Tyrosine-protein phosphatase SIW14

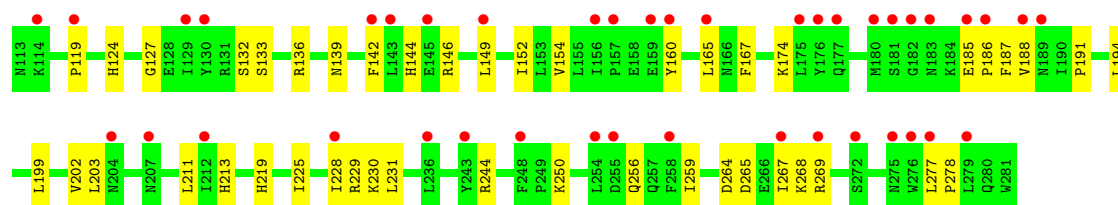


- Molecule 1: Tyrosine-protein phosphatase SIW14

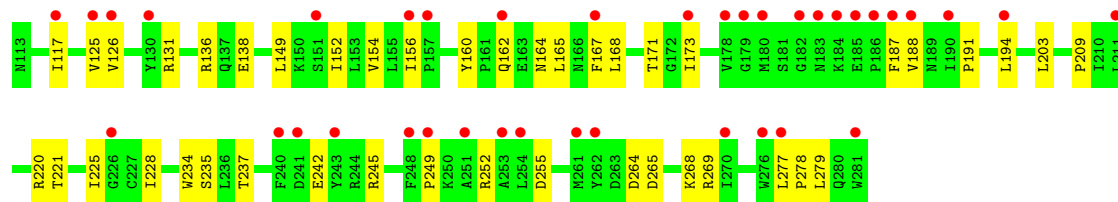
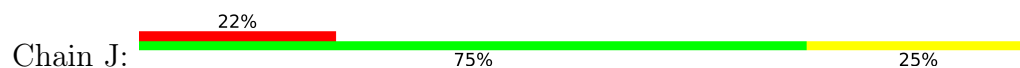


- Molecule 1: Tyrosine-protein phosphatase SIW14

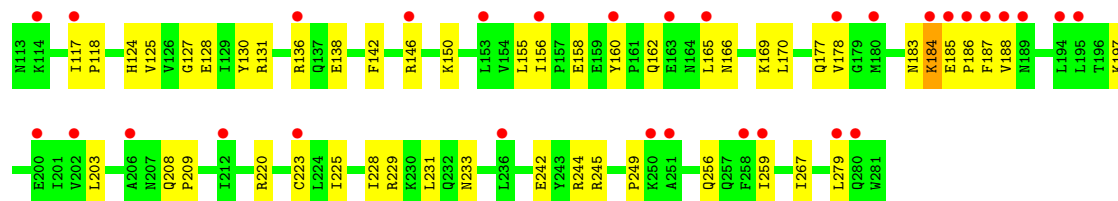




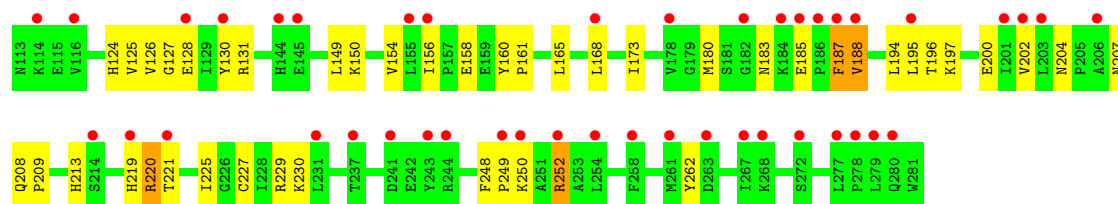
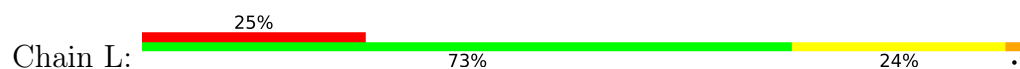
• Molecule 1: Tyrosine-protein phosphatase SIW14



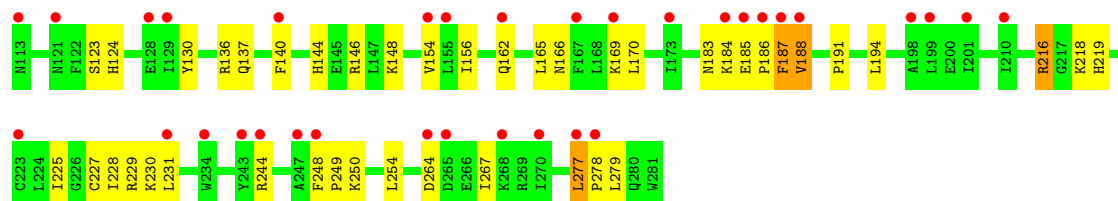
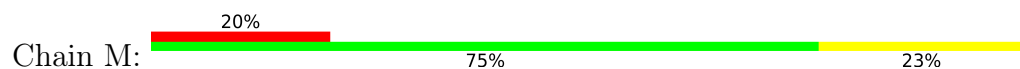
• Molecule 1: Tyrosine-protein phosphatase SIW14



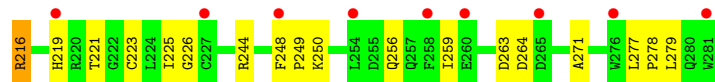
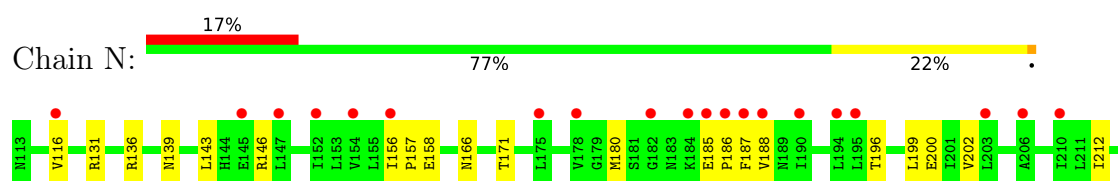
• Molecule 1: Tyrosine-protein phosphatase SIW14



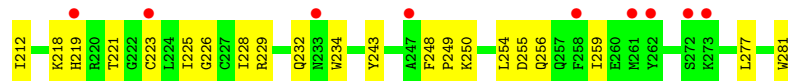
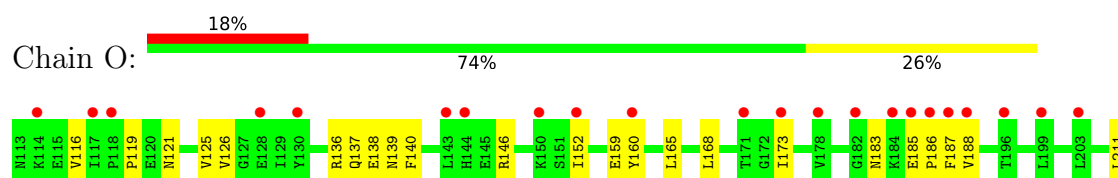
• Molecule 1: Tyrosine-protein phosphatase SIW14



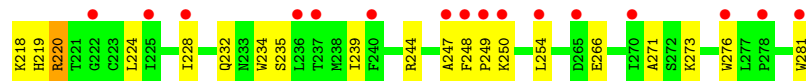
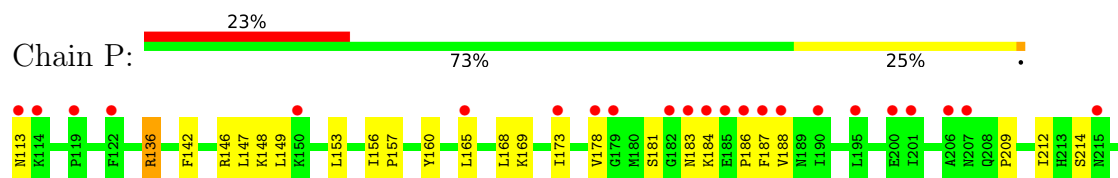
• Molecule 1: Tyrosine-protein phosphatase SIW14



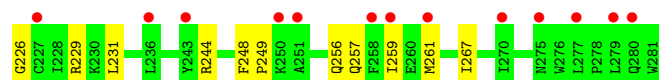
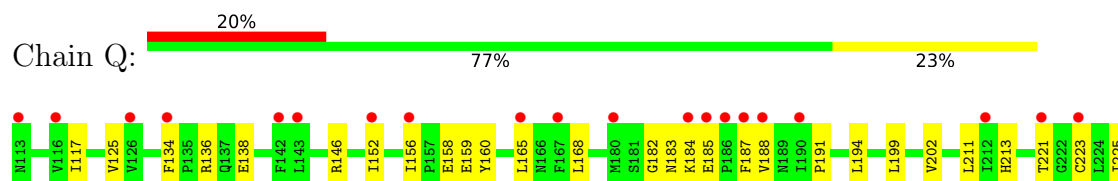
• Molecule 1: Tyrosine-protein phosphatase SIW14



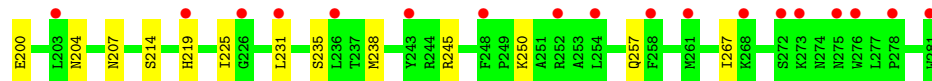
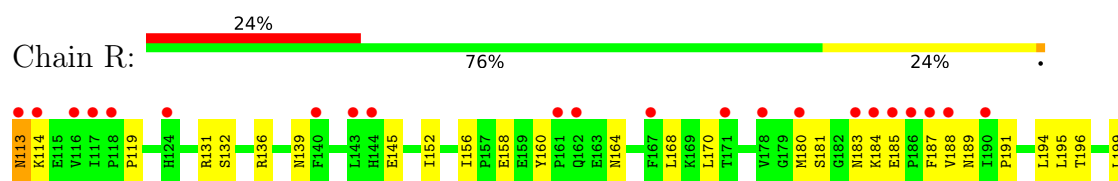
• Molecule 1: Tyrosine-protein phosphatase SIW14



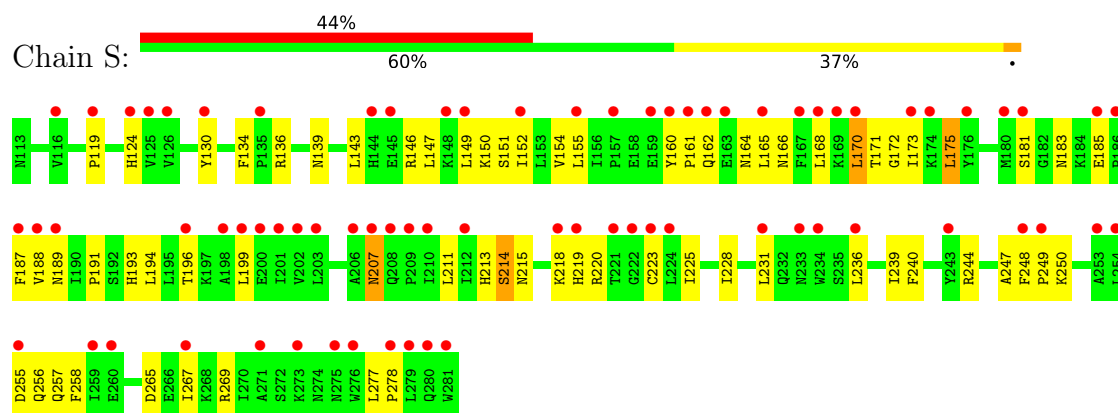
• Molecule 1: Tyrosine-protein phosphatase SIW14



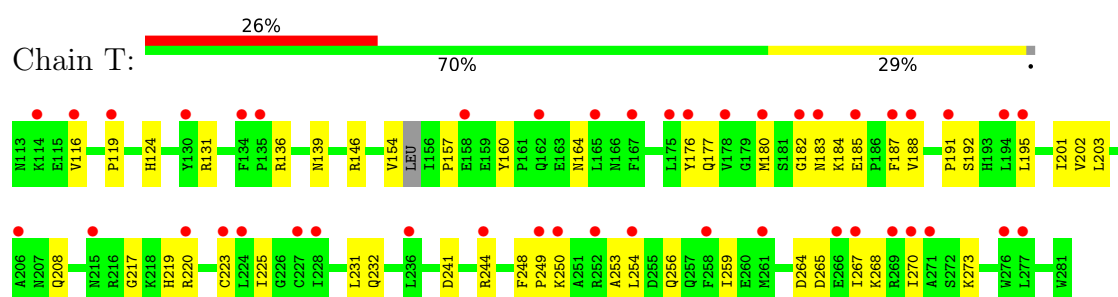
• Molecule 1: Tyrosine-protein phosphatase SIW14



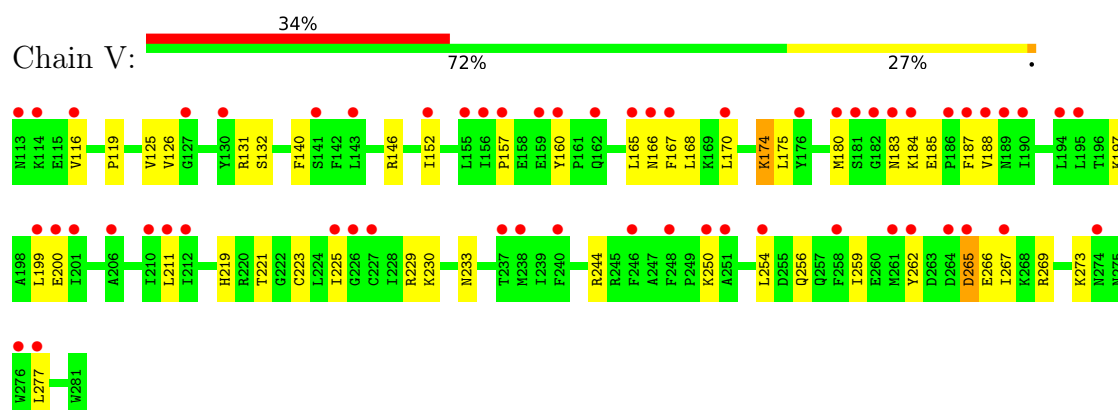
- Molecule 1: Tyrosine-protein phosphatase SIW14



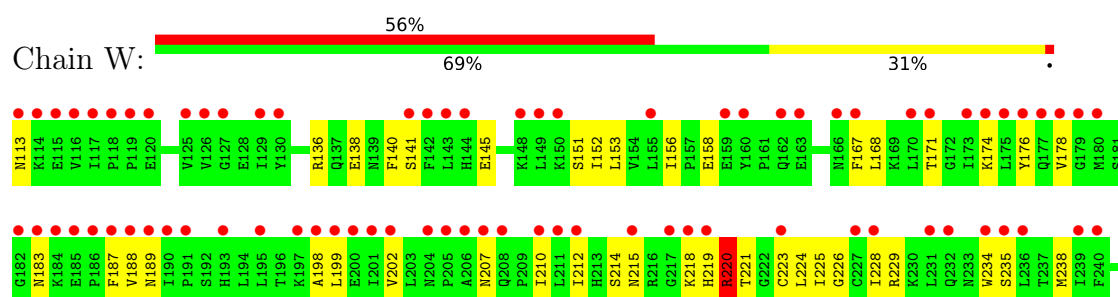
- Molecule 1: Tyrosine-protein phosphatase SIW14

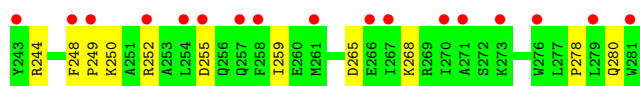


- Molecule 1: Tyrosine-protein phosphatase SIW14

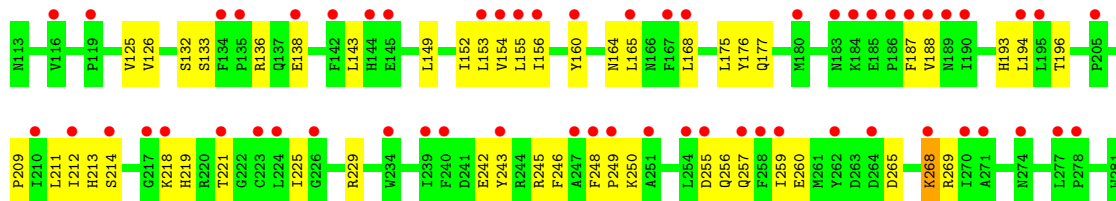


- Molecule 1: Tyrosine-protein phosphatase SIW14

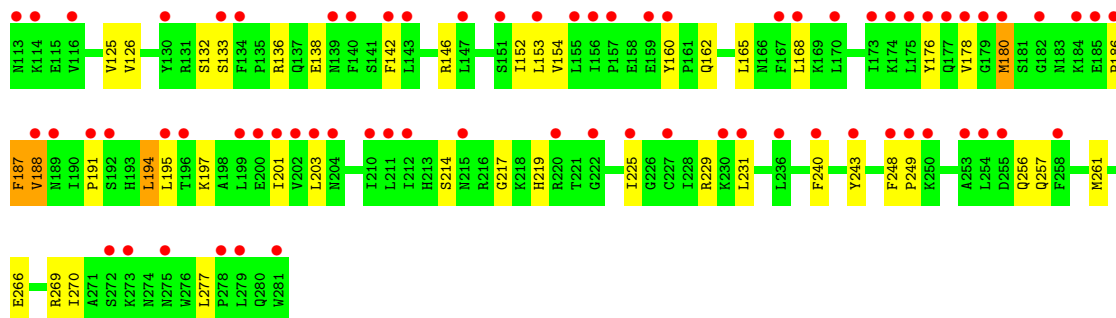
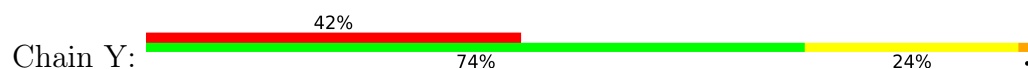




- Molecule 1: Tyrosine-protein phosphatase SIW14



- Molecule 1: Tyrosine-protein phosphatase SIW14



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.47Å 160.49Å 360.57Å 90.00° 90.32° 90.00°	Depositor
Resolution (Å)	14.99 – 2.50 39.79 – 2.49	Depositor EDS
% Data completeness (in resolution range)	90.4 (14.99-2.50) 83.1 (39.79-2.49)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.01	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.62 (at 2.48Å)	Xtriage
Refinement program	PHENIX 1.14 _3260	Depositor
R, $R_{free}$	0.230 , 0.246 0.211 , 0.217	Depositor DCC
$R_{free}$ test set	2019 reflections (1.21%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.4	Xtriage
Anisotropy	0.019	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 24.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.21$	Xtriage
Estimated twinning fraction	0.336 for -1/2*h-1/2*k,-3/2*h+1/2*k,-l 0.357 for -1/2*h+1/2*k,3/2*h+1/2*k,-l 0.290 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.340 for 1/2*h+1/2*k,3/2*h-1/2*k,-l 0.329 for h,-k,-l	Xtriage
Reported twinning fraction	0.450 for h,-k,-l	Depositor
Outliers	1 of 166746 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	34047	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.20% 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: SO4

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	0.27	0/1437	0.49	0/1943
1	B	0.30	0/1437	0.53	1/1943 (0.1%)
1	C	0.28	0/1437	0.49	0/1943
1	D	0.28	0/1437	0.49	0/1943
1	E	0.34	0/1437	0.59	1/1943 (0.1%)
1	F	0.35	0/1431	0.57	3/1936 (0.2%)
1	G	0.30	0/1437	0.51	0/1943
1	H	0.31	0/1431	0.58	2/1936 (0.1%)
1	I	0.31	0/1431	0.52	0/1936
1	J	0.29	0/1437	0.50	0/1943
1	K	0.31	0/1437	0.58	1/1943 (0.1%)
1	L	0.32	0/1431	0.57	2/1936 (0.1%)
1	M	0.29	0/1437	0.56	3/1943 (0.2%)
1	N	0.28	0/1437	0.50	0/1943
1	O	0.28	0/1431	0.49	0/1936
1	P	0.27	0/1437	0.51	0/1943
1	Q	0.29	0/1431	0.53	0/1936
1	R	0.31	0/1437	0.57	1/1943 (0.1%)
1	S	0.30	0/1413	0.52	1/1900 (0.1%)
1	T	0.29	0/1428	0.54	0/1929
1	V	0.33	0/1437	0.62	2/1943 (0.1%)
1	W	0.33	0/1437	0.59	1/1943 (0.1%)
1	X	0.28	0/1423	0.52	0/1927
1	Y	0.31	0/1437	0.59	3/1943 (0.2%)
All	All	0.30	0/34405	0.54	21/46517 (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
1	B	0	1
1	D	0	1
1	E	0	2
1	H	0	2
1	K	0	1
1	M	0	1
1	O	0	1
1	P	0	2
1	W	0	1
1	Y	0	2
All	All	0	14

There are no bond length outliers.

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	V	199	LEU	CA-CB-CG	6.75	130.83	115.30
1	H	188	VAL	N-CA-C	-6.46	93.55	111.00
1	R	113	ASN	N-CA-C	-6.11	94.51	111.00
1	F	146	ARG	CA-CB-CG	-5.83	100.58	113.40
1	L	188	VAL	N-CA-C	-5.71	95.57	111.00

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	186	PRO	Peptide
1	D	146	ARG	Sidechain
1	E	186	PRO	Peptide
1	E	252	ARG	Sidechain
1	H	186	PRO	Peptide

## 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	1401	0	1416	29	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1401	0	1416	35	0
1	C	1401	0	1416	26	0
1	D	1401	0	1416	16	0
1	E	1401	0	1416	44	0
1	F	1395	0	1405	30	0
1	G	1401	0	1416	29	0
1	H	1395	0	1405	31	0
1	I	1395	0	1405	26	0
1	J	1401	0	1416	29	0
1	K	1401	0	1416	34	0
1	L	1395	0	1405	29	0
1	M	1401	0	1416	42	0
1	N	1401	0	1416	33	0
1	O	1395	0	1405	32	0
1	P	1401	0	1416	39	0
1	Q	1395	0	1405	31	0
1	R	1401	0	1416	28	0
1	S	1377	0	1405	43	0
1	T	1393	0	1404	36	0
1	V	1401	0	1416	37	0
1	W	1401	0	1416	47	0
1	X	1387	0	1390	30	0
1	Y	1401	0	1416	27	0
2	A	5	0	0	1	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
2	G	5	0	0	1	0
2	H	5	0	0	1	0
2	I	5	0	0	0	0
2	J	5	0	0	0	0
2	K	5	0	0	0	0
2	L	5	0	0	1	0
2	M	5	0	0	0	0
2	N	5	0	0	0	0
2	O	5	0	0	0	0
2	P	5	0	0	3	0
2	Q	5	0	0	0	0
2	R	5	0	0	0	0
2	S	5	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	T	5	0	0	1	0
2	V	5	0	0	0	0
2	W	5	0	0	5	0
2	X	5	0	0	0	0
2	Y	5	0	0	2	0
3	A	14	0	0	0	0
3	B	17	0	0	0	0
3	C	17	0	0	0	0
3	D	17	0	0	0	0
3	E	13	0	0	1	0
3	F	23	0	0	3	0
3	G	15	0	0	0	0
3	H	6	0	0	0	0
3	I	12	0	0	0	0
3	J	18	0	0	2	0
3	K	17	0	0	1	0
3	L	21	0	0	2	0
3	M	18	0	0	2	0
3	N	15	0	0	2	0
3	O	18	0	0	0	0
3	P	16	0	0	3	0
3	Q	23	0	0	1	0
3	R	13	0	0	1	0
3	S	12	0	0	0	0
3	T	17	0	0	2	0
3	V	9	0	0	1	0
3	W	15	0	0	2	0
3	X	25	0	0	0	0
3	Y	14	0	0	0	0
All	All	34047	0	33869	719	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 11.

The worst 5 of 719 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:V:166:ASN:O	1:V:170:LEU:HD12	1.44	1.16
1:E:244:ARG:HD3	1:M:244:ARG:HE	1.24	1.02
1:M:230:LYS:HE3	1:M:264:ASP:OD2	1.60	1.02
1:M:230:LYS:HE3	1:M:264:ASP:CG	1.85	0.97

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:M:187:PHE:HA	1:M:188:VAL:HG13	1.49	0.93

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	167/169 (99%)	158 (95%)	9 (5%)	0	100	100
1	B	167/169 (99%)	156 (93%)	11 (7%)	0	100	100
1	C	167/169 (99%)	156 (93%)	10 (6%)	1 (1%)	27	46
1	D	167/169 (99%)	157 (94%)	10 (6%)	0	100	100
1	E	167/169 (99%)	155 (93%)	10 (6%)	2 (1%)	14	26
1	F	167/169 (99%)	159 (95%)	8 (5%)	0	100	100
1	G	167/169 (99%)	157 (94%)	10 (6%)	0	100	100
1	H	167/169 (99%)	158 (95%)	9 (5%)	0	100	100
1	I	167/169 (99%)	157 (94%)	10 (6%)	0	100	100
1	J	167/169 (99%)	158 (95%)	9 (5%)	0	100	100
1	K	167/169 (99%)	157 (94%)	10 (6%)	0	100	100
1	L	167/169 (99%)	155 (93%)	12 (7%)	0	100	100
1	M	167/169 (99%)	156 (93%)	11 (7%)	0	100	100
1	N	167/169 (99%)	157 (94%)	10 (6%)	0	100	100
1	O	167/169 (99%)	157 (94%)	10 (6%)	0	100	100
1	P	167/169 (99%)	155 (93%)	12 (7%)	0	100	100
1	Q	167/169 (99%)	157 (94%)	10 (6%)	0	100	100
1	R	167/169 (99%)	158 (95%)	9 (5%)	0	100	100
1	S	167/169 (99%)	157 (94%)	10 (6%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	T	164/169 (97%)	153 (93%)	11 (7%)	0	100	100
1	V	167/169 (99%)	157 (94%)	9 (5%)	1 (1%)	27	46
1	W	167/169 (99%)	156 (93%)	11 (7%)	0	100	100
1	X	167/169 (99%)	156 (93%)	11 (7%)	0	100	100
1	Y	167/169 (99%)	156 (93%)	11 (7%)	0	100	100
All	All	4005/4056 (99%)	3758 (94%)	243 (6%)	4 (0%)	53	75

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	V	184	LYS
1	C	184	LYS
1	E	184	LYS
1	E	188	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	156/156 (100%)	153 (98%)	3 (2%)	60	83
1	B	156/156 (100%)	156 (100%)	0	100	100
1	C	156/156 (100%)	154 (99%)	2 (1%)	71	89
1	D	156/156 (100%)	156 (100%)	0	100	100
1	E	156/156 (100%)	155 (99%)	1 (1%)	87	96
1	F	155/156 (99%)	155 (100%)	0	100	100
1	G	156/156 (100%)	154 (99%)	2 (1%)	71	89
1	H	155/156 (99%)	153 (99%)	2 (1%)	71	89
1	I	155/156 (99%)	154 (99%)	1 (1%)	87	96
1	J	156/156 (100%)	155 (99%)	1 (1%)	87	96
1	K	156/156 (100%)	155 (99%)	1 (1%)	87	96

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	L	155/156 (99%)	153 (99%)	2 (1%)	71	89
1	M	156/156 (100%)	154 (99%)	2 (1%)	71	89
1	N	156/156 (100%)	155 (99%)	1 (1%)	87	96
1	O	155/156 (99%)	155 (100%)	0	100	100
1	P	156/156 (100%)	155 (99%)	1 (1%)	87	96
1	Q	155/156 (99%)	155 (100%)	0	100	100
1	R	156/156 (100%)	155 (99%)	1 (1%)	87	96
1	S	155/156 (99%)	150 (97%)	5 (3%)	42	69
1	T	155/156 (99%)	155 (100%)	0	100	100
1	V	156/156 (100%)	153 (98%)	3 (2%)	60	83
1	W	156/156 (100%)	155 (99%)	1 (1%)	87	96
1	X	153/156 (98%)	152 (99%)	1 (1%)	85	95
1	Y	156/156 (100%)	155 (99%)	1 (1%)	87	96
All	All	3733/3744 (100%)	3702 (99%)	31 (1%)	83	94

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	252	ARG
1	N	216	ARG
1	W	220	ARG
1	M	123	SER
1	P	220	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	J	162	GLN
1	K	256	GLN
1	O	177	GLN
1	I	256	GLN
1	Q	213	HIS

### 5.3.3 RNA ⓘ

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](#)

24 ligands 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	$\# Z  > 2$
2	SO4	A	301	-	4,4,4	0.17	0	6,6,6	0.15	0
2	SO4	B	301	-	4,4,4	0.15	0	6,6,6	0.08	0
2	SO4	C	301	-	4,4,4	0.19	0	6,6,6	0.08	0
2	SO4	D	301	-	4,4,4	0.17	0	6,6,6	0.08	0
2	SO4	E	301	-	4,4,4	0.16	0	6,6,6	0.16	0
2	SO4	F	301	-	4,4,4	0.16	0	6,6,6	0.12	0
2	SO4	G	301	-	4,4,4	0.15	0	6,6,6	0.13	0
2	SO4	H	301	-	4,4,4	0.16	0	6,6,6	0.12	0
2	SO4	I	301	-	4,4,4	0.15	0	6,6,6	0.11	0
2	SO4	J	301	-	4,4,4	0.15	0	6,6,6	0.10	0
2	SO4	K	301	-	4,4,4	0.14	0	6,6,6	0.09	0
2	SO4	L	301	-	4,4,4	0.15	0	6,6,6	0.13	0
2	SO4	M	301	-	4,4,4	0.16	0	6,6,6	0.18	0
2	SO4	N	301	-	4,4,4	0.14	0	6,6,6	0.08	0
2	SO4	O	301	-	4,4,4	0.16	0	6,6,6	0.15	0
2	SO4	P	301	-	4,4,4	0.16	0	6,6,6	0.19	0
2	SO4	Q	301	-	4,4,4	0.16	0	6,6,6	0.11	0
2	SO4	R	301	-	4,4,4	0.17	0	6,6,6	0.11	0
2	SO4	S	301	-	4,4,4	0.17	0	6,6,6	0.16	0
2	SO4	T	301	-	4,4,4	0.14	0	6,6,6	0.11	0
2	SO4	V	301	-	4,4,4	0.17	0	6,6,6	0.12	0
2	SO4	W	301	-	4,4,4	0.11	0	6,6,6	0.39	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SO4	X	301	-	4,4,4	0.17	0	6,6,6	0.08	0
2	SO4	Y	301	-	4,4,4	0.17	0	6,6,6	0.13	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	SO4	A	301	-	-	0/0/0/0	0/0/0/0
2	SO4	B	301	-	-	0/0/0/0	0/0/0/0
2	SO4	C	301	-	-	0/0/0/0	0/0/0/0
2	SO4	D	301	-	-	0/0/0/0	0/0/0/0
2	SO4	E	301	-	-	0/0/0/0	0/0/0/0
2	SO4	F	301	-	-	0/0/0/0	0/0/0/0
2	SO4	G	301	-	-	0/0/0/0	0/0/0/0
2	SO4	H	301	-	-	0/0/0/0	0/0/0/0
2	SO4	I	301	-	-	0/0/0/0	0/0/0/0
2	SO4	J	301	-	-	0/0/0/0	0/0/0/0
2	SO4	K	301	-	-	0/0/0/0	0/0/0/0
2	SO4	L	301	-	-	0/0/0/0	0/0/0/0
2	SO4	M	301	-	-	0/0/0/0	0/0/0/0
2	SO4	N	301	-	-	0/0/0/0	0/0/0/0
2	SO4	O	301	-	-	0/0/0/0	0/0/0/0
2	SO4	P	301	-	-	0/0/0/0	0/0/0/0
2	SO4	Q	301	-	-	0/0/0/0	0/0/0/0
2	SO4	R	301	-	-	0/0/0/0	0/0/0/0
2	SO4	S	301	-	-	0/0/0/0	0/0/0/0
2	SO4	T	301	-	-	0/0/0/0	0/0/0/0
2	SO4	V	301	-	-	0/0/0/0	0/0/0/0
2	SO4	W	301	-	-	0/0/0/0	0/0/0/0
2	SO4	X	301	-	-	0/0/0/0	0/0/0/0
2	SO4	Y	301	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

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 15 short contacts:



Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	SO4	1	0
2	G	301	SO4	1	0
2	H	301	SO4	1	0
2	L	301	SO4	1	0
2	P	301	SO4	3	0
2	T	301	SO4	1	0
2	W	301	SO4	5	0
2	Y	301	SO4	2	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	169/169 (100%)	1.28	34 (20%) 1 1	34, 52, 69, 84	0
1	B	169/169 (100%)	1.28	32 (18%) 1 1	32, 47, 70, 92	0
1	C	169/169 (100%)	1.31	35 (20%) 1 0	27, 46, 77, 107	0
1	D	169/169 (100%)	1.20	31 (18%) 1 1	29, 48, 70, 100	0
1	E	169/169 (100%)	1.26	31 (18%) 1 1	34, 45, 75, 103	0
1	F	169/169 (100%)	1.44	36 (21%) 1 0	36, 48, 70, 99	0
1	G	169/169 (100%)	1.47	40 (23%) 0 0	35, 53, 75, 97	0
1	H	169/169 (100%)	1.47	44 (26%) 0 0	44, 55, 76, 99	0
1	I	169/169 (100%)	1.43	41 (24%) 0 0	42, 54, 80, 94	0
1	J	169/169 (100%)	1.42	38 (22%) 0 0	35, 50, 81, 108	0
1	K	169/169 (100%)	1.49	31 (18%) 1 1	40, 53, 79, 110	0
1	L	169/169 (100%)	1.41	43 (25%) 0 0	39, 53, 67, 81	0
1	M	169/169 (100%)	1.26	33 (19%) 1 1	38, 50, 74, 87	0
1	N	169/169 (100%)	1.23	29 (17%) 1 1	35, 48, 66, 102	0
1	O	169/169 (100%)	1.26	31 (18%) 1 1	35, 48, 70, 83	0
1	P	169/169 (100%)	1.45	39 (23%) 0 0	37, 52, 72, 107	0
1	Q	169/169 (100%)	1.41	33 (19%) 1 1	33, 47, 78, 124	0
1	R	169/169 (100%)	1.35	40 (23%) 0 0	32, 47, 72, 110	0
1	S	169/169 (100%)	1.90	74 (43%) 0 0	42, 69, 92, 100	0
1	T	168/169 (99%)	1.61	44 (26%) 0 0	34, 57, 86, 102	0
1	V	169/169 (100%)	1.84	58 (34%) 0 0	33, 57, 95, 126	0
1	W	169/169 (100%)	2.68	94 (55%) 0 0	50, 72, 112, 131	0
1	X	169/169 (100%)	1.92	58 (34%) 0 0	36, 63, 115, 154	0
1	Y	169/169 (100%)	1.97	71 (42%) 0 0	33, 63, 84, 92	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
All	All	4055/4056 (99%)	1.51	1040 (25%) 0 0	27, 52, 84, 154	0

The worst 5 of 1040 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	V	182	GLY	24.0
1	J	186	PRO	21.1
1	W	187	PHE	15.6
1	X	185	GLU	13.2
1	F	186	PRO	11.0

## 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. 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
2	SO4	L	301	5/5	0.82	0.21	55,55,56,56	0
2	SO4	X	301	5/5	0.83	0.21	58,58,58,59	0
2	SO4	W	301	5/5	0.86	0.18	78,80,81,83	0
2	SO4	H	301	5/5	0.90	0.19	55,55,55,56	0
2	SO4	J	301	5/5	0.90	0.20	42,42,43,43	0
2	SO4	Q	301	5/5	0.91	0.22	51,51,51,52	0
2	SO4	F	301	5/5	0.91	0.17	42,42,42,43	0
2	SO4	N	301	5/5	0.92	0.22	47,47,47,48	0
2	SO4	B	301	5/5	0.92	0.21	49,50,50,50	0
2	SO4	E	301	5/5	0.92	0.19	32,33,33,33	0
2	SO4	G	301	5/5	0.93	0.16	62,62,62,62	0
2	SO4	O	301	5/5	0.93	0.15	46,46,46,46	0
2	SO4	Y	301	5/5	0.93	0.15	53,55,56,56	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	D	301	5/5	0.94	0.21	39,39,40,40	0
2	SO4	S	301	5/5	0.94	0.16	47,47,47,48	0
2	SO4	V	301	5/5	0.94	0.18	51,51,51,51	0
2	SO4	I	301	5/5	0.94	0.15	42,42,43,43	0
2	SO4	M	301	5/5	0.94	0.15	41,42,42,43	0
2	SO4	K	301	5/5	0.95	0.16	40,41,41,42	0
2	SO4	R	301	5/5	0.95	0.15	26,28,28,29	0
2	SO4	C	301	5/5	0.96	0.15	39,39,39,39	0
2	SO4	P	301	5/5	0.96	0.16	45,45,45,45	0
2	SO4	A	301	5/5	0.96	0.19	30,30,31,31	0
2	SO4	T	301	5/5	0.97	0.09	50,50,50,50	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.