



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 6, 2024 – 08:54 pm GMT

PDB ID : 5MG5  
Title : A multi-component acyltransferase PhlABC from *Pseudomonas protegens* soaked with the monoacetylphloroglucinol (MAPG)  
Authors : Pavkov-Keller, T.; Schmidt, N.G.; Kroutil, W.; Gruber, K.  
Deposited on : 2016-11-20  
Resolution : 3.44 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

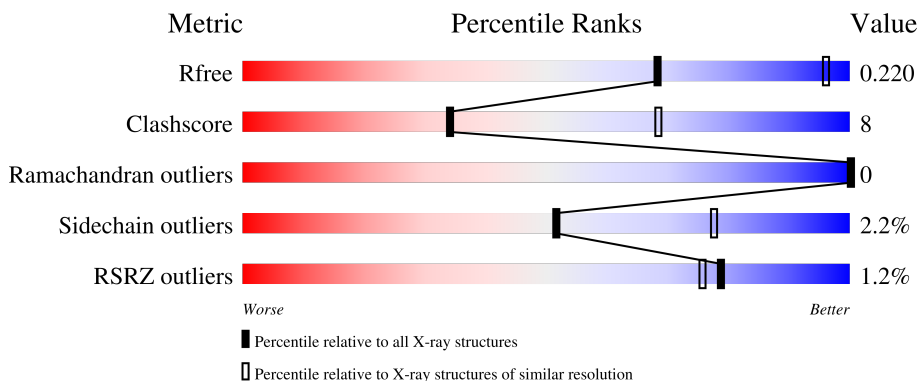
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.44 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))
$R_{free}$	130704	1278 (3.50-3.38)
Clashscore	141614	1361 (3.50-3.38)
Ramachandran outliers	138981	1327 (3.50-3.38)
Sidechain outliers	138945	1328 (3.50-3.38)
RSRZ outliers	127900	1192 (3.50-3.38)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	362	81% 17% ..
1	D	362	81% 17% ..
1	G	362	80% 19% ..
1	J	362	83% 15% ..
1	M	362	78% 20% ..

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	P	362	 3% 78% 20% ..
1	S	362	 78% 20% ..
1	V	362	 81% 17% ..
2	B	146	 3% 84% 16% .
2	E	146	 % 81% 18% ..
2	H	146	 % 77% 22% ..
2	K	146	 % 78% 21% ..
2	N	146	 4% 80% 18% ..
2	Q	146	 2% 86% 14% .
2	T	146	 5% 74% 24% ..
2	W	146	 4% 73% 25% ..
3	C	398	 % 83% 16% .
3	F	398	 % 76% 23% ..
3	I	398	 % 73% 27% .
3	L	398	 76% 23% ..
3	O	398	 2% 77% 22% ..
3	R	398	 2% 74% 25% ..
3	U	398	 2% 78% 22% .
3	X	398	 2% 81% 17% ..

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 54522 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 Hydroxymethylglutaryl-CoA synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	358	Total 2677	C 1687	N 452	O 529	S 9	0	0	0
1	D	359	Total 2686	C 1693	N 454	O 530	S 9	0	0	0
1	G	359	Total 2686	C 1693	N 454	O 530	S 9	0	0	0
1	J	359	Total 2686	C 1693	N 454	O 530	S 9	0	0	0
1	M	359	Total 2686	C 1693	N 454	O 530	S 9	0	0	0
1	P	358	Total 2677	C 1687	N 452	O 529	S 9	0	0	0
1	S	358	Total 2687	C 1693	N 455	O 530	S 9	0	1	0
1	V	359	Total 2686	C 1693	N 454	O 530	S 9	0	0	0

- Molecule 2 is a protein called 2,4-diacetylphloroglucinol biosynthesis protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	145	Total 1150	C 721	N 208	O 207	S 14	0	0	0
2	E	145	Total 1150	C 721	N 208	O 207	S 14	0	0	0
2	H	145	Total 1150	C 721	N 208	O 207	S 14	0	0	0
2	K	145	Total 1150	C 721	N 208	O 207	S 14	0	0	0
2	N	145	Total 1150	C 721	N 208	O 207	S 14	0	0	0
2	Q	145	Total 1150	C 721	N 208	O 207	S 14	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	145	Total	C	N	O	S	0	0	0
			1150	721	208	207	14			
2	W	145	Total	C	N	O	S	0	0	0
			1150	721	208	207	14			

- Molecule 3 is a protein called 2,4-diacetylphloroglucinol biosynthesis protein PhlC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	395	Total	C	N	O	S	0	0	0
			2965	1862	512	571	20			
3	F	395	Total	C	N	O	S	0	0	0
			2965	1862	512	571	20			
3	I	396	Total	C	N	O	S	0	0	0
			2970	1865	513	572	20			
3	L	396	Total	C	N	O	S	0	0	0
			2970	1865	513	572	20			
3	O	396	Total	C	N	O	S	0	0	0
			2970	1865	513	572	20			
3	R	395	Total	C	N	O	S	0	0	0
			2965	1862	512	571	20			
3	U	396	Total	C	N	O	S	0	0	0
			2970	1865	513	572	20			
3	X	396	Total	C	N	O	S	0	0	0
			2970	1865	513	572	20			

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

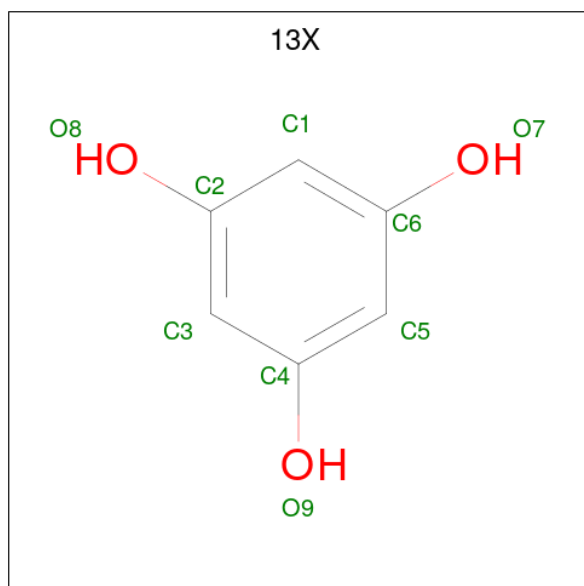
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Zn	0	0
			1	1		
4	E	1	Total	Zn	0	0
			1	1		
4	H	1	Total	Zn	0	0
			1	1		
4	K	1	Total	Zn	0	0
			1	1		
4	N	1	Total	Zn	0	0
			1	1		
4	Q	1	Total	Zn	0	0
			1	1		
4	T	1	Total	Zn	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	W	1	Total Zn 1 1	0	0

- Molecule 5 is benzene-1,3,5-triol (three-letter code: 13X) (formula: C<sub>6</sub>H<sub>6</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total C O 9 6 3	0	0
5	L	1	Total C O 9 6 3	0	0
5	O	1	Total C O 9 6 3	0	0
5	R	1	Total C O 9 6 3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	5	Total O 5 5	0	0
6	B	3	Total O 3 3	0	0
6	C	5	Total O 5 5	0	0
6	D	3	Total O 3 3	0	0

Continued on next page...

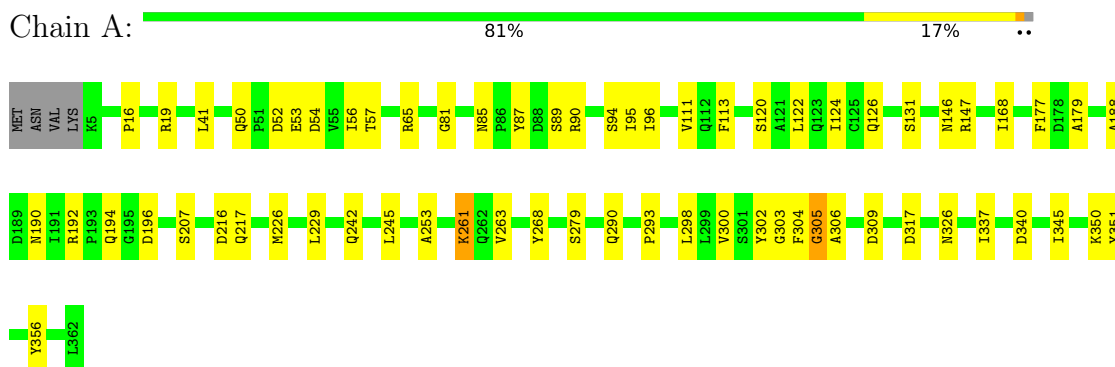
*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	E	2	Total O 2 2	0	0
6	F	2	Total O 2 2	0	0
6	G	2	Total O 2 2	0	0
6	H	3	Total O 3 3	0	0
6	I	2	Total O 2 2	0	0
6	J	5	Total O 5 5	0	0
6	K	1	Total O 1 1	0	0
6	L	1	Total O 1 1	0	0
6	M	4	Total O 4 4	0	0
6	N	3	Total O 3 3	0	0
6	O	1	Total O 1 1	0	0
6	P	1	Total O 1 1	0	0
6	Q	2	Total O 2 2	0	0
6	R	3	Total O 3 3	0	0
6	S	2	Total O 2 2	0	0
6	T	1	Total O 1 1	0	0
6	U	1	Total O 1 1	0	0
6	V	6	Total O 6 6	0	0
6	X	4	Total O 4 4	0	0

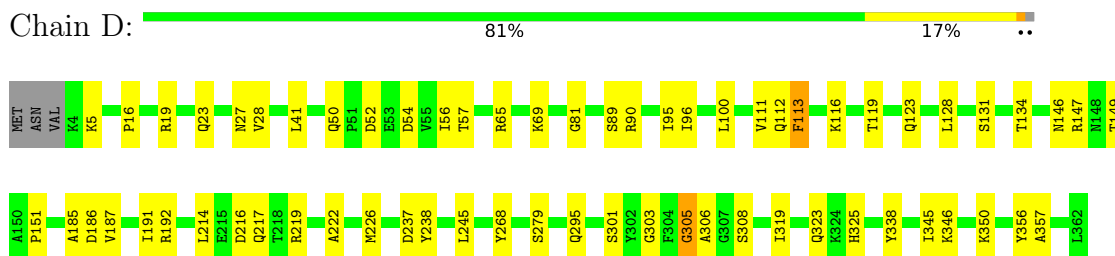
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

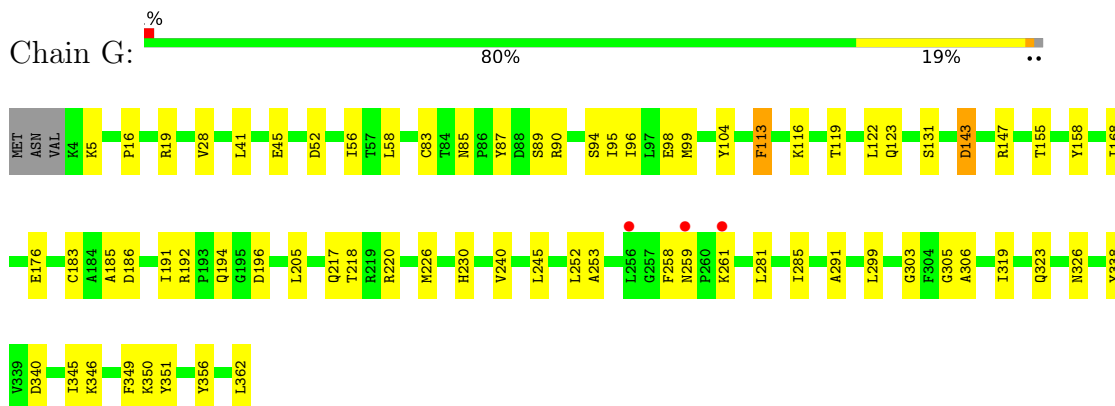
- Molecule 1: Hydroxymethylglutaryl-CoA synthase



- Molecule 1: Hydroxymethylglutaryl-CoA synthase




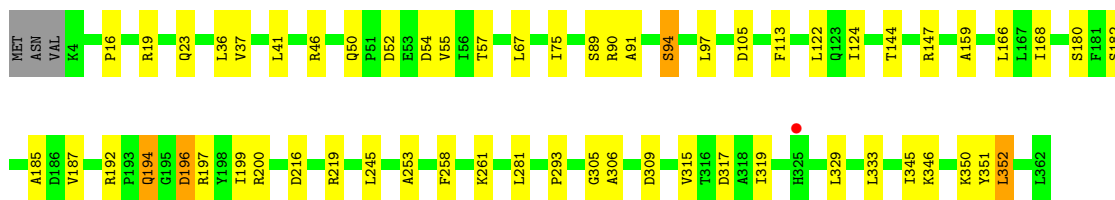
- Molecule 1: Hydroxymethylglutaryl-CoA synthase




- Molecule 1: Hydroxymethylglutaryl-CoA synthase

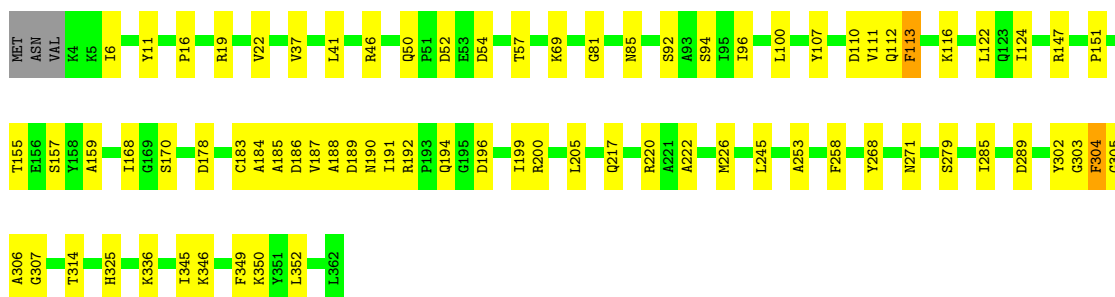


Chain J:  83% 15% ..




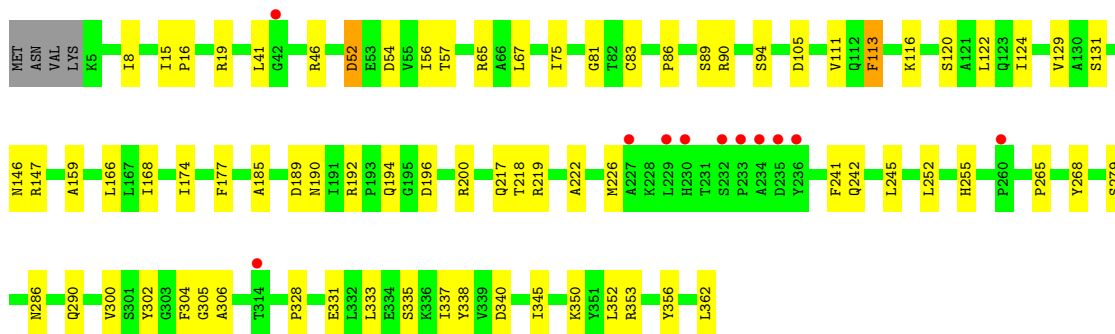
• Molecule 1: Hydroxymethylglutaryl-CoA synthase

Chain M:  78% 20% ..




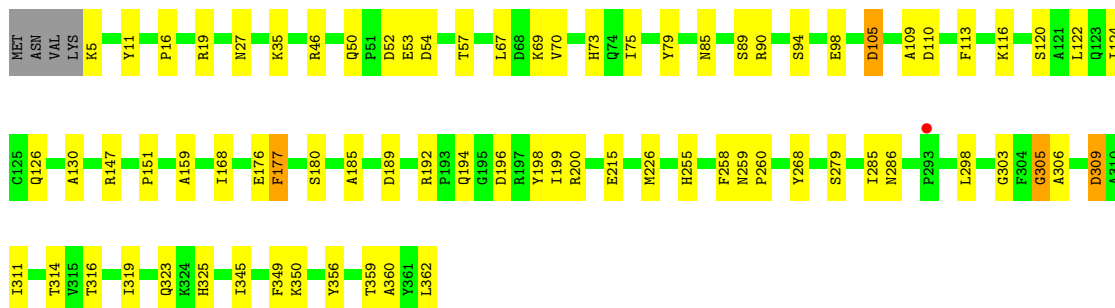
• Molecule 1: Hydroxymethylglutaryl-CoA synthase

Chain P:  3% 78% 20% ..

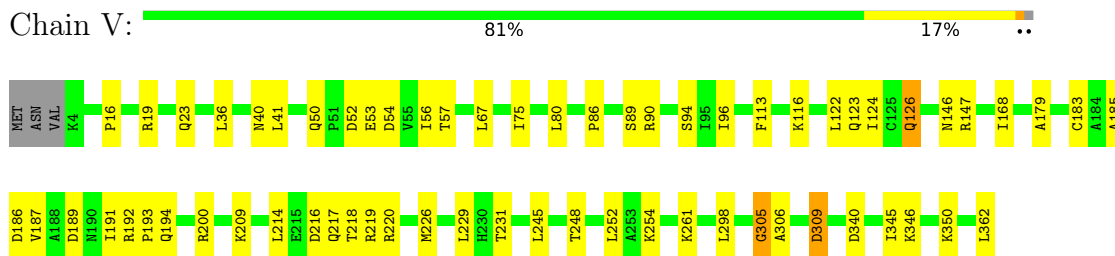


• Molecule 1: Hydroxymethylglutaryl-CoA synthase

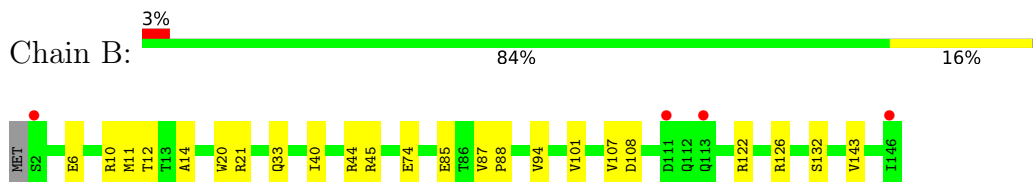
Chain S:  78% 20% ..



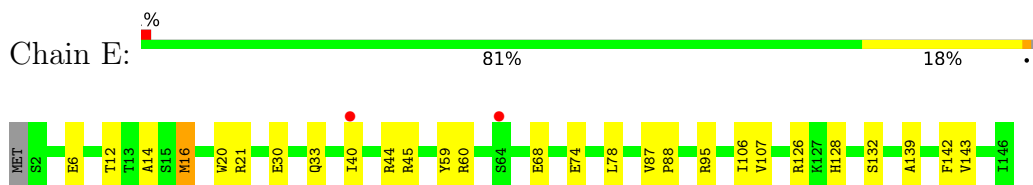
- Molecule 1: Hydroxymethylglutaryl-CoA synthase



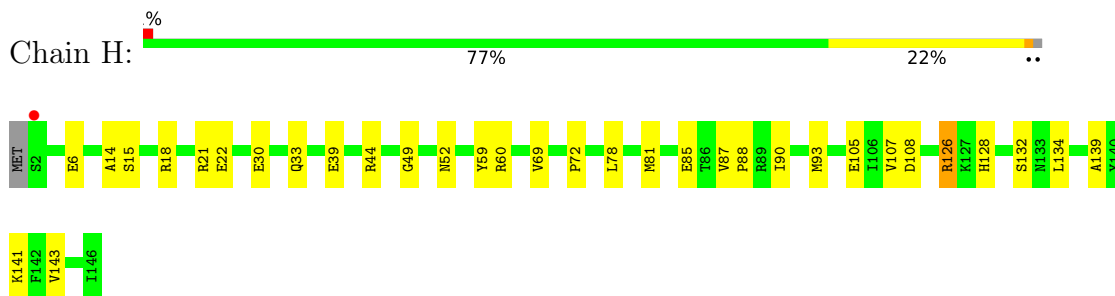
- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein



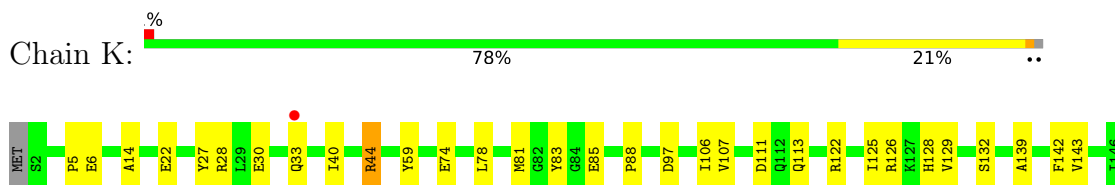
- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein



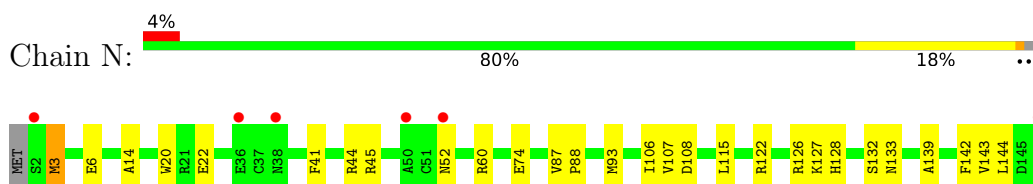
- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein



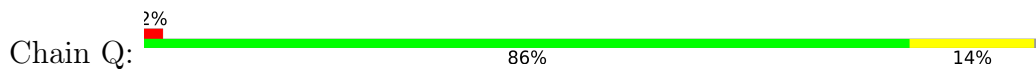
- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein



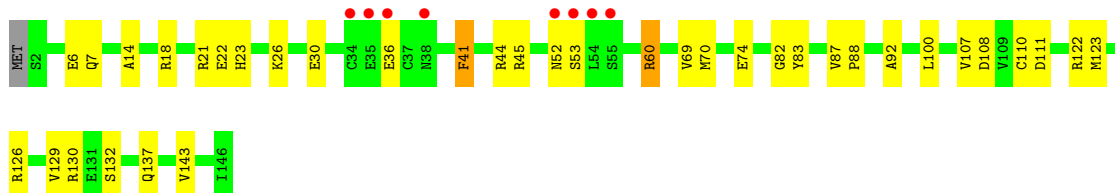
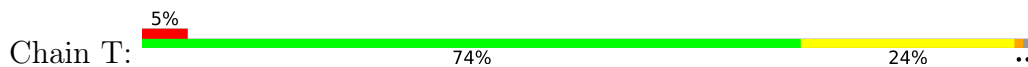
- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein



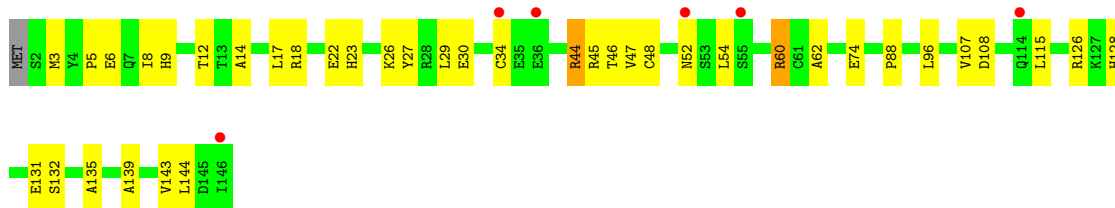
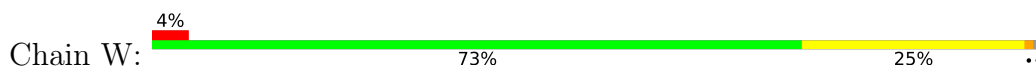
- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein



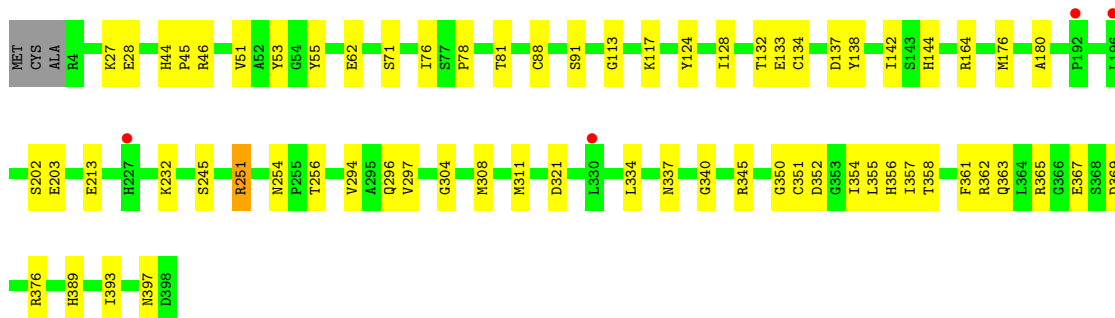
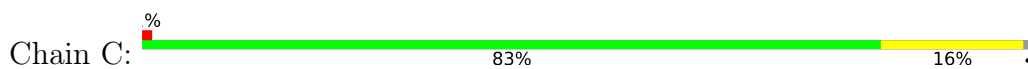
- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein



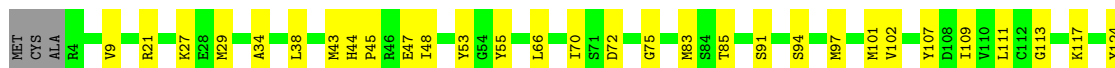
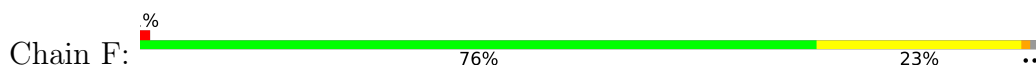
- Molecule 2: 2,4-diacetylphloroglucinol biosynthesis protein

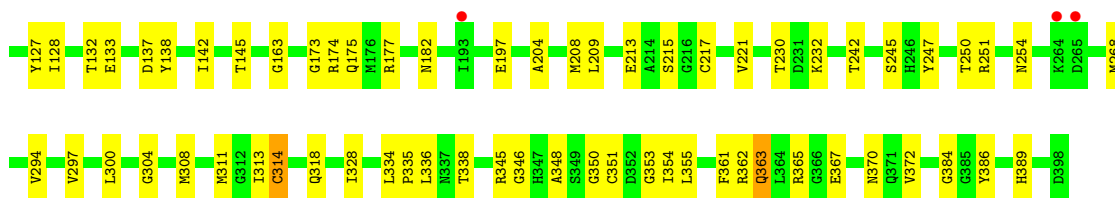


- Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein PhlC

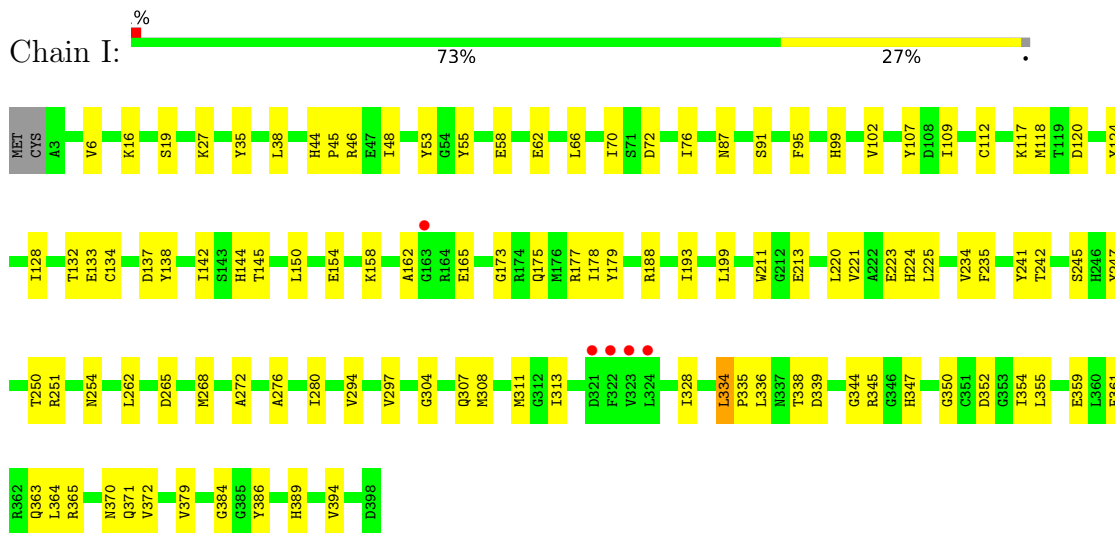


- Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein PhlC

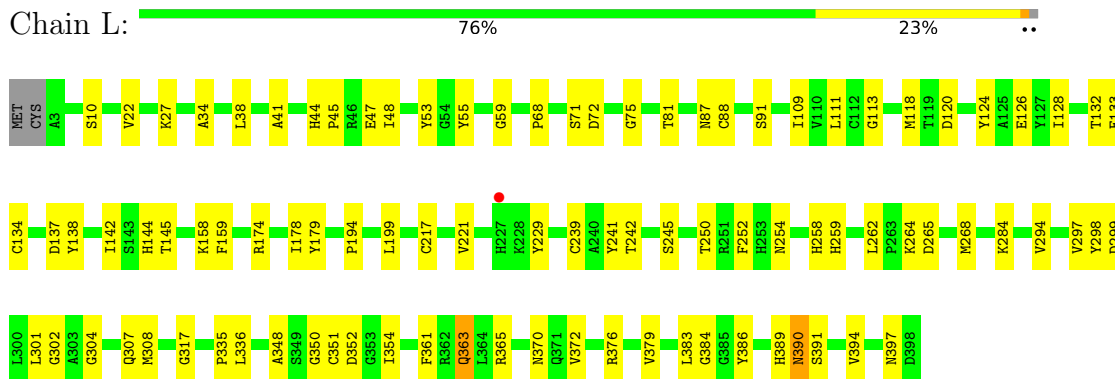




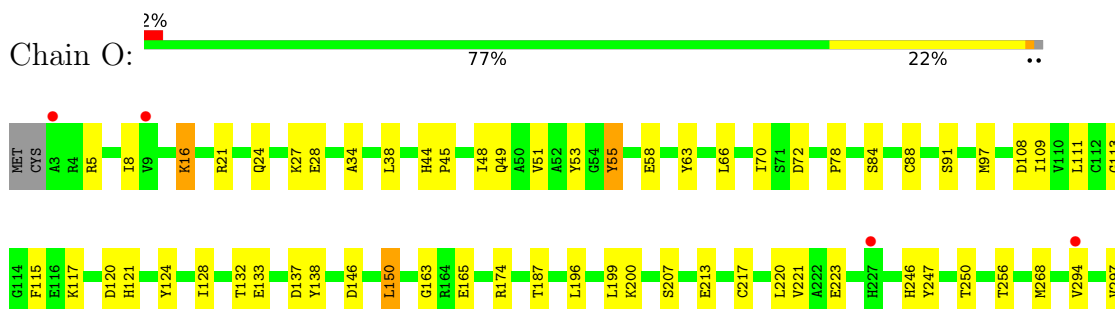
• Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein PhIC



• Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein PhIC

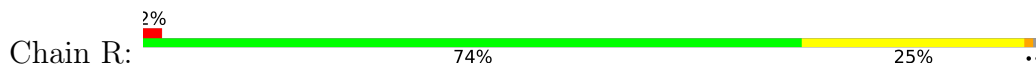


• Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein PhIC

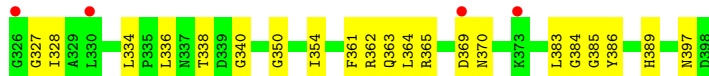
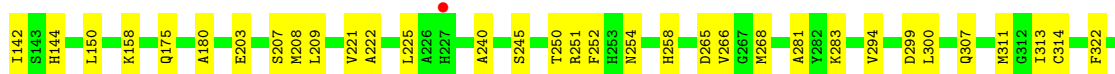
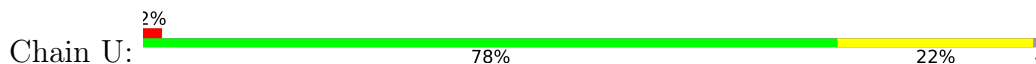




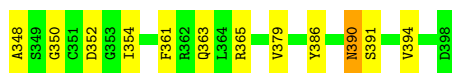
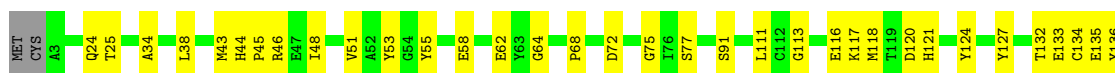
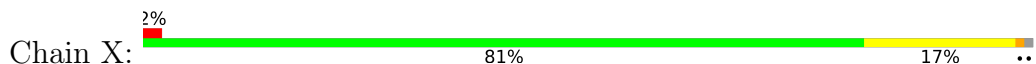
● Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein PhIC



● Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein PhIC



● Molecule 3: 2,4-diacetylphloroglucinol biosynthesis protein PhIC



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.79Å 229.78Å 311.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.73 – 3.44 49.72 – 3.44	Depositor EDS
% Data completeness (in resolution range)	99.2 (49.73-3.44) 99.2 (49.72-3.44)	Depositor EDS
$R_{merge}$	0.27	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.37 (at 3.40Å)	Xtrriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, $R_{free}$	0.165 , 0.221 0.165 , 0.220	Depositor DCC
$R_{free}$ test set	4986 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.9	Xtrriage
Anisotropy	0.743	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 50.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	54522	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% 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: ZN, 13X, SCY

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	0.25	0/2721	0.46	0/3695
1	D	0.25	0/2730	0.45	0/3706
1	G	0.26	0/2730	0.47	0/3706
1	J	0.25	0/2730	0.45	0/3706
1	M	0.25	0/2730	0.47	0/3706
1	P	0.25	0/2721	0.46	0/3695
1	S	0.25	0/2732	0.46	0/3710
1	V	0.25	0/2730	0.48	0/3706
2	B	0.25	0/1172	0.46	0/1580
2	E	0.25	0/1172	0.44	0/1580
2	H	0.24	0/1172	0.45	0/1580
2	K	0.25	0/1172	0.45	0/1580
2	N	0.24	0/1172	0.44	0/1580
2	Q	0.26	0/1172	0.46	0/1580
2	T	0.25	0/1172	0.45	0/1580
2	W	0.25	0/1172	0.46	0/1580
3	C	0.26	0/3019	0.45	0/4082
3	F	0.25	0/3019	0.44	0/4082
3	I	0.26	0/3024	0.44	0/4089
3	L	0.26	0/3024	0.45	0/4089
3	O	0.26	0/3024	0.45	0/4089
3	R	0.26	0/3019	0.45	0/4082
3	U	0.26	0/3024	0.44	0/4089
3	X	0.26	0/3024	0.47	0/4089
All	All	0.26	0/55377	0.45	0/74961

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	A	0	1
1	D	0	1
1	G	0	1
1	M	0	1
1	P	0	1
1	S	0	1
1	V	0	1
All	All	0	7

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	305	GLY	Peptide
1	D	305	GLY	Peptide
1	G	305	GLY	Peptide
1	M	305	GLY	Peptide
1	P	305	GLY	Peptide

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2677	0	2662	46	0
1	D	2686	0	2675	51	0
1	G	2686	0	2675	50	0
1	J	2686	0	2675	45	0
1	M	2686	0	2675	57	0
1	P	2677	0	2662	55	0
1	S	2687	0	2668	57	0
1	V	2686	0	2675	49	0
2	B	1150	0	1149	21	0
2	E	1150	0	1149	21	0
2	H	1150	0	1149	24	0
2	K	1150	0	1149	22	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	N	1150	0	1149	19	0
2	Q	1150	0	1149	16	0
2	T	1150	0	1149	30	0
2	W	1150	0	1149	30	0
3	C	2965	0	2869	46	0
3	F	2965	0	2869	65	0
3	I	2970	0	2874	79	0
3	L	2970	0	2874	67	0
3	O	2970	0	2874	59	0
3	R	2965	0	2869	65	0
3	U	2970	0	2874	59	0
3	X	2970	0	2874	50	0
4	B	1	0	0	0	0
4	E	1	0	0	0	0
4	H	1	0	0	0	0
4	K	1	0	0	0	0
4	N	1	0	0	0	0
4	Q	1	0	0	0	0
4	T	1	0	0	0	0
4	W	1	0	0	0	0
5	C	9	0	6	2	0
5	L	9	0	6	2	0
5	O	9	0	6	1	0
5	R	9	0	6	0	0
6	A	5	0	0	0	0
6	B	3	0	0	0	0
6	C	5	0	0	0	0
6	D	3	0	0	0	0
6	E	2	0	0	0	0
6	F	2	0	0	0	0
6	G	2	0	0	0	0
6	H	3	0	0	0	0
6	I	2	0	0	0	0
6	J	5	0	0	0	0
6	K	1	0	0	0	0
6	L	1	0	0	0	0
6	M	4	0	0	0	0
6	N	3	0	0	0	0
6	O	1	0	0	1	0
6	P	1	0	0	0	0
6	Q	2	0	0	0	0
6	R	3	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	S	2	0	0	0	0
6	T	1	0	0	0	0
6	U	1	0	0	0	0
6	V	6	0	0	0	0
6	X	4	0	0	0	0
All	All	54522	0	53560	902	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 8.

The worst 5 of 902 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)
3:U:132:THR:HG22	3:U:133:GLU:H	1.21	1.04
1:D:238:TYR:OH	1:D:295:GLN:OE1	1.76	1.03
3:L:132:THR:HG22	3:L:133:GLU:H	1.24	1.02
3:C:132:THR:HG22	3:C:133:GLU:H	1.25	0.99
3:X:132:THR:HG22	3:X:133:GLU:H	1.27	0.98

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	356/362 (98%)	345 (97%)	11 (3%)	0	100	100
1	D	357/362 (99%)	345 (97%)	12 (3%)	0	100	100
1	G	357/362 (99%)	340 (95%)	17 (5%)	0	100	100
1	J	357/362 (99%)	347 (97%)	10 (3%)	0	100	100
1	M	357/362 (99%)	347 (97%)	10 (3%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	P	356/362 (98%)	345 (97%)	11 (3%)	0	100	100
1	S	357/362 (99%)	347 (97%)	10 (3%)	0	100	100
1	V	357/362 (99%)	346 (97%)	11 (3%)	0	100	100
2	B	143/146 (98%)	139 (97%)	4 (3%)	0	100	100
2	E	143/146 (98%)	138 (96%)	5 (4%)	0	100	100
2	H	143/146 (98%)	138 (96%)	5 (4%)	0	100	100
2	K	143/146 (98%)	140 (98%)	3 (2%)	0	100	100
2	N	143/146 (98%)	139 (97%)	4 (3%)	0	100	100
2	Q	143/146 (98%)	140 (98%)	3 (2%)	0	100	100
2	T	143/146 (98%)	140 (98%)	3 (2%)	0	100	100
2	W	143/146 (98%)	139 (97%)	4 (3%)	0	100	100
3	C	392/398 (98%)	376 (96%)	16 (4%)	0	100	100
3	F	392/398 (98%)	377 (96%)	15 (4%)	0	100	100
3	I	393/398 (99%)	380 (97%)	13 (3%)	0	100	100
3	L	393/398 (99%)	381 (97%)	12 (3%)	0	100	100
3	O	393/398 (99%)	382 (97%)	11 (3%)	0	100	100
3	R	392/398 (98%)	378 (96%)	14 (4%)	0	100	100
3	U	393/398 (99%)	378 (96%)	15 (4%)	0	100	100
3	X	393/398 (99%)	381 (97%)	12 (3%)	0	100	100
All	All	7139/7248 (98%)	6908 (97%)	231 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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	279/283 (99%)	272 (98%)	7 (2%)	47	75
1	D	280/283 (99%)	275 (98%)	5 (2%)	59	81

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	280/283 (99%)	276 (99%)	4 (1%)	67	85
1	J	280/283 (99%)	272 (97%)	8 (3%)	42	71
1	M	280/283 (99%)	275 (98%)	5 (2%)	59	81
1	P	279/283 (99%)	274 (98%)	5 (2%)	59	81
1	S	280/283 (99%)	274 (98%)	6 (2%)	53	78
1	V	280/283 (99%)	276 (99%)	4 (1%)	67	85
2	B	124/125 (99%)	122 (98%)	2 (2%)	62	83
2	E	124/125 (99%)	123 (99%)	1 (1%)	81	92
2	H	124/125 (99%)	120 (97%)	4 (3%)	39	69
2	K	124/125 (99%)	122 (98%)	2 (2%)	62	83
2	N	124/125 (99%)	120 (97%)	4 (3%)	39	69
2	Q	124/125 (99%)	123 (99%)	1 (1%)	81	92
2	T	124/125 (99%)	120 (97%)	4 (3%)	39	69
2	W	124/125 (99%)	121 (98%)	3 (2%)	49	76
3	C	303/305 (99%)	292 (96%)	11 (4%)	35	66
3	F	303/305 (99%)	296 (98%)	7 (2%)	50	77
3	I	303/305 (99%)	299 (99%)	4 (1%)	69	86
3	L	303/305 (99%)	296 (98%)	7 (2%)	50	77
3	O	303/305 (99%)	296 (98%)	7 (2%)	50	77
3	R	303/305 (99%)	291 (96%)	12 (4%)	31	63
3	U	303/305 (99%)	299 (99%)	4 (1%)	69	86
3	X	303/305 (99%)	293 (97%)	10 (3%)	38	68
All	All	5654/5704 (99%)	5527 (98%)	127 (2%)	52	77

5 of 127 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	245	SER
1	V	113	PHE
3	O	16	LYS
3	U	383	LEU
3	X	55	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	M	271	ASN
1	P	23	GLN
1	V	23	GLN
1	S	323	GLN
1	J	255	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

8 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	$\# Z  > 2$
3	SCY	C	88	3	7,8,9	0.85	0	3,9,11	0.80	0
3	SCY	I	88	3	7,8,9	0.85	0	3,9,11	0.79	0
3	SCY	O	88	3	7,8,9	0.85	0	3,9,11	0.83	0
3	SCY	R	88	3	7,8,9	0.82	0	3,9,11	0.82	0
3	SCY	X	88	3	7,8,9	0.87	0	3,9,11	0.81	0
3	SCY	F	88	3	7,8,9	0.87	0	3,9,11	0.79	0
3	SCY	U	88	3	7,8,9	0.89	0	3,9,11	0.70	0
3	SCY	L	88	3	7,8,9	0.86	0	3,9,11	0.86	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
3	SCY	C	88	3	-	2/5/7/9	-
3	SCY	I	88	3	-	2/5/7/9	-
3	SCY	O	88	3	-	2/5/7/9	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SCY	R	88	3	-	2/5/7/9	-
3	SCY	X	88	3	-	2/5/7/9	-
3	SCY	F	88	3	-	2/5/7/9	-
3	SCY	U	88	3	-	2/5/7/9	-
3	SCY	L	88	3	-	2/5/7/9	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	88	SCY	OCD-CD-SG-CB
3	C	88	SCY	CE-CD-SG-CB
3	F	88	SCY	OCD-CD-SG-CB
3	F	88	SCY	CE-CD-SG-CB
3	I	88	SCY	OCD-CD-SG-CB

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	88	SCY	2	0
3	O	88	SCY	1	0
3	L	88	SCY	2	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 8 are monoatomic - leaving 4 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
5	13X	L	500	-	9,9,9	0.87	0	12,12,12	0.41	0
5	13X	O	500	-	9,9,9	0.88	0	12,12,12	0.39	0
5	13X	C	500	-	9,9,9	0.88	0	12,12,12	0.41	0
5	13X	R	500	-	9,9,9	0.89	0	12,12,12	0.39	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	13X	L	500	-	-	-	0/1/1/1
5	13X	O	500	-	-	-	0/1/1/1
5	13X	C	500	-	-	-	0/1/1/1
5	13X	R	500	-	-	-	0/1/1/1

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.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	L	500	13X	2	0
5	O	500	13X	1	0
5	C	500	13X	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	358/362 (98%)	-0.42	0 <a href="#">100</a> <a href="#">100</a>	25, 37, 57, 169	0
1	D	359/362 (99%)	-0.44	0 <a href="#">100</a> <a href="#">100</a>	27, 38, 55, 213	0
1	G	359/362 (99%)	-0.06	3 (0%) <a href="#">86</a> <a href="#">83</a>	36, 53, 71, 93	0
1	J	359/362 (99%)	-0.32	1 (0%) <a href="#">94</a> <a href="#">92</a>	27, 39, 59, 153	0
1	M	359/362 (99%)	-0.34	0 <a href="#">100</a> <a href="#">100</a>	30, 39, 57, 87	0
1	P	358/362 (98%)	0.09	11 (3%) <a href="#">49</a> <a href="#">48</a>	34, 49, 65, 84	0
1	S	358/362 (98%)	-0.30	1 (0%) <a href="#">94</a> <a href="#">92</a>	29, 41, 58, 85	0
1	V	359/362 (99%)	-0.29	0 <a href="#">100</a> <a href="#">100</a>	34, 46, 64, 116	0
2	B	145/146 (99%)	-0.11	4 (2%) <a href="#">53</a> <a href="#">51</a>	30, 44, 67, 87	0
2	E	145/146 (99%)	0.04	2 (1%) <a href="#">75</a> <a href="#">72</a>	32, 45, 69, 79	0
2	H	145/146 (99%)	-0.30	1 (0%) <a href="#">87</a> <a href="#">85</a>	32, 48, 70, 96	0
2	K	145/146 (99%)	-0.20	1 (0%) <a href="#">87</a> <a href="#">85</a>	30, 45, 65, 74	0
2	N	145/146 (99%)	0.02	6 (4%) <a href="#">37</a> <a href="#">36</a>	33, 46, 72, 92	0
2	Q	145/146 (99%)	-0.05	3 (2%) <a href="#">63</a> <a href="#">62</a>	35, 48, 78, 90	0
2	T	145/146 (99%)	0.01	8 (5%) <a href="#">25</a> <a href="#">26</a>	33, 47, 72, 85	0
2	W	145/146 (99%)	0.18	6 (4%) <a href="#">37</a> <a href="#">36</a>	38, 55, 88, 95	0
3	C	394/398 (98%)	-0.02	4 (1%) <a href="#">82</a> <a href="#">79</a>	32, 47, 59, 75	0
3	F	394/398 (98%)	-0.17	3 (0%) <a href="#">86</a> <a href="#">83</a>	37, 49, 65, 88	0
3	I	395/398 (99%)	0.00	5 (1%) <a href="#">77</a> <a href="#">74</a>	40, 56, 71, 155	0
3	L	395/398 (99%)	-0.16	1 (0%) <a href="#">94</a> <a href="#">92</a>	36, 51, 69, 88	0
3	O	395/398 (99%)	0.08	9 (2%) <a href="#">60</a> <a href="#">58</a>	37, 50, 64, 81	0
3	R	394/398 (98%)	0.00	6 (1%) <a href="#">73</a> <a href="#">71</a>	38, 52, 66, 94	0
3	U	395/398 (99%)	-0.05	6 (1%) <a href="#">73</a> <a href="#">71</a>	33, 48, 63, 88	0
3	X	395/398 (99%)	-0.01	7 (1%) <a href="#">68</a> <a href="#">66</a>	42, 57, 76, 93	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	7186/7248 (99%)	-0.13	88 (1%) 79 76	25, 48, 68, 213	0

The worst 5 of 88 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	T	35	GLU	3.7
3	R	265	ASP	3.7
3	X	265	ASP	3.6
2	T	38	ASN	3.6
2	T	55	SER	3.5

## 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
3	SCY	U	88	9/10	0.92	0.23	37,48,62,64	0
3	SCY	O	88	9/10	0.93	0.20	37,44,52,53	0
3	SCY	I	88	9/10	0.93	0.30	51,53,69,75	0
3	SCY	X	88	9/10	0.93	0.27	48,51,59,63	0
3	SCY	L	88	9/10	0.94	0.26	40,44,54,57	0
3	SCY	C	88	9/10	0.94	0.21	44,45,56,57	0
3	SCY	F	88	9/10	0.95	0.29	41,44,70,77	0
3	SCY	R	88	9/10	0.95	0.17	48,52,58,58	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides 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
5	13X	O	500	9/9	0.85	0.29	50,56,60,63	0
5	13X	R	500	9/9	0.85	0.26	52,56,61,63	0
5	13X	C	500	9/9	0.90	0.23	42,56,59,61	0
5	13X	L	500	9/9	0.91	0.23	50,56,58,60	0
4	ZN	B	200	1/1	0.94	0.04	116,116,116,116	0
4	ZN	T	200	1/1	0.94	0.14	132,132,132,132	0
4	ZN	H	200	1/1	0.95	0.04	81,81,81,81	0
4	ZN	W	200	1/1	0.95	0.16	135,135,135,135	0
4	ZN	Q	200	1/1	0.97	0.07	162,162,162,162	0
4	ZN	K	200	1/1	0.99	0.09	104,104,104,104	0
4	ZN	N	200	1/1	0.99	0.14	97,97,97,97	0
4	ZN	E	200	1/1	0.99	0.08	96,96,96,96	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.