



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 21, 2016 – 01:43 PM EDT

PDB ID : 5DX0  
Title : Crystal structure of CARM1, sinefungin, and H3 peptide (R17)  
Authors : Boriack-Sjodin, P.A.  
Deposited on : 2015-09-23  
Resolution : 2.05 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027107  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0122  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027107

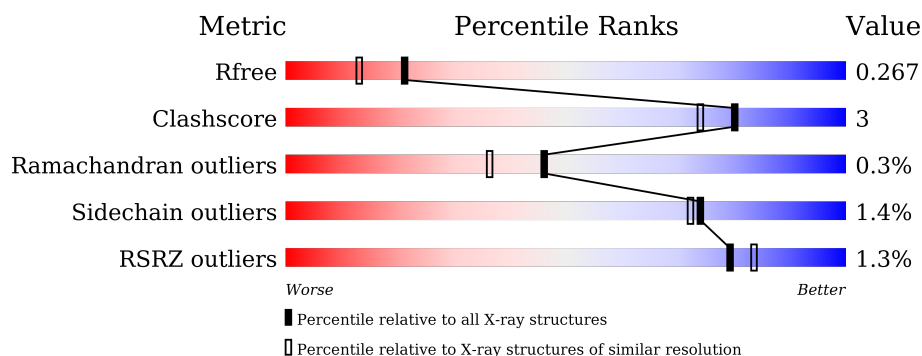
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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))
$R_{free}$	91344	1192 (2.04-2.04)
Clashscore	102246	1269 (2.04-2.04)
Ramachandran outliers	100387	1258 (2.04-2.04)
Sidechain outliers	100360	1258 (2.04-2.04)
RSRZ outliers	91569	1194 (2.04-2.04)

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	349	<div> <div></div> <div>90%</div> <div>8%</div> <div>.</div> </div>
1	B	349	<div> <div></div> <div>89%</div> <div>9%</div> <div>.</div> </div>
1	C	349	<div> <div></div> <div>90%</div> <div>8%</div> <div>.</div> </div>
1	D	349	<div> <div></div> <div>88%</div> <div>10%</div> <div>..</div> </div>
2	F	20	<div> <div>10%</div> <div>30%</div> <div>15%</div> <div>55%</div> </div>
2	G	20	<div> <div>40%</div> <div>10%</div> <div>50%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	20	
2	I	20	

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
5	SO4	C	503	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11892 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 Histone-arginine methyltransferase CARM1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	343	Total	C	N	O	S	0	1	0
			2756	1779	454	509	14			
1	B	343	Total	C	N	O	S	0	2	0
			2761	1783	454	509	15			
1	C	343	Total	C	N	O	S	0	0	0
			2751	1776	454	507	14			
1	D	343	Total	C	N	O	S	0	1	0
			2756	1779	454	509	14			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	131	SER	-	expression tag	UNP Q86X55
A	132	ILE	-	expression tag	UNP Q86X55
A	133	ALA	-	expression tag	UNP Q86X55
B	131	SER	-	expression tag	UNP Q86X55
B	132	ILE	-	expression tag	UNP Q86X55
B	133	ALA	-	expression tag	UNP Q86X55
C	131	SER	-	expression tag	UNP Q86X55
C	132	ILE	-	expression tag	UNP Q86X55
C	133	ALA	-	expression tag	UNP Q86X55
D	131	SER	-	expression tag	UNP Q86X55
D	132	ILE	-	expression tag	UNP Q86X55
D	133	ALA	-	expression tag	UNP Q86X55

- Molecule 2 is a protein called H3 peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	F	9	Total	C	N	O	0	0	0
			65	41	14	10			
2	G	10	Total	C	N	O	0	0	0
			70	44	15	11			

*Continued on next page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	H	9	Total	C	N	O	0	0	0
			65	41	14	10			
2	I	10	Total	C	N	O	0	0	0
			70	44	15	11			

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	ACE	-	acetylation	UNP P84243
F	19	NH2	-	amidation	UNP P84243
G	0	ACE	-	acetylation	UNP P84243
G	19	NH2	-	amidation	UNP P84243
H	0	ACE	-	acetylation	UNP P84243
H	19	NH2	-	amidation	UNP P84243
I	0	ACE	-	acetylation	UNP P84243
I	19	NH2	-	amidation	UNP P84243

- # SFG
- 
- The chemical structure of SFG (Sarcosine Glycylglycine) is shown. It features a pyrazole ring system (C1-C8, N1-N2) linked to a ribose sugar (C1' to C5'). The ribose sugar is further linked to a glycine residue (C6, C7, C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19, C20, C21, C22, C23, C24, C25, C26, C27, C28, C29, C30, C31, C32, C33, C34, C35, C36, C37, C38, C39, C40, C41, C42, C43, C44, C45, C46, C47, C48, C49, C50, C51, C52, C53, C54, C55, C56, C57, C58, C59, C60, C61, C62, C63, C64, C65, C66, C67, C68, C69, C70, C71, C72, C73, C74, C75, C76, C77, C78, C79, C80, C81, C82, C83, C84, C85, C86, C87, C88, C89, C90, C91, C92, C93, C94, C95, C96, C97, C98, C99, C100, C101, C102, C103, C104, C105, C106, C107, C108, C109, C110, C111, C112, C113, C114, C115, C116, C117, C118, C119, C120, C121, C122, C123, C124, C125, C126, C127, C128, C129, C130, C131, C132, C133, C134, C135, C136, C137, C138, C139, C140, C141, C142, C143, C144, C145, C146, C147, C148, C149, C150, C151, C152, C153, C154, C155, C156, C157, C158, C159, C160, C161, C162, C163, C164, C165, C166, C167, C168, C169, C170, C171, C172, C173, C174, C175, C176, C177, C178, C179, C180, C181, C182, C183, C184, C185, C186, C187, C188, C189, C190, C191, C192, C193, C194, C195, C196, C197, C198, C199, C200, C201, C202, C203, C204, C205, C206, C207, C208, C209, C210, C211, C212, C213, C214, C215, C216, C217, C218, C219, C220, C221, C222, C223, C224, C225, C226, C227, C228, C229, C230, C231, C232, C233, C234, C235, C236, C237, C238, C239, C240, C241, C242, C243, C244, C245, C246, C247, C248, C249, C250, C251, C252, C253, C254, C255, C256, C257, C258, C259, C260, C261, C262, C263, C264, C265, C266, C267, C268, C269, C270, C271, C272, C273, C274, C275, C276, C277, C278, C279, C280, C281, C282, C283, C284, C285, C286, C287, C288, C289, C290, C291, C292, C293, C294, C295, C296, C297, C298, C299, C300, C301, C302, C303, C304, C305, C306, C307, C308, C309, C310, C311, C312, C313, C314, C315, C316, C317, C318, C319, C320, C321, C322, C323, C324, C325, C326, C327, C328, C329, C330, C331, C332, C333, C334, C335, C336, C337, C338, C339, C340, C341, C342, C343, C344, C345, C346, C347, C348, C349, C350, C351, C352, C353, C354, C355, C356, C357, C358, C359, C360, C361, C362, C363, C364, C365, C366, C367, C368, C369, C370, C371, C372, C373, C374, C375, C376, C377, C378, C379, C380, C381, C382, C383, C384, C385, C386, C387, C388, C389, C390, C391, C392, C393, C394, C395, C396, C397, C398, C399, C400, C401, C402, C403, C404, C405, C406, C407, C408, C409, C410, C411, C412, C413, C414, C415, C416, C417, C418, C419, C420, C421, C422, C423, C424, C425, C426, C427, C428, C429, C430, C431, C432, C433, C434, C435, C436, C437, C438, C439, C440, C441, C442, C443, C444, C445, C446, C447, C448, C449, C450, C451, C452, C453, C454, C455, C456, C457, C458, C459, C460, C461, C462, C463, C464, C465, C466, C467, C468, C469, C470, C471, C472, C473, C474, C475, C476, C477, C478, C479, C480, C481, C482, C483, C484, C485, C486, C487, C488, C489, C490, C491, C492, C493, C494, C495, C496, C497, C498, C499, C500, C501, C502, C503, C504, C505, C506, C507, C508, C509, C510, C511, C512, C513, C514, C515, C516, C517, C518, C519, C520, C521, C522, C523, C524, C525, C526, C527, C528, C529, C530, C531, C532, C533, C534, C535, C536, C537, C538, C539, C540, C541, C542, C543, C544, C545, C546, C547, C548, C549, C550, C551, C552, C553, C554, C555, C556, C557, C558, C559, C560, C561, C562, C563, C564, C565, C566, C567, C568, C569, C570, C571, C572, C573, C574, C575, C576, C577, C578, C579, C580, C581, C582, C583, C584, C585, C586, C587, C588, C589, C590, C591, C592, C593, C594, C595, C596, C597, C598, C599, C600, C601, C602, C603, C604, C605, C606, C607, C608, C609, C610, C611, C612, C613, C614, C615, C616, C617, C618, C619, C620, C621, C622, C623, C624, C625, C626, C627, C628, C629, C630, C631, C632, C633, C634, C635, C636, C637, C638, C639, C640, C641, C642, C643, C644, C645, C646, C647, C648, C649, C650, C651, C652, C653, C654, C655, C656, C657, C658, C659, C660, C661, C662, C663, C664, C665, C666, C667, C668, C669, C670, C671, C672, C673, C674, C675, C676, C677, C678, C679, C680, C681, C682, C683, C684, C685, C686, C687, C688, C689, C690, C691, C692, C693, C694, C695, C696, C697, C698, C699, C700, C701, C702, C703, C704, C705, C706, C707, C708, C709, C710, C711, C712, C713, C714, C715, C716, C717, C718, C719, C720, C721, C722, C723, C724, C725, C726, C727, C728, C729, C730, C731, C732, C733, C734, C735, C736, C737, C738, C739, C740, C741, C742, C743, C744, C745, C746, C747, C748, C749, C750, C751, C752, C753, C754, C755, C756, C757, C758, C759, C760, C761, C762, C763, C764, C765, C766, C767, C768, C769, C770, C771, C772, C773, C774, C775, C776, C777, C778, C779, C780, C781, C782, C783, C784, C785, C786, C787, C788, C789, C790, C791, C792, C793, C794, C795, C796, C797, C798, C799, C800, C801, C802, C803, C804, C805, C806, C807, C808, C809, C810, C811, C812, C813, C814, C815, C816, C817, C81

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 27	C 15	N 7	O 5	0	0
3	B	1	Total 27	C 15	N 7	O 5	0	0
3	C	1	Total 27	C 15	N 7	O 5	0	0

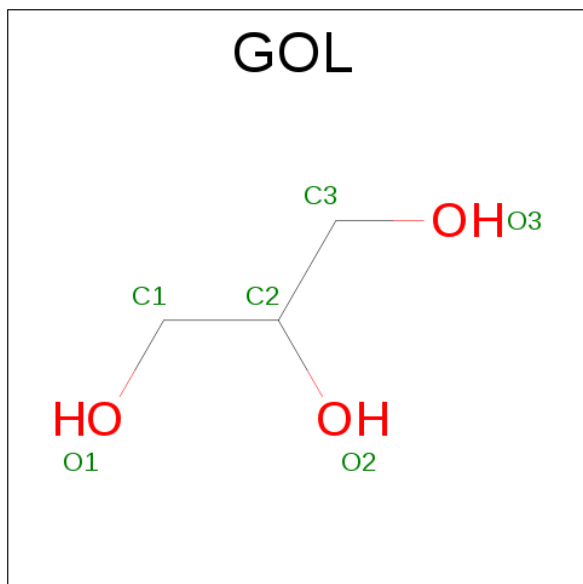


WORLD WIDE  
PDB  
PROTEIN DATA BANK

*Continued from previous page...*

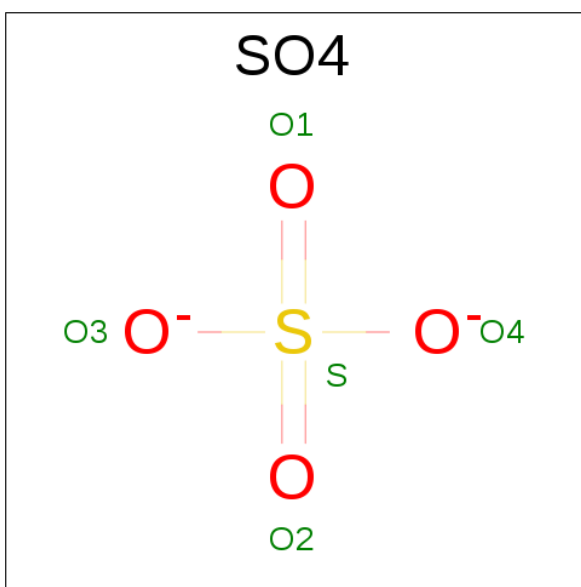
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	D	1	Total	C	N	O	0	0
			27	15	7	5		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		

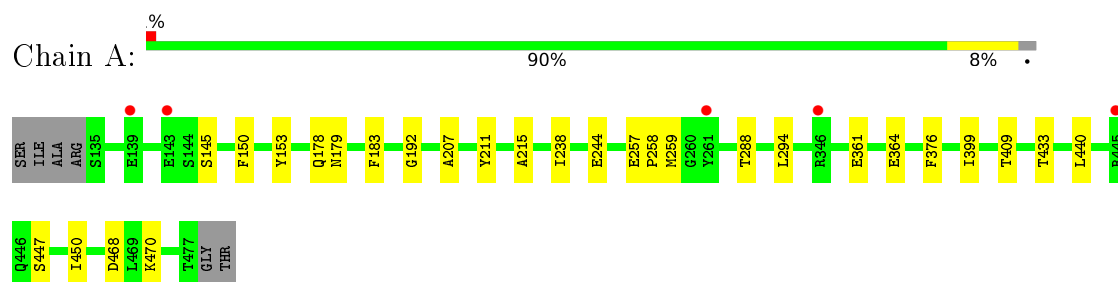
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	94	Total	O	0	0
			94	94		
6	B	120	Total	O	0	0
			120	120		
6	C	109	Total	O	0	0
			109	109		
6	D	113	Total	O	0	0
			113	113		
6	F	2	Total	O	0	0
			2	2		
6	G	8	Total	O	0	0
			8	8		
6	H	3	Total	O	0	0
			3	3		
6	I	9	Total	O	0	0
			9	9		

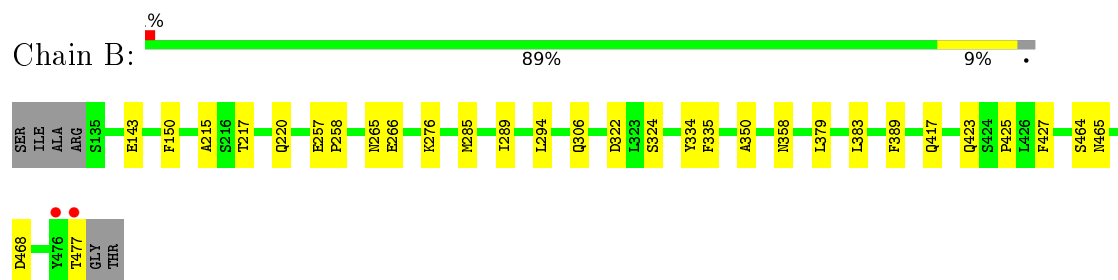
### 3 Residue-property plots [i](#)

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.

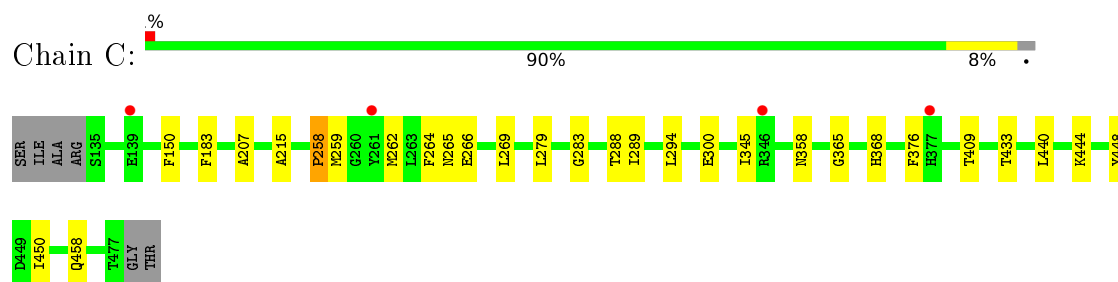
- Molecule 1: Histone-arginine methyltransferase CARM1



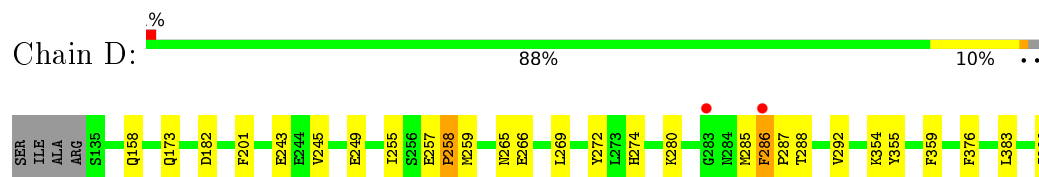
- Molecule 1: Histone-arginine methyltransferase CARM1



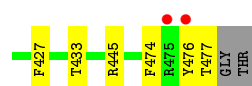
- Molecule 1: Histone-arginine methyltransferase CARM1



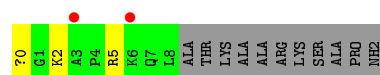
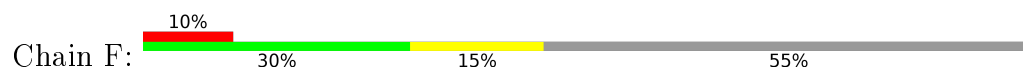
- Molecule 1: Histone-arginine methyltransferase CARM1



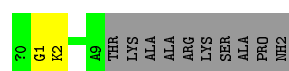




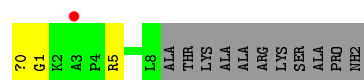
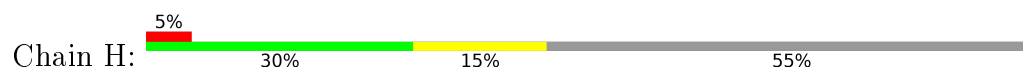
- Molecule 2: H3 peptide



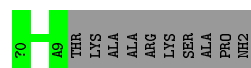
- Molecule 2: H3 peptide



- Molecule 2: H3 peptide



- Molecule 2: H3 peptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.72Å 98.85Å 208.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.05 46.14 – 2.05	Depositor EDS
% Data completeness (in resolution range)	89.8 (50.00-2.05) 89.8 (46.14-2.05)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.27 (at 2.05Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.211 , 0.263 0.217 , 0.267	Depositor DCC
$R_{free}$ test set	4324 reflections (5.11%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.2	Xtriage
Anisotropy	0.100	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 32.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	8 of 88891 reflections (0.009%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11892	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 79.70 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 5.2649e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, SO4, ACE, SFG

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	0.52	0/2829	0.70	0/3834
1	B	0.51	0/2837	0.68	0/3844
1	C	0.54	0/2821	0.68	0/3823
1	D	0.52	0/2829	0.68	0/3834
2	F	0.47	0/63	0.78	0/82
2	G	0.52	0/68	0.75	0/89
2	H	0.55	0/63	0.72	0/82
2	I	0.44	0/68	0.64	0/89
All	All	0.52	0/11578	0.69	0/15677

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2756	0	2703	19	0
1	B	2761	0	2712	16	0
1	C	2751	0	2699	16	0
1	D	2756	0	2703	24	0
2	F	65	0	76	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	G	70	0	81	2	0
2	H	65	0	76	4	0
2	I	70	0	81	0	0
3	A	27	0	22	2	0
3	B	27	0	22	0	0
3	C	27	0	22	0	0
3	D	27	0	22	0	0
4	A	6	0	8	0	0
4	C	6	0	8	0	0
5	A	5	0	0	0	0
5	B	5	0	0	1	0
5	C	5	0	0	1	0
5	D	5	0	0	1	0
6	A	94	0	0	4	0
6	B	120	0	0	3	0
6	C	109	0	0	1	0
6	D	113	0	0	4	0
6	F	2	0	0	0	0
6	G	8	0	0	0	0
6	H	3	0	0	0	0
6	I	9	0	0	0	0
All	All	11892	0	11235	78	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 3.

All (78) 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:B:468[A]:ASP:OD1	6:B:601:HOH:O	1.82	0.97
1:C:265:ASN:HD21	2:H:0:ACE:H3	1.42	0.85
1:D:272:TYR:OH	1:D:287:PRO:O	2.09	0.71
1:A:294:LEU:HD12	1:A:376:PHE:HZ	1.55	0.70
1:D:474:PHE:HB3	6:D:603:HOH:O	1.94	0.66
1:A:364:GLU:OE1	6:A:601:HOH:O	2.14	0.66
1:D:383:LEU:CD2	1:D:427:PHE:CD2	2.79	0.65
1:D:445:ARG:NH1	5:D:502:SO4:O4	2.32	0.62
1:D:383:LEU:HD21	1:D:427:PHE:CE2	2.35	0.61
1:B:217:THR:O	1:B:220:GLN:HG2	2.00	0.61
1:A:145:SER:HA	6:A:674:HOH:O	2.03	0.59
1:D:476:TYR:O	6:D:601:HOH:O	2.16	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:468[A]:ASP:OD1	1:A:470:LYS:HG2	2.07	0.54
1:D:383:LEU:HD21	1:D:427:PHE:CD2	2.42	0.53
1:A:440:LEU:HD22	1:A:450:ILE:HG12	1.91	0.53
1:A:294:LEU:HD12	1:A:376:PHE:CZ	2.42	0.53
1:A:150:PHE:CE1	1:A:215:ALA:HB3	2.44	0.52
1:B:383:LEU:HD21	1:B:427:PHE:CE2	2.45	0.51
2:F:2:LYS:HD2	2:G:2:LYS:HB3	1.92	0.51
1:A:244:GLU:HB2	6:A:671:HOH:O	2.10	0.51
1:B:294:LEU:HG	1:B:389:PHE:CE2	2.45	0.50
5:B:502:SO4:O3	2:G:1:GLY:HA2	2.12	0.50
1:C:365:GLY:HA2	1:C:368:HIS:CE1	2.47	0.50
1:C:289:ILE:HG22	1:C:358:ASN:HA	1.93	0.50
1:D:285:MET:O	1:D:288:THR:OG1	2.25	0.50
1:A:294:LEU:CD1	1:A:376:PHE:HZ	2.25	0.49
5:C:503:SO4:O1	2:H:1:GLY:HA2	2.13	0.49
1:B:265:ASN:O	1:B:266:GLU:HB2	2.12	0.48
1:D:265:ASN:O	1:D:266:GLU:HB2	2.14	0.48
1:B:464:SER:C	1:B:465:ASN:ND2	2.67	0.48
1:B:289:ILE:HG22	1:B:358:ASN:HA	1.96	0.47
1:D:383:LEU:CD2	1:D:427:PHE:CE2	2.98	0.47
1:A:211:TYR:HB3	1:A:238:ILE:HD12	1.95	0.47
1:A:447:SER:HB2	2:F:0:ACE:H3	1.96	0.47
1:D:182:ASP:OD1	1:D:398:SER:OG	2.33	0.47
1:A:433:THR:HG22	6:A:686:HOH:O	2.15	0.46
1:C:150:PHE:CE1	1:C:215:ALA:HB3	2.51	0.46
1:A:376:PHE:O	1:A:433:THR:HA	2.15	0.46
1:D:258:PRO:HD2	1:D:259:MET:SD	2.56	0.46
1:A:447:SER:HB2	2:F:0:ACE:CH3	2.46	0.45
1:C:376:PHE:O	1:C:433:THR:HA	2.15	0.45
1:C:258:PRO:HG2	1:C:269:LEU:HD23	1.99	0.45
1:D:255:ILE:HA	1:D:286:PHE:O	2.16	0.45
1:C:440:LEU:CD2	1:C:450:ILE:HG12	2.47	0.45
1:A:153:TYR:CD1	3:A:501:SFG:H3'	2.53	0.44
1:A:183:PHE:O	1:A:207:ALA:HA	2.17	0.44
1:D:396:ILE:HA	1:D:401:THR:OG1	2.17	0.44
1:D:376:PHE:O	1:D:433:THR:HA	2.18	0.44
1:B:276:LYS:HD3	1:B:285[A]:MET:SD	2.57	0.44
1:C:279:LEU:HD11	1:C:283:GLY:HA3	2.00	0.44
1:B:465:ASN:HB3	6:B:696:HOH:O	2.17	0.43
1:B:306:GLN:HE21	1:B:334:TYR:HB3	1.83	0.43
1:C:458:GLN:OE1	1:C:458:GLN:N	2.49	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:183:PHE:O	1:C:207:ALA:HA	2.19	0.43
1:C:259:MET:O	2:H:5:ARG:NH1	2.52	0.43
1:C:444:LYS:HG2	6:C:676:HOH:O	2.19	0.43
1:D:292:VAL:O	1:D:354:LYS:HA	2.19	0.43
1:B:423:GLN:HB3	6:B:619:HOH:O	2.18	0.42
1:D:245:VAL:O	6:D:602:HOH:O	2.21	0.42
1:B:322:ASP:OD1	1:B:324:SER:OG	2.30	0.42
1:D:269:LEU:HD22	1:D:359:PHE:CE1	2.54	0.42
1:A:259:MET:O	2:F:5:ARG:NH1	2.52	0.42
1:C:264:PHE:HB3	1:C:448:TYR:CE1	2.54	0.42
1:D:158:GLN:NE2	1:D:158:GLN:HA	2.35	0.42
1:D:243:GLU:HB3	1:D:274:HIS:CD2	2.54	0.42
1:B:150:PHE:CE1	1:B:215:ALA:HB3	2.55	0.42
1:B:383:LEU:HD22	1:B:425:PRO:HB2	2.01	0.42
1:D:173:GLN:HG3	1:D:201:PHE:CD1	2.55	0.42
1:C:265:ASN:O	1:C:266:GLU:HB2	2.19	0.41
1:A:179:ASN:HA	1:A:399:ILE:HD12	2.02	0.41
1:B:335:PHE:CE1	1:B:423:GLN:HA	2.55	0.41
1:D:292:VAL:CG1	1:D:355:TYR:HB3	2.51	0.41
1:A:192:GLY:HA3	3:A:501:SFG:O4'	2.20	0.41
1:D:280:LYS:HA	6:D:657:HOH:O	2.20	0.41
1:B:350:ALA:HA	1:B:379:LEU:HD12	2.03	0.41
1:C:345:ILE:HG21	1:C:409:THR:HG23	2.02	0.40
1:D:265:ASN:O	1:D:266:GLU:CB	2.69	0.40
1:C:265:ASN:ND2	2:H:0:ACE:H3	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	342/349 (98%)	329 (96%)	12 (4%)	1 (0%)	46 36

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	343/349 (98%)	332 (97%)	10 (3%)	1 (0%)	46	36
1	C	341/349 (98%)	328 (96%)	12 (4%)	1 (0%)	46	36
1	D	342/349 (98%)	328 (96%)	13 (4%)	1 (0%)	46	36
2	F	7/20 (35%)	7 (100%)	0	0	100	100
2	G	8/20 (40%)	8 (100%)	0	0	100	100
2	H	7/20 (35%)	7 (100%)	0	0	100	100
2	I	8/20 (40%)	8 (100%)	0	0	100	100
All	All	1398/1476 (95%)	1347 (96%)	47 (3%)	4 (0%)	46	36

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	258	PRO
1	B	258	PRO
1	D	258	PRO
1	A	258	PRO

### 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	300/303 (99%)	295 (98%)	5 (2%)	68	65
1	B	301/303 (99%)	297 (99%)	4 (1%)	76	74
1	C	299/303 (99%)	295 (99%)	4 (1%)	76	74
1	D	300/303 (99%)	296 (99%)	4 (1%)	76	74
2	F	6/12 (50%)	6 (100%)	0	100	100
2	G	6/12 (50%)	6 (100%)	0	100	100
2	H	6/12 (50%)	6 (100%)	0	100	100
2	I	6/12 (50%)	6 (100%)	0	100	100
All	All	1224/1260 (97%)	1207 (99%)	17 (1%)	74	72

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	178	GLN
1	A	257	GLU
1	A	288	THR
1	A	361	GLU
1	A	409	THR
1	B	143	GLU
1	B	257	GLU
1	B	417	GLN
1	B	477	THR
1	C	262	MET
1	C	288	THR
1	C	294	LEU
1	C	300	GLU
1	D	249	GLU
1	D	257	GLU
1	D	286	PHE
1	D	477	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (25) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	151	GLN
1	A	164	GLN
1	A	311	ASN
1	A	417	GLN
1	A	455	GLN
1	B	178	GLN
1	B	306	GLN
1	B	417	GLN
1	B	455	GLN
1	C	151	GLN
1	C	160	GLN
1	C	164	GLN
1	C	178	GLN
1	C	265	ASN
1	C	368	HIS
1	D	148	GLN
1	D	151	GLN
1	D	158	GLN
1	D	164	GLN
1	D	178	GLN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	D	250	GLN
1	D	306	GLN
1	D	455	GLN
1	D	471	ASN
2	F	7	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

10 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	$\# Z  > 2$
3	SFG	A	501	-	23,29,29	1.10	2 (8%)	20,42,42	2.29	3 (15%)
4	GOL	A	502	-	5,5,5	0.22	0	5,5,5	0.30	0
5	SO4	A	503	-	4,4,4	0.87	0	6,6,6	0.25	0
3	SFG	B	501	-	23,29,29	1.09	4 (17%)	20,42,42	2.15	3 (15%)
5	SO4	B	502	-	4,4,4	0.39	0	6,6,6	0.64	0
3	SFG	C	501	-	23,29,29	0.92	1 (4%)	20,42,42	1.98	3 (15%)
4	GOL	C	502	-	5,5,5	0.21	0	5,5,5	0.29	0
5	SO4	C	503	-	4,4,4	0.78	0	6,6,6	0.38	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	SFG	D	501	-	23,29,29	1.04	2 (8%)	20,42,42	2.20	4 (20%)
5	SO4	D	502	-	4,4,4	0.66	0	6,6,6	0.35	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	SFG	A	501	-	-	0/9/33/33	0/3/3/3
4	GOL	A	502	-	-	0/4/4/4	0/0/0/0
5	SO4	A	503	-	-	0/0/0/0	0/0/0/0
3	SFG	B	501	-	-	0/9/33/33	0/3/3/3
5	SO4	B	502	-	-	0/0/0/0	0/0/0/0
3	SFG	C	501	-	-	0/9/33/33	0/3/3/3
4	GOL	C	502	-	-	0/4/4/4	0/0/0/0
5	SO4	C	503	-	-	0/0/0/0	0/0/0/0
3	SFG	D	501	-	-	0/9/33/33	0/3/3/3
5	SO4	D	502	-	-	0/0/0/0	0/0/0/0

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	501	SFG	C2'-C1'	-2.55	1.49	1.53
3	B	501	SFG	C2'-C1'	-2.15	1.50	1.53
3	B	501	SFG	CD-NE	-2.06	1.41	1.47
3	B	501	SFG	C2-N3	2.08	1.35	1.32
3	A	501	SFG	O4'-C1'	2.33	1.44	1.41
3	C	501	SFG	C5-C4	2.45	1.46	1.40
3	D	501	SFG	C5-C4	2.51	1.46	1.40
3	B	501	SFG	C5-C4	2.55	1.46	1.40
3	A	501	SFG	C5-C4	2.75	1.46	1.40

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	501	SFG	N3-C2-N1	-8.68	122.06	128.87
3	B	501	SFG	N3-C2-N1	-7.67	122.84	128.87
3	D	501	SFG	N3-C2-N1	-7.39	123.07	128.87
3	C	501	SFG	N3-C2-N1	-7.25	123.17	128.87
3	D	501	SFG	C1'-N9-C4	-3.79	122.58	126.81
3	B	501	SFG	C1'-N9-C4	-3.40	123.01	126.81

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	501	SFG	O4'-C4'-C5'	-2.61	105.36	109.22
3	C	501	SFG	C1'-N9-C4	-2.21	124.33	126.81
3	C	501	SFG	C5'-CD-NE	-2.14	102.97	108.96
3	A	501	SFG	C2'-C3'-C4'	2.05	106.82	102.64
3	B	501	SFG	C5'-CD-CG	2.08	116.73	110.43
3	D	501	SFG	N6-C6-N1	2.32	122.41	118.52
3	A	501	SFG	N6-C6-N1	2.70	123.05	118.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	501	SFG	2	0
5	B	502	SO4	1	0
5	C	503	SO4	1	0
5	D	502	SO4	1	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	343/349 (98%)	-0.11	5 (1%) 76 81	24, 34, 50, 65	0
1	B	343/349 (98%)	-0.25	2 (0%) 90 92	23, 31, 45, 71	0
1	C	343/349 (98%)	-0.22	4 (1%) 81 84	22, 32, 44, 62	0
1	D	343/349 (98%)	-0.01	4 (1%) 81 84	21, 35, 54, 68	0
2	F	8/20 (40%)	1.21	2 (25%) 1 0	54, 63, 74, 79	0
2	G	9/20 (45%)	0.06	0 100 100	33, 40, 49, 56	0
2	H	8/20 (40%)	1.07	1 (12%) 5 5	47, 57, 61, 68	0
2	I	9/20 (45%)	-0.36	0 100 100	33, 36, 48, 49	0
All	All	1406/1476 (95%)	-0.13	18 (1%) 79 83	21, 33, 51, 79	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	283	GLY	4.2
1	A	143	GLU	4.1
1	D	476	TYR	3.9
1	A	139	GLU	3.4
1	B	476	TYR	2.9
2	H	3	ALA	2.8
1	C	139	GLU	2.7
1	C	346	ARG	2.7
1	D	475	ARG	2.7
1	D	286	PHE	2.7
1	A	346	ARG	2.6
2	F	6	LYS	2.4
1	C	261	TYR	2.4
1	C	377	HIS	2.3
1	B	477	THR	2.3
2	F	3	ALA	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	445	ARG	2.2
1	A	261	TYR	2.2

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

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	SO4	C	503	5/5	0.92	0.24	3.80	56,58,59,59	0
5	SO4	A	503	5/5	0.92	0.26	1.36	53,55,58,60	0
5	SO4	B	502	5/5	0.94	0.15	0.85	46,50,54,57	0
3	SFG	D	501	27/27	0.94	0.13	0.42	26,30,33,34	0
3	SFG	B	501	27/27	0.94	0.11	0.14	21,24,26,27	0
3	SFG	C	501	27/27	0.94	0.11	-0.20	25,29,33,34	0
3	SFG	A	501	27/27	0.94	0.11	-0.24	28,31,34,34	0
5	SO4	D	502	5/5	0.97	0.11	-0.54	46,46,52,52	0
4	GOL	C	502	6/6	0.93	0.14	-	52,54,54,58	0
4	GOL	A	502	6/6	0.92	0.10	-	54,55,56,58	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.