

Full wwPDB X-ray Structure Validation Report (i)

Nov 12, 2024 – 08:28 PM EST

PDB ID : 4FZV

Title : Crystal structure of the human MTERF4:NSUN4:SAM ternary complex Authors : Guja, K.E.; Yakubovskaya, E.; Mejia, E.; Castano, S.; Hambardjieva, E.; Choi,

W.S.; Garcia-Diaz, M.

Deposited on : 2012-07-08

Resolution : 2.00 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.orgA user guide is available at

https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

MolProbity: 4.02b-467

Mogul : 2022.3.0, CSD as543be (2022)

Xtriage (Phenix) : 1.20.1

EDS : 3.0

buster-report : 1.1.7 (2018)

Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)

CCP4 : 9.0.003 (Gargrove)

Density-Fitness : 1.0.11

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

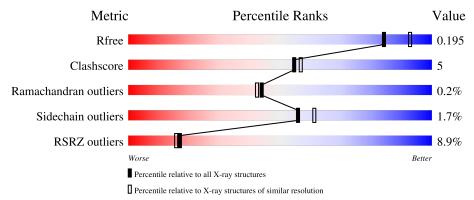
Validation Pipeline (wwPDB-VP) : 2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	Similar resolution
Metric	$(\# ext{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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 >=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 <=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	359	89%	6%	5%
2	В	239	16% 82%	14%	



2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 4635 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.

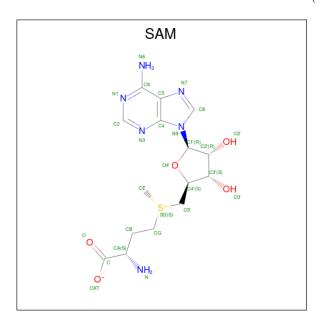
• Molecule 1 is a protein called Putative methyltransferase NSUN4.

Mol	Chain	Residues		Atoms					ZeroOcc	AltConf	Trace
1	A	342	Total 2709	C 1729	N 463	O 498	S 11	Se 8	0	12	0

• Molecule 2 is a protein called mTERF domain-containing protein 2.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
2	B	233	Total	С	N	О	S	0	0	0
		255	1591	1003	282	298	8	0	0	

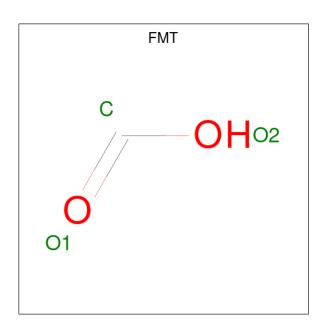
• Molecule 3 is S-ADENOSYLMETHIONINE (three-letter code: SAM) (formula: $C_{15}H_{22}N_6O_5S$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
9	Λ	1	Total	С	N	О	S	0	0
3	A	1	27	15	6	5	1		U

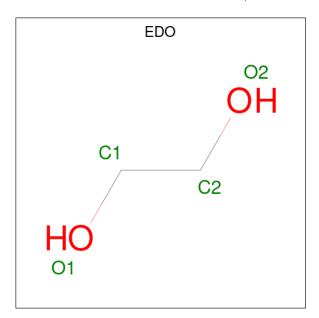
• Molecule 4 is FORMIC ACID (three-letter code: FMT) (formula: CH₂O₂).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C 1	O 2	0	0

 \bullet Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $\mathrm{C_2H_6O_2}).$



Mol	Chain	Residues	Aton	ns	ZeroOcc	AltConf
5	A	1	Total (C O 2	0	0

• Molecule 6 is water.



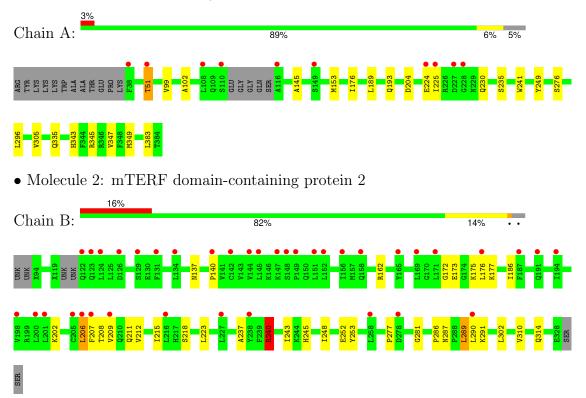
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	275	Total O 275 275	0	0
6	В	26	Total O 26 26	0	0



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: Putative methyltransferase NSUN4





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants	298.67Å 53.23Å 52.96Å	D
a, b, c, α , β , γ	90.00° 98.75° 90.00°	Depositor
Resolution (Å)	46.81 - 2.00	Depositor
Resolution (A)	46.81 - 2.00	EDS
% Data completeness	99.8 (46.81-2.00)	Depositor
(in resolution range)	95.5 (46.81-2.00)	EDS
R_{merge}	(Not available)	Depositor
R_{sum}	(Not available)	Depositor
$< I/\sigma(I) > 1$	1.97 (at 2.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
υ .	0.165 , 0.190	Depositor
R, R_{free}	0.169 , 0.195	DCC
R_{free} test set	2841 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	37.7	Xtriage
Anisotropy	0.297	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.33 , 74.4	EDS
L-test for twinning ²	$< L > = 0.48, < L^2> = 0.31$	Xtriage
	0.006 for -h-k-l,l,k	
Estimated twinning fraction	0.032 for -h+k-l,-l,-k	Xtriage
	0.024 for -h-2*l,-k,l	
F_o, F_c correlation	0.96	EDS
Total number of atoms	4635	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.71% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical values of <|L|>, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



¹Intensities estimated from amplitudes.

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: FMT, SAM, EDO

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		nd lengths	Bond angles		
MIOI		RMSZ	# Z > 5	RMSZ	# Z > 5	
1	A	0.59	$1/2799 \ (0.0\%)$	0.64	0/3796	
2	В	0.30	0/1486	0.57	1/2026 (0.0%)	
All	All	0.51	$1/4285 \ (0.0\%)$	0.62	1/5822 (0.0%)	

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\text{\AA})$	$\operatorname{Ideal}(ext{\AA})$
1	A	99	VAL	CB-CG1	5.15	1.63	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
2	В	240	ARG	NE-CZ-NH1	5.37	122.98	120.30

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	2709	0	2674	13	0
2	В	1591	0	1293	26	0
3	A	27	0	22	0	0
4	A	3	0	1	0	0

Continued on next page...



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	4	0	6	0	0
6	A	275	0	0	0	0
6	В	26	0	0	0	0
All	All	4635	0	3996	39	0

The all-atom clash score is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clash score for this structure is 5.

All (39) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic	Clash
		distance (Å)	overlap (Å)
2:B:240:ARG:HH11	2:B:240:ARG:HG2	1.41	0.85
2:B:137:ASN:HA	2:B:140:PRO:HG2	1.67	0.74
2:B:289:LEU:HD23	2:B:291:LYS:HE3	1.77	0.66
2:B:240:ARG:HG2	2:B:240:ARG:NH1	2.09	0.63
1:A:102:ALA:HB1	1:A:347[A]:VAL:HG12	1.84	0.59
1:A:225:ILE:HG23	1:A:230:GLN:OE1	2.03	0.58
2:B:237:ALA:HB2	2:B:253:TYR:CZ	2.41	0.55
1:A:345:ARG:HH11	1:A:349[A]:MSE:HE3	1.74	0.53
2:B:175:LYS:HA	2:B:177:LYS:N	2.25	0.52
2:B:186:ILE:HD13	2:B:223:LEU:HD13	1.92	0.52
2:B:237:ALA:HB2	2:B:253:TYR:CE2	2.45	0.52
2:B:277:PRO:HB2	2:B:281:GLY:HA2	1.92	0.52
1:A:204:ASP:O	1:A:235[B]:SER:HA	2.10	0.51
1:A:204:ASP:O	1:A:235[A]:SER:HA	2.10	0.50
2:B:172:GLY:HA2	2:B:173:GLU:C	2.30	0.50
2:B:290:LEU:H	2:B:290:LEU:HD12	1.78	0.49
2:B:175:LYS:N	2:B:176:LEU:HB3	2.28	0.48
2:B:207:PHE:CD2	2:B:245:HIS:HD2	2.32	0.47
2:B:286:PRO:O	2:B:287:ASN:HB2	2.14	0.47
1:A:176[A]:ILE:HG22	1:A:249:TYR:CD1	2.51	0.46
1:A:189:LEU:O	1:A:193[A]:GLN:HG2	2.16	0.45
2:B:310:VAL:O	2:B:314:GLN:HG2	2.16	0.45
2:B:207:PHE:CD2	2:B:245:HIS:CD2	3.05	0.45
2:B:218:SER:HB2	2:B:252:GLU:HG2	1.98	0.44
2:B:302:LEU:HD13	2:B:310:VAL:HA	1.98	0.44
2:B:175:LYS:HA	2:B:176:LEU:C	2.38	0.44
2:B:207:PHE:HB3	2:B:208:THR:H	1.50	0.44
2:B:162:ARG:NH1	2:B:186:ILE:O	2.48	0.44
2:B:208:THR:HG23	2:B:211:GLN:OE1	2.18	0.43
1:A:335:GLN:HG2	1:A:383:LEU:HD12	2.01	0.43
	·	Continu	ed on next page

Continued on next page...



Continued	trom	mmoninonic	maaa
COHABABACA		DIEUIUU	DUIUE
0 0 1000100000			

Atom-1	Atom-2	Interatomic	Clash
Atom-1	Atom-2	${ m distance} ({ m \AA})$	overlap (Å)
2:B:208:THR:O	2:B:212:VAL:HG23	2.19	0.43
2:B:243:ILE:HG22	2:B:248:ILE:HG13	2.01	0.43
1:A:145:ALA:HB3	1:A:153:MSE:HE3	2.02	0.41
1:A:176[B]:ILE:HG22	1:A:249:TYR:CD1	2.55	0.41
1:A:296:LEU:HD23	1:A:305:VAL:HG21	2.03	0.41
1:A:343:HIS:O	1:A:347[B]:VAL:HG13	2.20	0.41
2:B:202:LYS:HA	2:B:206:LEU:O	2.20	0.41
2:B:212:VAL:HA	2:B:215:ILE:HD12	2.03	0.40

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	Perce	ntiles
1	A	351/359 (98%)	336 (96%)	14 (4%)	1 (0%)	37	35
2	В	205/239~(86%)	192 (94%)	13 (6%)	0	100	100
All	All	556/598 (93%)	528 (95%)	27 (5%)	1 (0%)	44	42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	224	GLU

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	296/300 (99%)	293 (99%)	3 (1%)	73 78		
2	В	124/196 (63%)	120 (97%)	4 (3%)	34 35		
All	All	420/496 (85%)	413 (98%)	7 (2%)	56 61		

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51	THR
1	A	241	TRP
1	A	276	SER
2	В	206	LEU
2	В	209	VAL
2	В	240	ARG
2	В	289	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	282	GLN
2	В	217	HIS
2	В	245	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)

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

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

3 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	Dog	Dog	Pos	Pag	Dog	Dog	Ros	Ros	Dog	Dog	Link	Bond lengths			Bond angles		
MIOI	ioi Type Chain	LIIIK		Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2											
4	FMT	A	402	-	2,2,2	1.21	0	1,1,1	0.12	0										
5	EDO	A	403	-	3,3,3	0.61	0	2,2,2	0.38	0										
3	SAM	A	401	-	23,29,29	1.28	3 (13%)	20,42,42	2.38	5 (25%)										

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
5	EDO	A	403	-	-	1/1/1/1	-
3	SAM	A	401	-	-	0/13/33/33	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	Observed(A)	$\operatorname{Ideal}(\text{\AA})$
3	A	401	SAM	C2-N3	3.80	1.38	1.32
3	A	401	SAM	OXT-C	-2.90	1.21	1.30
3	A	401	SAM	C2-N1	2.02	1.37	1.33

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$Ideal(^{o})$
3	A	401	SAM	N3-C2-N1	-8.63	116.96	128.67
3	A	401	SAM	C4'-O4'-C1'	-2.85	107.31	109.92
3	A	401	SAM	C1'-N9-C4	-2.44	122.35	126.64
3	A	401	SAM	CB-CA-N	2.30	116.12	110.12
3	A	401	SAM	O4'-C1'-N9	-2.19	105.84	108.75

There are no chirality outliers.

All (1) torsion outliers are listed below:

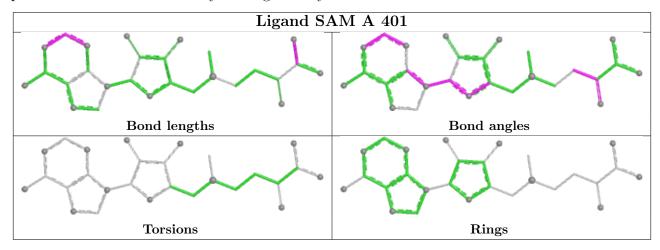


Mol	Chain	Res	Type	Atoms
5	A	403	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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^{th} 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>	# RSRZ > 2	$OWAB(Å^2)$	Q < 0.9
1	A	335/359~(93%)	-0.36	10 (2%) 52 51	10, 32, 67, 97	11 (3%)
2	В	207/239 (86%)	1.01	38 (18%) 4 3	41, 97, 158, 179	0
All	All	542/598 (90%)	0.16	48 (8%) 17 16	10, 45, 143, 179	11 (2%)

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	В	187	PHE	4.4
2	В	216	LEU	4.1
2	В	151	LEU	4.1
1	A	225	ILE	3.9
2	В	176	LEU	3.8
2	В	205	CYS	3.6
1	A	224	GLU	3.5
2	В	124	LEU	3.5
1	A	51	THR	3.4
2	В	131	PHE	3.2
1	A	38	PHE	3.2
1	A	108	LEU	3.1
2	В	126	ASP	3.0
2	В	290	LEU	3.0
2	В	194	ILE	3.0
1	A	228	GLY	2.9
2	В	144	VAL	2.8
1	A	116	ALA	2.8
2	В	169	LEU	2.8
2	В	156	ILE	2.6
2	В	140	PRO	2.6
2	В	209	VAL	2.6
1	A	227	ASP	2.6
2	В	148	SER	2.4

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	В	145	LEU	2.4
2	В	149	PRO	2.4
2	В	191	GLN	2.3
2	В	201	LEU	2.3
2	В	158	GLN	2.3
2	В	152	LEU	2.3
1	A	149	SER	2.3
2	В	123	GLN	2.3
2	В	278	ASP	2.3
2	В	206	LEU	2.3
2	В	171	LEU	2.2
2	В	227	LEU	2.2
2	В	238	TYR	2.2
2	В	134	LEU	2.2
2	В	200	LEU	2.2
2	В	122	GLN	2.1
2	В	198	VAL	2.1
2	В	207	PHE	2.1
1	A	110	SER	2.1
2	В	147	LYS	2.1
2	В	258	LEU	2.0
2	В	129	SER	2.0
2	В	165	TYR	2.0
2	В	142	CYS	2.0

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 monosaccharides in this entry.

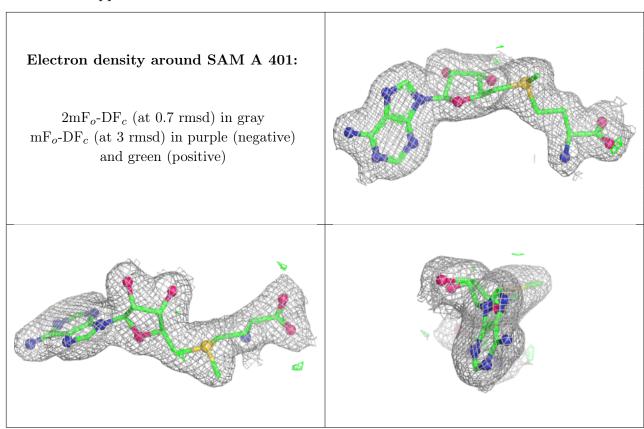
6.4 Ligands (i)

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^{th} 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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
5	EDO	A	403	4/4	0.84	0.15	57,63,68,71	0
4	FMT	A	402	3/3	0.90	0.12	26,26,40,42	0
3	SAM	A	401	27/27	0.98	0.05	19,24,30,32	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



6.5 Other polymers (i)

There are no such residues in this entry.

