wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 08:13 PM GMT

PDB ID	:	4R57
Title	:	Crystal structure of spermidine N-acetyltransferase from Vibrio cholerae in complex with acetyl-CoA
Authors	:	Filippova, E.V.; Minasov, G.; Kiryukhina, O.; Shuvalova, L.; Kuhn, M.L.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on	:	2014-08-20
Resolution	:	2.08 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

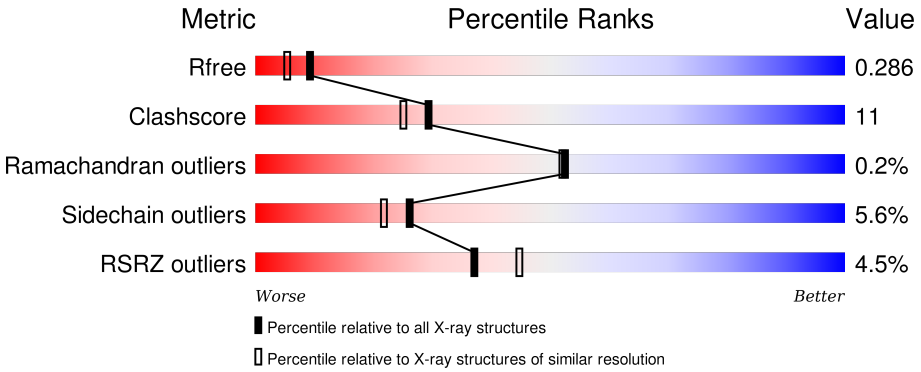
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.08 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	4546 (2.10-2.06)
Clashscore	102246	5101 (2.10-2.06)
Ramachandran outliers	100387	5048 (2.10-2.06)
Sidechain outliers	100360	5049 (2.10-2.06)
RSRZ outliers	91569	4556 (2.10-2.06)

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 >=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 <=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	176	5% 68% 28%
1	B	176	5% 64% 28%
1	C	176	3% 67% 28%
1	D	176	3% 77% 18%
1	E	176	3% 64% 30%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	176	2% 73% 21%
1	G	176	5% 71% 23%
1	H	176	5% 66% 30%
1	I	176	5% 72% 24%
1	J	176	9% 69% 23%
1	K	176	3% 73% 22%
1	L	176	6% 68% 27%

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	ACO	A	201	-	-	-	X
2	ACO	B	201	-	-	-	X
2	ACO	C	201	-	-	-	X
2	ACO	E	201	-	-	-	X
2	ACO	K	201	-	-	-	X
4	PEG	H	202	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 18214 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 Spermidine n1-acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	170	Total	C	N	O	S	0	2	0
			1457	931	258	265	3			
1	B	169	Total	C	N	O	S	0	2	0
			1448	926	254	265	3			
1	C	170	Total	C	N	O	S	0	1	0
			1447	925	255	264	3			
1	D	169	Total	C	N	O	S	0	0	0
			1428	915	249	261	3			
1	E	170	Total	C	N	O	S	0	1	0
			1447	925	255	264	3			
1	F	169	Total	C	N	O	S	0	2	0
			1449	930	252	264	3			
1	G	170	Total	C	N	O	S	0	1	0
			1448	926	255	264	3			
1	H	169	Total	C	N	O	S	0	0	0
			1428	915	249	261	3			
1	I	169	Total	C	N	O	S	0	0	0
			1428	915	249	261	3			
1	J	169	Total	C	N	O	S	0	1	0
			1436	919	251	263	3			
1	K	170	Total	C	N	O	S	0	0	0
			1439	921	253	262	3			
1	L	169	Total	C	N	O	S	0	1	0
			1437	921	251	262	3			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
A	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
A	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
B	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
B	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03

Continued on next page...

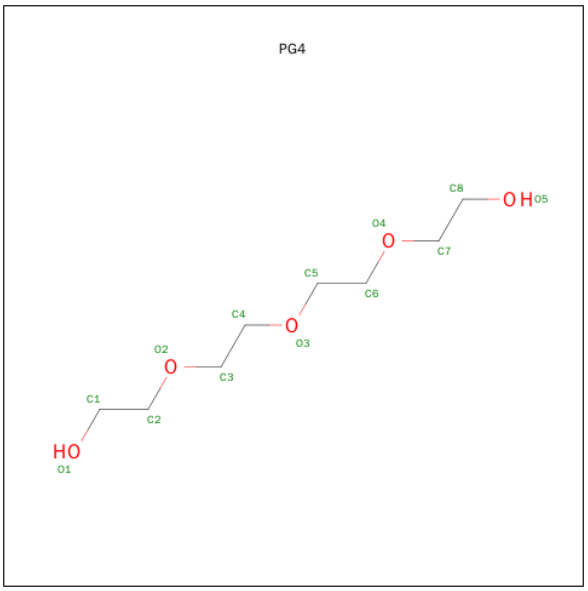
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
C	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
C	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
C	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
D	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
D	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
D	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
E	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
E	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
E	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
F	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
F	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
F	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
G	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
G	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
G	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
H	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
H	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
H	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
I	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
I	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
I	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
J	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
J	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
J	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
K	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
K	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
K	0	ALA	-	EXPRESSION TAG	UNP Q9KL03
L	-2	SER	-	EXPRESSION TAG	UNP Q9KL03
L	-1	ASN	-	EXPRESSION TAG	UNP Q9KL03
L	0	ALA	-	EXPRESSION TAG	UNP Q9KL03

- Molecule 2 is ACETYL COENZYME *A (three-letter code: ACO) (formula: $C_{23}H_{38}N_7O_{17}P_3S$).

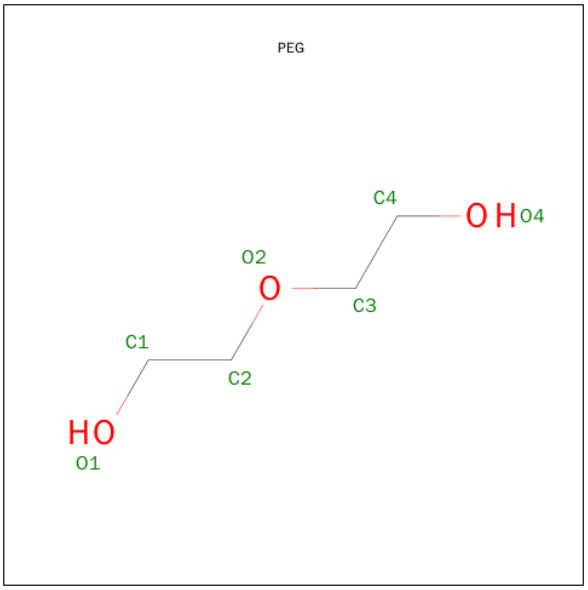
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	B	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	C	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	D	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	E	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	F	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	G	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	H	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	I	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	J	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	K	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
2	L	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		

- Molecule 3 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C O 10 6 4	0	0

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	1	Total C O 7 4 3	0	0
4	G	1	Total C O 7 4 3	0	0
4	H	1	Total C O 7 4 3	0	0

- Molecule 5 is water.

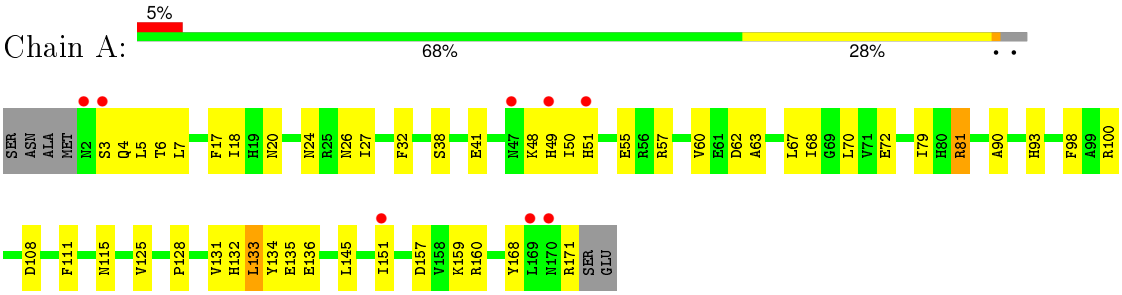
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	27	Total O 27 27	0	0
5	B	29	Total O 30 30	0	1
5	C	25	Total O 25 25	0	0
5	D	20	Total O 20 20	0	0
5	E	36	Total O 36 36	0	0
5	F	29	Total O 29 29	0	0
5	G	23	Total O 23 23	0	0
5	H	23	Total O 23 23	0	0
5	I	14	Total O 14 14	0	0
5	J	26	Total O 26 26	0	0
5	K	11	Total O 11 11	0	0
5	L	15	Total O 15 15	0	0

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.

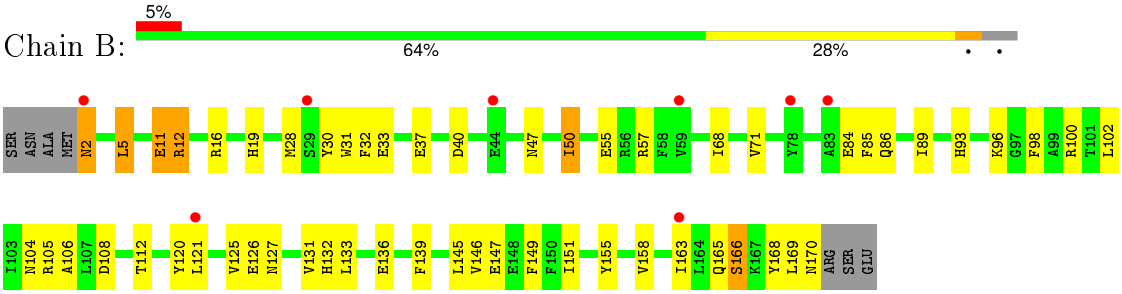
- Molecule 1: Spermidine n1-acetyltransferase

Chain A:



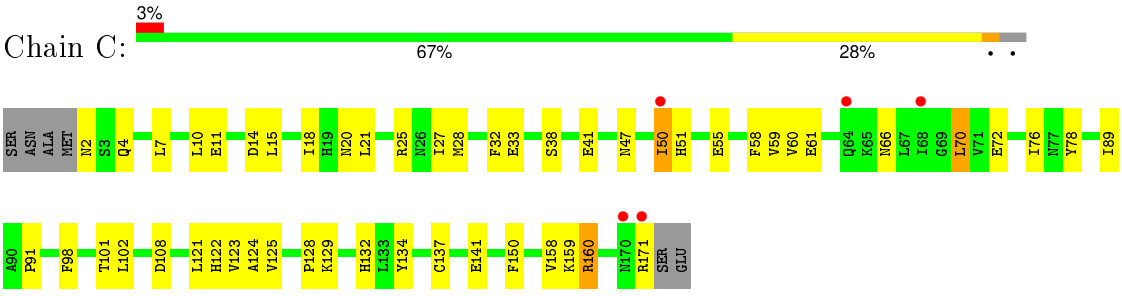
- Molecule 1: Spermidine n1-acetyltransferase

Chain B:



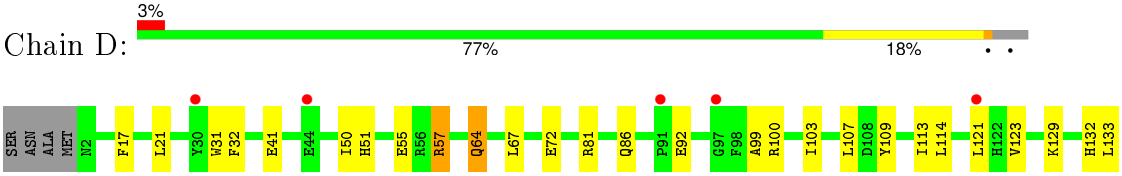
- Molecule 1: Spermidine n1-acetyltransferase

Chain C:



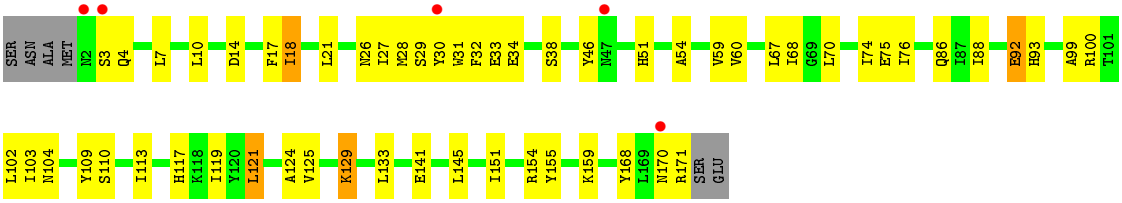
- Molecule 1: Spermidine n1-acetyltransferase

Chain D:



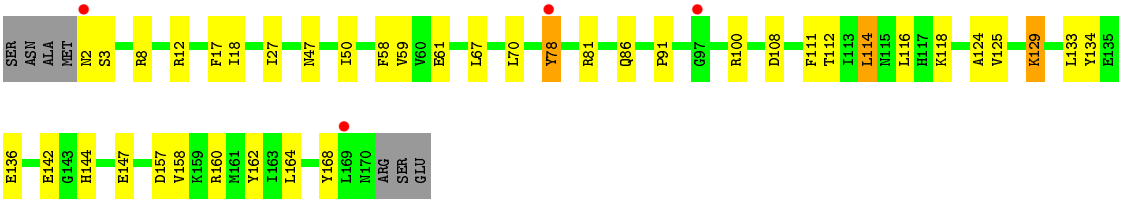
• Molecule 1: Spermidine n1-acetyltransferase

Chain E:



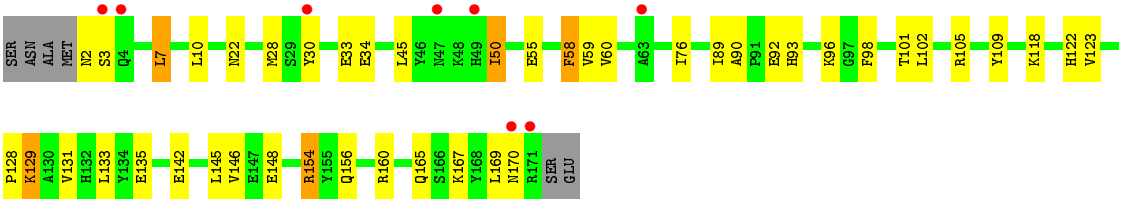
• Molecule 1: Spermidine n1-acetyltransferase

Chain F:



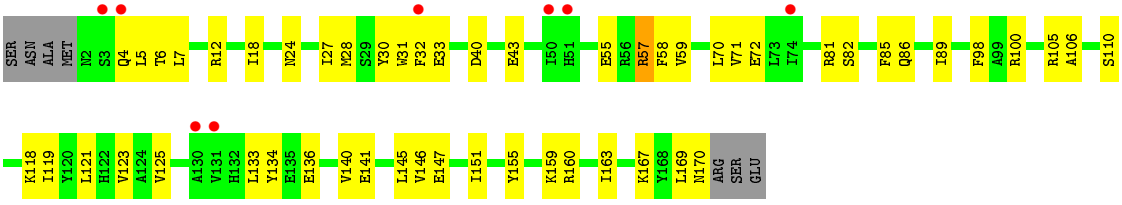
• Molecule 1: Spermidine n1-acetyltransferase

Chain G:



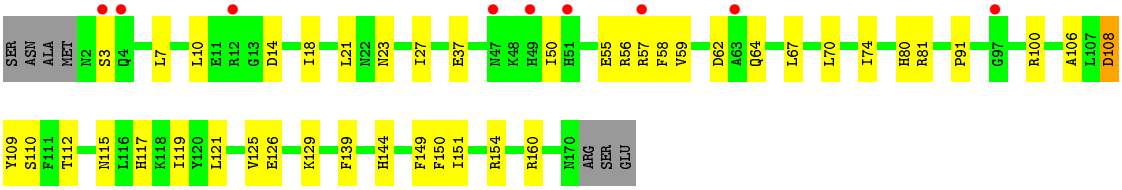
• Molecule 1: Spermidine n1-acetyltransferase

Chain H:

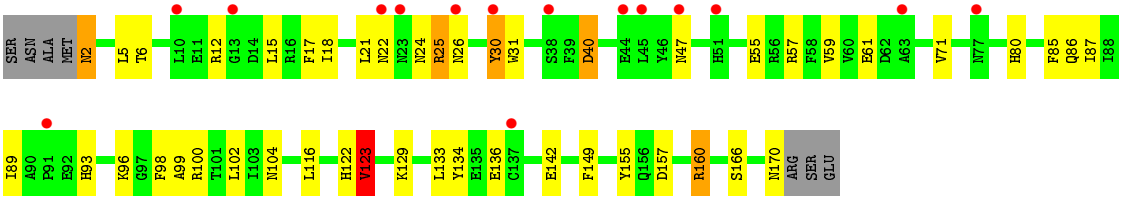


• Molecule 1: Spermidine n1-acetyltransferase

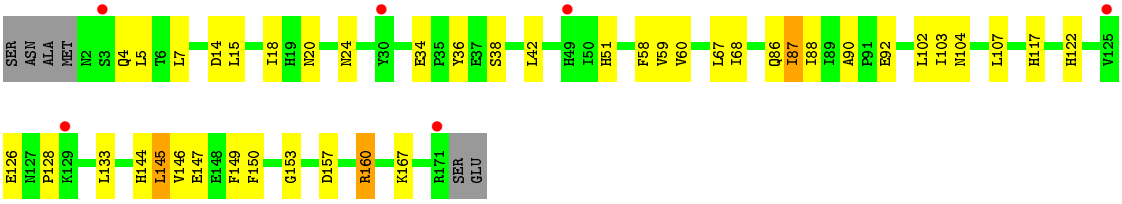
Chain I:



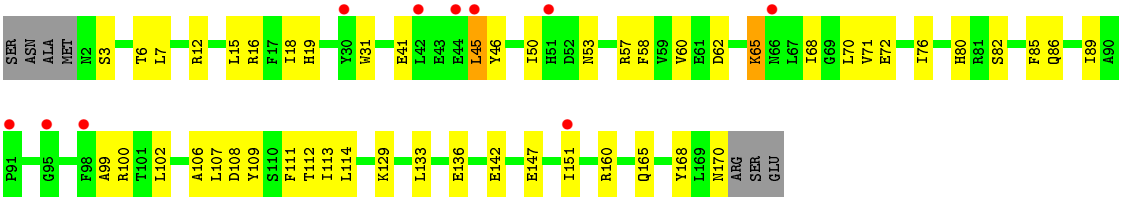
• Molecule 1: Spermidine n1-acetyltransferase



• Molecule 1: Spermidine n1-acetyltransferase



• Molecule 1: Spermidine n1-acetyltransferase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α, β, γ	176.73Å 176.73Å 67.05Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.60 – 2.08 40.60 – 2.08	Depositor EDS
% Data completeness (in resolution range)	99.4 (40.60-2.08) 99.7 (40.60-2.08)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.52 (at 2.08Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.171 , 0.240 0.222 , 0.286	Depositor DCC
R_{free} test set	7225 reflections (5.42%)	DCC
Wilson B-factor (Å^2)	33.8	Xtriage
Anisotropy	0.024	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 38.4	EDS
Estimated twinning fraction	0.224 for H, K, L 0.274 for -K, -H, -L 0.292 for K, H, -L 0.209 for -H, -K, L 0.020 for -h,-k,l 0.116 for h,-h-k,-l 0.000 for -k,-h,-l	Xtriage
Reported twinning fraction	0.224 for H, K, L 0.274 for -K, -H, -L 0.292 for K, H, -L 0.209 for -H, -K, L	Depositor
L-test for twinning[2]	$< \|L\| > = 0.54$, $< L^2 > = 0.38$	Xtriage
Outliers	0 of 140479 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	18214	wwPDB-VP
Average B, all atoms (Å^2)	37.0	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 29.88 % 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 1.4383e-03.The detected translational NCS is most likely*

[1] Intensities estimated from amplitudes.

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

also responsible for the elevated intensity ratio.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PG4, PEG, ACO

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	0.72	0/1491	0.89	4/2013 (0.2%)
1	B	0.71	0/1481	0.82	0/1999
1	C	0.68	0/1480	0.83	0/1998
1	D	0.68	0/1461	0.87	0/1973
1	E	0.80	1/1480 (0.1%)	0.92	0/1998
1	F	0.73	0/1483	0.88	1/2002 (0.0%)
1	G	0.78	1/1481 (0.1%)	0.92	1/1999 (0.1%)
1	H	0.72	0/1461	0.94	3/1973 (0.2%)
1	I	0.62	0/1461	0.82	0/1973
1	J	0.64	0/1469	0.82	2/1984 (0.1%)
1	K	0.64	0/1472	0.82	1/1987 (0.1%)
1	L	0.62	0/1470	0.83	0/1984
All	All	0.70	2/17690 (0.0%)	0.86	12/23883 (0.1%)

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

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	38	SER	CB-OG	-5.24	1.35	1.42
1	G	148	GLU	CD-OE1	-5.08	1.20	1.25

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	157	ASP	CB-CG-OD1	7.51	125.06	118.30
1	A	160	ARG	NE-CZ-NH2	5.89	123.24	120.30
1	H	57	ARG	NE-CZ-NH2	-5.63	117.49	120.30
1	F	160	ARG	NE-CZ-NH2	5.55	123.08	120.30
1	J	116	LEU	CA-CB-CG	-5.50	102.65	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	32	PHE	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	1457	0	1408	35	0
1	B	1448	0	1401	47	0
1	C	1447	0	1402	34	0
1	D	1428	0	1384	26	0
1	E	1447	0	1402	46	0
1	F	1449	0	1404	28	0
1	G	1448	0	1404	30	0
1	H	1428	0	1384	42	0
1	I	1428	0	1384	30	0
1	J	1436	0	1389	40	0
1	K	1439	0	1397	24	0
1	L	1437	0	1396	38	0
2	A	51	0	34	4	0
2	B	51	0	34	6	0
2	C	51	0	34	6	0
2	D	51	0	34	1	0
2	E	51	0	34	7	0
2	F	51	0	34	2	0
2	G	51	0	34	3	0
2	H	51	0	34	6	0
2	I	51	0	34	1	0
2	J	51	0	34	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	K	51	0	34	1	0
2	L	51	0	34	4	0
3	B	10	0	13	0	0
4	C	7	0	10	0	0
4	G	7	0	10	0	0
4	H	7	0	10	0	0
5	A	27	0	0	1	0
5	B	30	0	0	6	0
5	C	25	0	0	0	0
5	D	20	0	0	2	0
5	E	36	0	0	1	0
5	F	29	0	0	0	0
5	G	23	0	0	1	0
5	H	23	0	0	2	0
5	I	14	0	0	0	0
5	J	26	0	0	2	0
5	K	11	0	0	1	0
5	L	15	0	0	0	0
All	All	18214	0	17206	397	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 397 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:D:140:VAL:HG11	1:E:125:VAL:HG12	1.27	1.07
1:F:129[A]:LYS:HE3	1:F:129[A]:LYS:HA	1.41	1.01
1:F:78[A]:TYR:N	1:F:78[A]:TYR:CD1	2.28	0.96
1:A:26[B]:ASN:ND2	1:A:26[B]:ASN:O	2.02	0.93
1:F:78[A]:TYR:H	1:F:78[A]:TYR:HD1	1.15	0.92

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	170/176 (97%)	160 (94%)	10 (6%)	0	100	100
1	B	169/176 (96%)	163 (96%)	5 (3%)	1 (1%)	30	24
1	C	169/176 (96%)	161 (95%)	7 (4%)	1 (1%)	30	24
1	D	167/176 (95%)	158 (95%)	9 (5%)	0	100	100
1	E	169/176 (96%)	165 (98%)	4 (2%)	0	100	100
1	F	169/176 (96%)	163 (96%)	6 (4%)	0	100	100
1	G	169/176 (96%)	158 (94%)	11 (6%)	0	100	100
1	H	167/176 (95%)	155 (93%)	12 (7%)	0	100	100
1	I	167/176 (95%)	158 (95%)	9 (5%)	0	100	100
1	J	168/176 (96%)	159 (95%)	9 (5%)	0	100	100
1	K	168/176 (96%)	160 (95%)	8 (5%)	0	100	100
1	L	168/176 (96%)	161 (96%)	5 (3%)	2 (1%)	16	9
All	All	2020/2112 (96%)	1921 (95%)	95 (5%)	4 (0%)	52	52

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	136	GLU
1	C	50	ILE
1	L	65	LYS
1	B	50	ILE

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	156/159 (98%)	150 (96%)	6 (4%)	40	39
1	B	155/159 (98%)	145 (94%)	10 (6%)	21	17
1	C	155/159 (98%)	144 (93%)	11 (7%)	18	13

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	153/159 (96%)	146 (95%)	7 (5%)	33	30
1	E	155/159 (98%)	145 (94%)	10 (6%)	21	17
1	F	155/159 (98%)	144 (93%)	11 (7%)	18	13
1	G	155/159 (98%)	143 (92%)	12 (8%)	16	11
1	H	153/159 (96%)	146 (95%)	7 (5%)	33	30
1	I	153/159 (96%)	146 (95%)	7 (5%)	33	30
1	J	154/159 (97%)	145 (94%)	9 (6%)	25	20
1	K	154/159 (97%)	144 (94%)	10 (6%)	21	17
1	L	154/159 (97%)	147 (96%)	7 (4%)	34	31
All	All	1852/1908 (97%)	1745 (94%)	107 (6%)	26	20

5 of 107 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	78[A]	TYR
1	G	58	PHE
1	K	145	LEU
1	F	78[B]	TYR
1	F	133	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	152	ASN
1	G	93	HIS
1	K	104	ASN
1	G	20	ASN
1	B	170	ASN

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)

16 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	ACO	A	201	-	43,53,53	0.93	2 (4%)	55,79,79	1.85	11 (20%)
2	ACO	B	201	-	43,53,53	0.88	2 (4%)	55,79,79	1.78	6 (10%)
3	PG4	B	202	-	9,9,12	0.48	0	8,8,11	0.66	0
2	ACO	C	201	-	43,53,53	0.94	2 (4%)	55,79,79	1.76	6 (10%)
4	PEG	C	202	-	6,6,6	0.51	0	5,5,5	0.52	0
2	ACO	D	201	-	43,53,53	0.96	3 (6%)	55,79,79	1.67	7 (12%)
2	ACO	E	201	-	43,53,53	0.85	2 (4%)	55,79,79	2.08	11 (20%)
2	ACO	F	201	-	43,53,53	0.92	2 (4%)	55,79,79	1.76	8 (14%)
2	ACO	G	201	-	43,53,53	0.90	1 (2%)	55,79,79	1.74	8 (14%)
4	PEG	G	202	-	6,6,6	0.44	0	5,5,5	0.75	0
2	ACO	H	201	-	43,53,53	0.87	1 (2%)	55,79,79	1.86	13 (23%)
4	PEG	H	202	-	6,6,6	0.48	0	5,5,5	0.51	0
2	ACO	I	201	-	43,53,53	0.95	2 (4%)	55,79,79	1.76	7 (12%)
2	ACO	J	201	-	43,53,53	0.91	2 (4%)	55,79,79	1.88	8 (14%)
2	ACO	K	201	-	43,53,53	0.89	1 (2%)	55,79,79	1.63	5 (9%)
2	ACO	L	201	-	43,53,53	0.91	2 (4%)	55,79,79	1.71	8 (14%)

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	ACO	A	201	-	-	2/47/67/67	0/3/3/3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ACO	B	201	-	-	0/47/67/67	0/3/3/3
3	PG4	B	202	-	-	0/7/7/10	0/0/0/0
2	ACO	C	201	-	-	0/47/67/67	0/3/3/3
4	PEG	C	202	-	-	0/4/4/4	0/0/0/0
2	ACO	D	201	-	-	0/47/67/67	0/3/3/3
2	ACO	E	201	-	-	0/47/67/67	0/3/3/3
2	ACO	F	201	-	-	0/47/67/67	0/3/3/3
2	ACO	G	201	-	-	0/47/67/67	0/3/3/3
4	PEG	G	202	-	-	0/4/4/4	0/0/0/0
2	ACO	H	201	-	-	0/47/67/67	0/3/3/3
4	PEG	H	202	-	-	0/4/4/4	0/0/0/0
2	ACO	I	201	-	-	0/47/67/67	0/3/3/3
2	ACO	J	201	-	-	2/47/67/67	0/3/3/3
2	ACO	K	201	-	-	1/47/67/67	0/3/3/3
2	ACO	L	201	-	-	0/47/67/67	0/3/3/3

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	201	ACO	C2A-N3A	2.00	1.35	1.32
2	B	201	ACO	O4B-C1B	2.01	1.43	1.41
2	E	201	ACO	O4B-C1B	2.44	1.44	1.41
2	A	201	ACO	O4B-C1B	2.51	1.44	1.41
2	C	201	ACO	O4B-C1B	2.51	1.44	1.41

The worst 5 of 98 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	201	ACO	N3A-C2A-N1A	-8.70	122.23	128.89
2	L	201	ACO	N3A-C2A-N1A	-8.59	122.32	128.89
2	G	201	ACO	N3A-C2A-N1A	-8.44	122.44	128.89
2	J	201	ACO	N3A-C2A-N1A	-8.35	122.50	128.89
2	C	201	ACO	N3A-C2A-N1A	-8.31	122.53	128.89

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	K	201	ACO	O-C-S1P-C2P
2	A	201	ACO	CH3-C-S1P-C2P
2	A	201	ACO	O-C-S1P-C2P

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	J	201	ACO	CH3-C-S1P-C2P
2	J	201	ACO	O-C-S1P-C2P

There are no ring outliers.

12 monomers are involved in 49 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	201	ACO	4	0
2	B	201	ACO	6	0
2	C	201	ACO	6	0
2	D	201	ACO	1	0
2	E	201	ACO	7	0
2	F	201	ACO	2	0
2	G	201	ACO	3	0
2	H	201	ACO	6	0
2	I	201	ACO	1	0
2	J	201	ACO	8	0
2	K	201	ACO	1	0
2	L	201	ACO	4	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{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>2			OWAB(Å^2)	Q<0.9
1	A	170/176 (96%)	0.49	8 (4%)	35	43	24, 35, 52, 69	0
1	B	169/176 (96%)	0.49	8 (4%)	35	43	24, 36, 51, 64	0
1	C	170/176 (96%)	0.35	5 (2%)	55	63	24, 34, 48, 59	0
1	D	169/176 (96%)	0.46	6 (3%)	46	55	23, 35, 53, 61	0
1	E	170/176 (96%)	0.36	5 (2%)	55	63	22, 31, 46, 67	0
1	F	169/176 (96%)	0.30	4 (2%)	62	67	21, 32, 47, 60	0
1	G	170/176 (96%)	0.38	8 (4%)	35	43	20, 32, 47, 59	0
1	H	169/176 (96%)	0.38	8 (4%)	35	43	20, 32, 46, 61	0
1	I	169/176 (96%)	0.52	9 (5%)	30	37	21, 37, 57, 71	0
1	J	169/176 (96%)	0.66	15 (8%)	12	16	23, 39, 57, 66	0
1	K	170/176 (96%)	0.48	6 (3%)	48	56	26, 38, 55, 79	0
1	L	169/176 (96%)	0.69	10 (5%)	26	33	24, 40, 58, 71	0
All	All	2033/2112 (96%)	0.46	92 (4%)	37	45	20, 35, 54, 79	0

The worst 5 of 92 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	3	SER	5.3
1	I	63	ALA	5.2
1	K	49	HIS	4.5
1	F	78[A]	TYR	4.0
1	L	95	GLY	4.0

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 (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^{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	LLDF	B-factors(Å^2)	Q<0.9
2	ACO	C	201	51/51	0.78	0.27	6.22	39,60,95,96	0
2	ACO	B	201	51/51	0.81	0.28	3.27	34,57,94,95	0
2	ACO	E	201	51/51	0.85	0.24	2.67	31,55,86,90	0
4	PEG	H	202	7/7	0.88	0.23	2.60	34,36,40,40	0
2	ACO	K	201	51/51	0.79	0.28	2.19	34,57,77,83	0
2	ACO	A	201	51/51	0.84	0.21	2.12	34,50,91,94	0
2	ACO	L	201	51/51	0.78	0.28	1.93	37,59,78,79	0
2	ACO	D	201	51/51	0.80	0.24	1.56	41,55,85,88	0
4	PEG	G	202	7/7	0.89	0.20	1.29	35,36,39,41	0
2	ACO	F	201	51/51	0.84	0.22	1.27	31,51,91,97	0
2	ACO	I	201	51/51	0.83	0.21	1.13	33,47,66,68	0
2	ACO	J	201	51/51	0.83	0.19	1.01	44,58,72,74	0
4	PEG	C	202	7/7	0.91	0.16	0.59	24,31,34,35	0
2	ACO	H	201	51/51	0.87	0.16	0.38	29,40,50,53	0
3	PG4	B	202	10/13	0.88	0.17	0.22	37,39,42,48	0
2	ACO	G	201	51/51	0.90	0.14	-0.25	27,37,52,54	0

6.5 Other polymers (i)

There are no such residues in this entry.