Phenylalanyl tRNA synthetase (PheRS) substrate mimics: Design, synthesis, molecular dynamics and antimicrobial evaluation

Nada A. Noureldin^{a,b,*}, Jennifer Richards^c, Hend Kothayer^b, Mohammed M. Baraka^b, Sobhy M. Eladl^b, Mandy Wootton^c, Claire Simons^a

^aSchool of Pharmacy and Pharmaceutical Sciences, Cardiff University, Cardiff CF10 3NB, U.K. ^bDepartment of Medicinal Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig P.C. 44519, Egypt ^cSpecialist Antimicrobial Chemotherapy Unit, University Hospital of Wales, Heath Park, Cardiff CF14 4XW, U.K.

^{*}Correspondence: <u>NANoureddine@pharmacy.zu.edu.eg</u>

¹ School of Pharmacy and Pharmaceutical Sciences, Cardiff University, Cardiff CF10 3NB, United Kingdom

² Department of Medicinal Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig P.C. 44519, Egyp

Electronic supplementary information

1. Microbiological evaluation results:

Compound	S. aureus	E. faecalis	Р.	E. coli	К.
	ATCC	ATCC	aeruginosa	ATCC	pneumoniae
	29213	29212	ATCC	25922	ATCC
			29853		700603
5a	>128	128	128	>128	>128
5b	>128	128	128	>128	128
6a	128	128	128	>128	128
6b	128	128	128	>128	128
6c	128	128	128	>128	128
6d	128	128	128	>128	128
6e	128	128	128	128	128
7a	128	128	128	128	128
7b	128	128	128	>128	128
7c	>128	128	128	>128	128
7d	128	128	128	128	128
9a	128	128	128	>128	128
9b	128	64	128	>128	128
9c	128	128	128	>128	>128
9d	128	64	128	>128	128
9e	64	32	128	>128	128
ciprofloxacin	0.25	1	0.5	0.008	0.25

Table. S1| MIC (µg/mL) for the synthesised compounds (5a-b), (6a-e), (7a-d) and (9a-e)

Compound	S. aureus	MRSA	E. faecalis	E. faecium
	ATCC 29213	NCTC 12493	ATCC 29212	16568
16a	128	64	64	32
16b	128	64	128	64
16c	128	64	32	32
ciprofloxacin	0.25	0.5	0.5	1

Table. S2| MIC (µg/mL) for the synthesised compounds (16a-c)

2. Computational studies supplementary information:

2.1. Flexible alignment studies:

Flexible alignment and docking studies were performed using MOE 2015.10 software¹. Flexible alignment was performed using MMFF94 forcefield, flexible alignment mode and the resulted conformations were examined according to their grand alignment score (S). The latter is the sum of the similiraity measure of configuration (F) and the average strain energy of the molecules in the alignment in kcal/mol (U). The lower S value indicates better alignment.

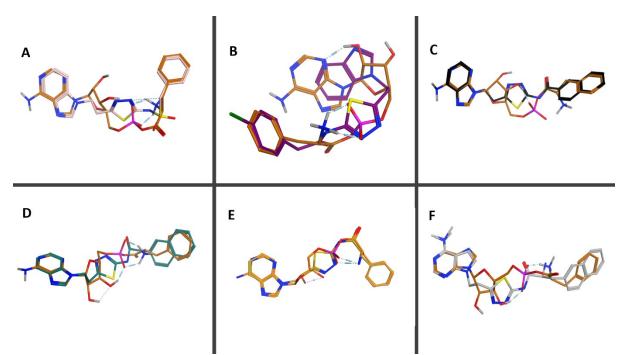


Fig. S1. 3D molecular alignment results. Phe-AMP is presented in orange colour. (A) **5a** (pink colour) S = -170.1604, (B) **6c** (purple colour) S = -154.4615, (C) **6e** (black colour) S = -159.4926, (D) **7d** (green colour) S = -162.6272, (E) **9a** (yellow colour) S = -190.7823 and (F) **16c** (grey colour) S = -206.69

2.2. Multiple sequence alignment:

CLUSTAL 0(1.2.4) multiple sequence alignment

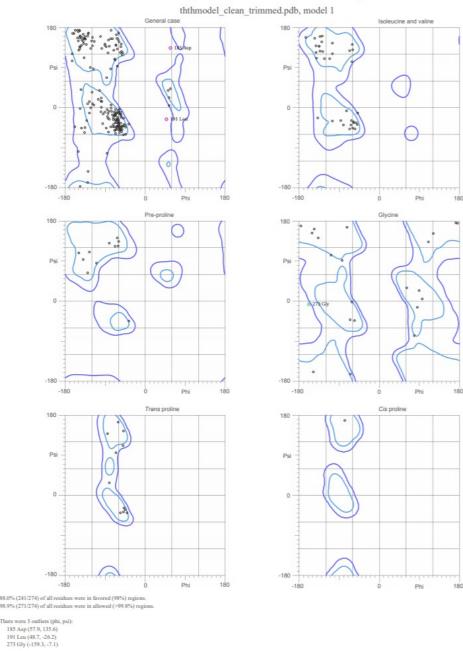
sp Q9I0A3 SYFA_PSEAE sp A6TAI3 SYFA_KLEP7 sp P08312 SYFA_ECOLI sp P68849 SYFA_STAAU sp Q836J6 SYFA_ENTFA	MENLDALVSQALEAVRHTEDVNALEQIRVHYLGKKGELTQVMKTLGDLPAEERPKVG MSHLAELVASAKAAINEASDVAALDNVRVEYLGKKGHLTLQMTTLRELPPEERPAAG MSHLAELVASAKAAISQASDVAALDNVRVEYLGKKGHLTLQMTTLRELPPEERPAAG MSEQQTMSELKQQALVDINEANDERALQEVKVKYLGKKGSVSGLMKLMKDLPNEEKPAFG MTLQAQLEALRDNTLKEIAQVATLKELNQIRVETLGKKGPITEVLRGMKNLSPEERPVVG : * .: : *::::*. ***** :: : : : * **:* *	57 57 57 60 60
sp Q9I0A3 SYFA_PSEAE sp A6TAI3 SYFA_KLEP7 sp P08312 SYFA_ECOLI sp P68849 SYFA_STAAU sp Q836J6 SYFA_ENTFA	ALINVAKEKVQDVLNARKTELEGAALAARLAAERIDVTLPGRGQLSGGLHPVTRTLERIE AVINEAKEQVQQALNARKAELEGAALNARLAAETIDVSLPGRRIENGGLHPVTRTIDRIE AVINEAKEQVQQALNARKAELESAALNARLAAETIDVSLPGRRIENGGLHPVTRTIDRIE QKVNELRQTIQNELDERQQMLVKEKLNKQLAEETIDVSLPGRHIEIGSKHPLTRTIEEIE GFANEIRDLLTEAIEARKVVLENEALNAALKEESLDVTLPGKQMPQGTRHILTQVMEEIE * :: : : : * * * * * * * * : **:***: * * : * * : *	117 117 117 120 120
sp Q9I0A3 SYFA_PSEAE sp A6TAI3 SYFA_KLEP7 sp P08312 SYFA_ECOLI sp P68849 SYFA_STAAU sp Q836J6 SYFA_ENTFA	QCFSRIGYEVAEGPEVEDDYHNFEALNIPGHHPARAMHDTFYFNANMLLRTHTSPVQVRT SFFGELGFTVATGPEIEDDYHNFDALNIPGHHPARADHDTFWFDATRLLRTQTSGVQIRT SFFGELGFTVATGPEIEDDYHNFDALNIPGHHPARADHDTFWFDTTRLLRTQTSGVQIRT DLFLGLGYEIVNGYEVEQDHYNFEMLNLPKSHPARDMQDSFYITDEILLRTHTSPVQART NF ::::::::::::::::::::::::::::::::::::	177 177 177 180 180
sp Q9I0A3 SYFA_PSEAE sp A6TAI3 SYFA_KLEP7 sp P08312 SYFA_ECOLI sp P68849 SYFA_STAAU sp Q836J6 SYFA_ENTFA	MESQQPPIRIVCPGRVYRCDSD-LTHSPMFHQVEGLLVDEGVSFADLKGTIEEFLRA MENQQPPIRIIAPGRVYRNDYD-QTHTPMFHQMEGLIVDKNISFTNLKGTLHDFLNN MKAQQPPIRIIAPGRVYRNDYD-QTHTPMFHQMEGLIVDTNISFTNLKGTLHDFLRN MESR-HGQGPVKIICPGKVYRRDSDDATHSHQFTQIEGLVVDKNVKMSDLKGTLELLAKK MEKHDFSKGALRMISPGKVFRRDTDDATHSHQFHQIEGLVVDKNVTMGDLKGTLEVMMKK *::::::::::::::::::::::::::::::::::	233 233 233 239 240
sp Q910A3 SYFA_PSEAE sp A6TA13 SYFA_KLEP7 sp P08312 SYFA_ECOLI sp P68849 SYFA_STAAU sp Q836J6 SYFA_ENTFA sp Q910A3 SYFA_ENTFA sp A6TA13 SYFA_KLEP7 sp P08312 SYFA_ECOLI sp P68849 SYFA_STAAU sp Q836J6 SYFA_ENTFA	FFEKQLEVRERPSFEPFTEPSAEVDIQCVICSGNGCRVCKQTGWLEVMGCGMVHPNVLRMFFEEDLQVRERPSYEPFTEPSAEVDVMGKNGKWLEVLGCGMVHPNVLRNFFEEDLQIRERPSYEPFTEPSAEVDVMGKNGKWLEVLGCGMVHPNVLRNLFGADREIRLRPSYEPFTEPSVEVDVSCFKCKGKGCNVCKHTGWIEILGAGMVHPNVLEMMFGEDRKIRLRPSYEPFTEPSVEVDVSCFKCGGAGCNVCKHTGWIEILGAGMVHPDVLQM:* :::::::::::::::::::::::::::::::::::	293 282 282 299 300

Fig. S2. Multiple sequence alignment² of PheRS α subunit amino acid sequences from *P. aeruginosa* (Q9I0A3), *K. pneumoniae* (A6TAI3), *E. coli* (P08312), *S. aureus* (P68849) and *E. faecalis* (Q836J6). Yellow boxes show residues responsible for hydrophobic interactions. While green boxes show residues responsible for hydrophilic interaction. A high level of conservation among sequences could be observed.

2.3. Homology model evaluation:

2.3.1. Ramachandran plot:

Almost 98% of the amino acids are in the allowed region with only 3 amino acids observed in the outlier region which are Asp 185, Leu 191 and Gly 273 and all of them are not a part of the enzyme's active site.

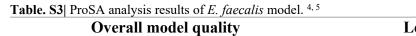


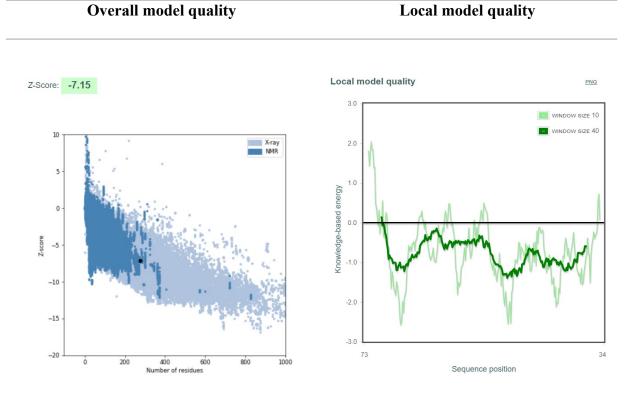
MolProbity Ramachandran analysis

Fig. S3| Ramachandran plot of E. faecalis model.³

2.3.2. ProSA analysis:

ProSA analysis revealed good quality model as its Z-score is among the same sized solved protein structures and equal to -7.15 and the internal energy of each amino acid in the protein 3D structure is with negative value (Table. S3).





2.3.3. Verify 3D:

The newly constructed model passed the 3D verification as 86.32% of its amino acids have scored more than or equal to 0.2.

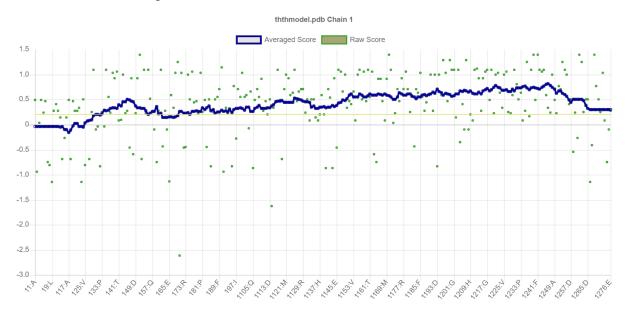


Fig. S4| 3D profile of *E. faecalis* model. ⁶

3. Chemistry supplementary information:

3.1. Preparation methods for compounds 3,4 and 12:

3.1.1. 3-(1H-Benzo[d]imidazol-1-yl) propanehydrazide (3) ^{7,8}

To a solution of the methyl propanoate ester (2) (3 g, 14.69 mmol) in EtOH (125 mL) was added hydrazine monohydrate (10 equivalents) and the reaction mixture was stirred at room temperature overnight. The formed solid was collected by filtration and recrystallised from EtOAc and petroleum ether to give the pure hydrazide compound (3) as white crystals; yield 2.5 g (84%); m.p. 105-110 °C; TLC CH₂Cl₂-MeOH 9:1 v/v, $R_f = 0.29$; ¹H NMR (DMSO-*d*₆) δ : 2.60 (t, *J* = 6.7 Hz, 2H, CH₂CO), 4.19 (s, 2H, NH₂), 4.47 (t, *J* = 6.6 Hz, 2H, CH₂CH₂), 7.21 (t, *J* = 7.2 Hz, 1H, Ar), 7.26 (t, *J* = 8.1 Hz, 1H, Ar), 7.58 (d, *J* = 8.00 Hz, 1H, Ar), 7.62 (d, *J* = 7.9 Hz, 1H, Ar), 8.11 (s, 1H, H-imidazole), 9.06 (s, 1H, NH). ¹³C NMR (DMSO-*d*₆) δ : 34.1 (CH₂CO), 40.9 (NCH₂CH₂), 110.9 (CH), 119.8 (CH), 122.1 (CH), 122.9 (CH), 134.0 (C), 143.6 (C), 144.4 (CH-imidazole), 169.5 (CO).

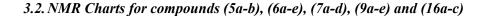
3.1.2. 5-(2-(1H-Benzo[d]imidazol-1-yl)ethyl)-1,3,4-thiadiazol-2-amine (4)⁷

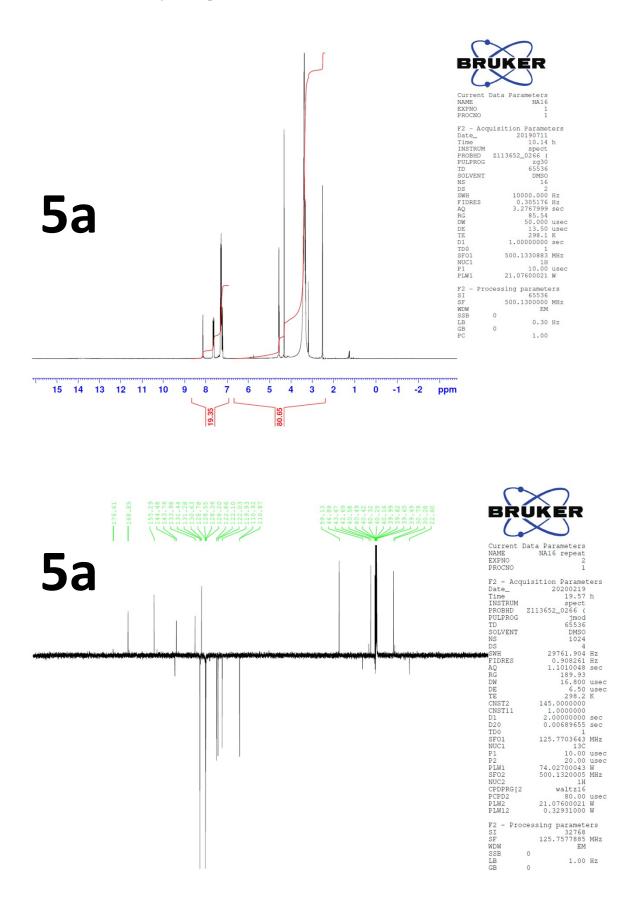
To a stirred mixture of the hydrazide (0.5 g, 2.45 mmol) in dry IPA (20 mL), was added (trimethylsilyl)isothiocyanate (TMSNCS) (1.38mL, 9.8 mmol), then the mixture was heated under reflux overnight. The solvent was then concentrated under vacuum and c.H₂SO₄ (10 mL) was added, and the reaction stirred at room temperature for 2 h. The reaction was poured into crushed ice and neutralised using NH₄OH in -70°C dry ice/acetone bath. The resulting solid was collected by filtration, washed with H₂O and petroleum ether and dried *in vacuo* at 40 °C to afford the product as a yellow solid. Yield 0.3 g (53%); m.p. 208- 210°C (Lit. m.p. 182-184²); TLC CH₂Cl₂-MeOH 9:1 v/v, R_f 0.49. ¹H NMR (DMSO-*d*₆) δ : CH₂ signal is obscured by DMSO-*d*₆ signal, 4.59 (t, *J* = 6.8 Hz, 2H, CH₂CH₂), 7.03 (s, 2H, NH₂), 7.21 (t, *J* = 7.3 Hz, 1H, Ar), 7.26 (t, *J* = 7.6 Hz, 1H, Ar), 7.63 (t, *J* = 8.6 Hz, 2H, Ar), 8.15 (s, 1H, H-imidazole). ¹³C NMR (DMSO-*d*₆) δ : 30.3 (NCH₂<u>C</u>H₂), 43.7 (N<u>C</u>H₂), 110.9 (CH), 119.9 (CH), 122.0 (CH), 122.8 (CH), 134.1 (C), 143.9 (C), 144.6 (CH-imidazole), 154.7 (C), 169.1 (C).

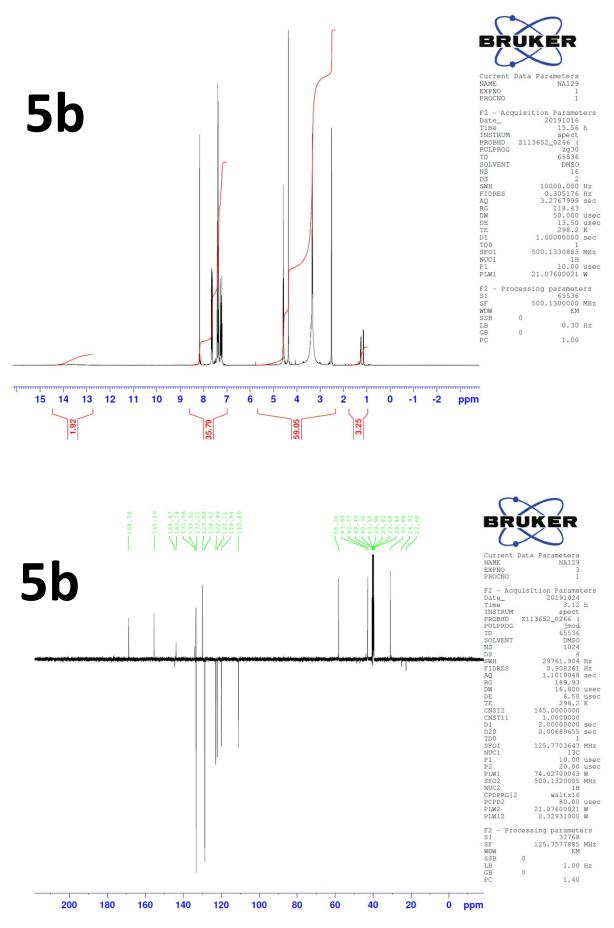
3.1.3. Methyl 3-(4-(dimethylamino)-7H-pyrrolo[2,3-d]pyrimidin-7-yl)propanoate (12) 9

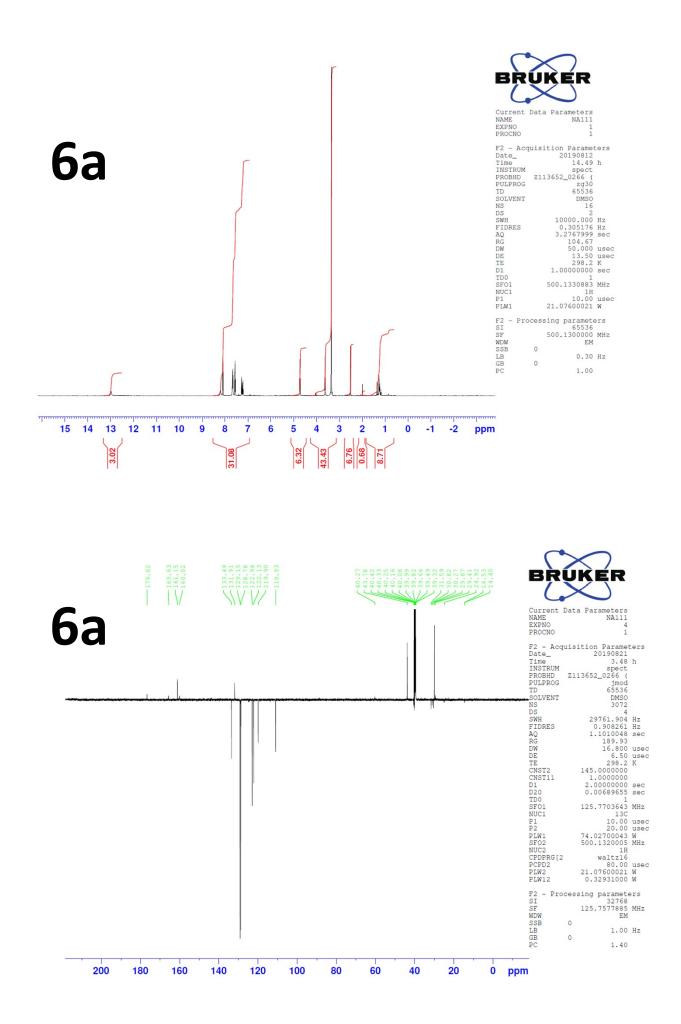
Prepared using *N*,*N*-dimethyl-7*H*-pyrrolo[2,3-*d*]pyrimidin-4-amine (11) (2 g, 12.33 mmol). The product was pure enough to be used in next steps with any further purification needed and separated as yellow low melting solid. Yield: 2.54 g, 83%; m.p. 52-54 °C; TLC: CH₂Cl₂-MeOH

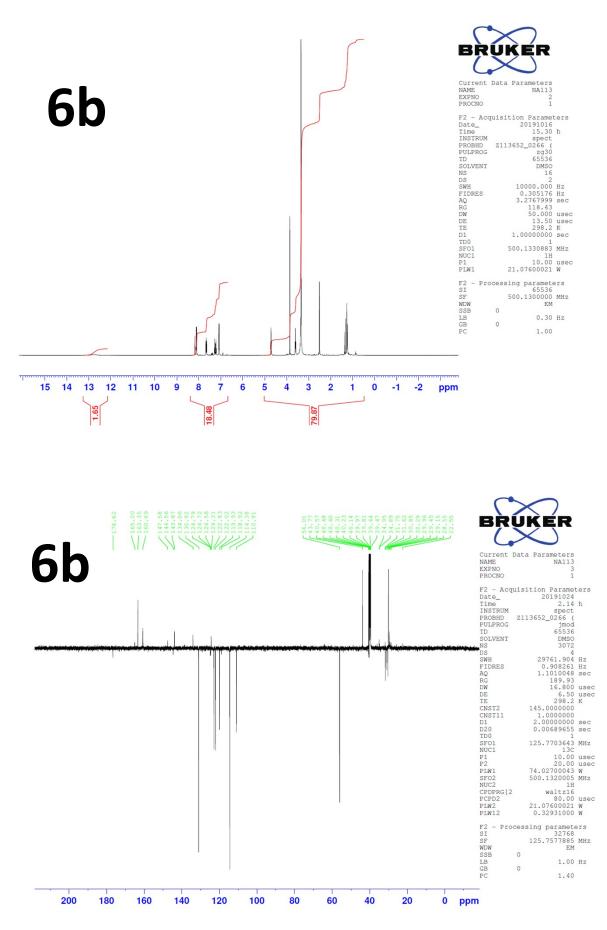
9:1 v/v, R_f 0.75. ¹H NMR (DMSO- d_6) δ : 2.85 (t, J = 7.0 Hz, 2H, CH₂CO), 3.28 (s, 6H, N(CH₃)₂), 3.58 (s, 3H, OCH₃), 4.38 (t, J = 7.0 Hz, 2H, NCH₂), 6.62 (d, J = 3.6 Hz, 1H, Ar), 7.18 (d, J = 3.7 Hz, 1H, Ar), 8.14 (s, 1H, Ar). ¹³C NMR (DMSO- d_6) δ : 34.6 (<u>C</u>H₂CO), 39.1 (N(CH₃)₂), CH₂ peak is obscured by DMSO- d_6 peak , 52.0 (OCH3), 101.7 (CH), 102.9 (C), 124.2 (CH), 150.5 (C), 151.2 (CH), 157.4 (C), 171.7 (CO).



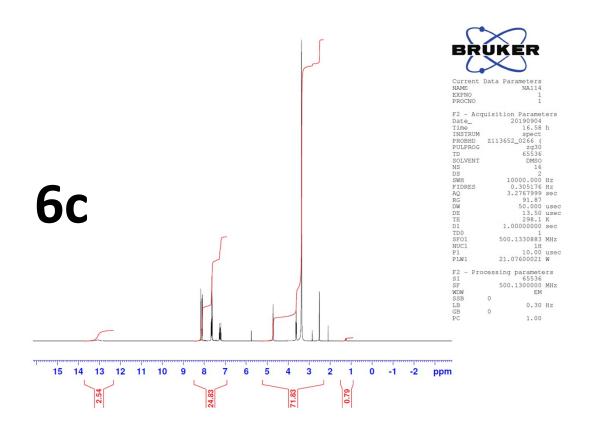


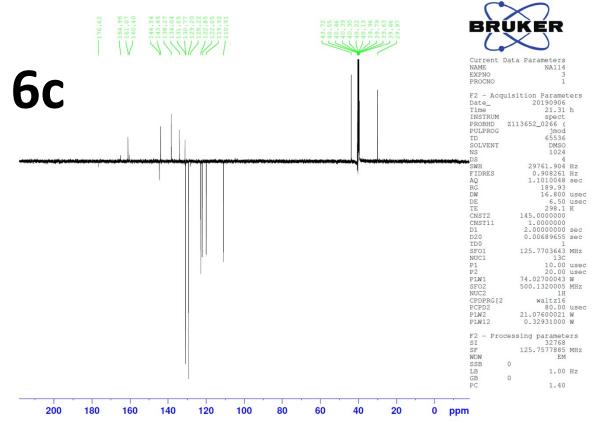


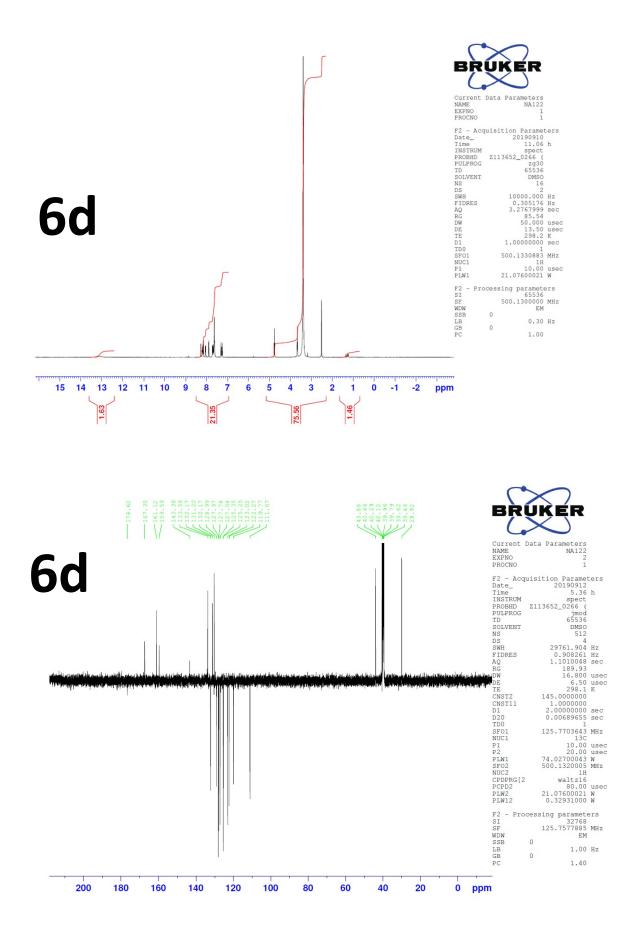


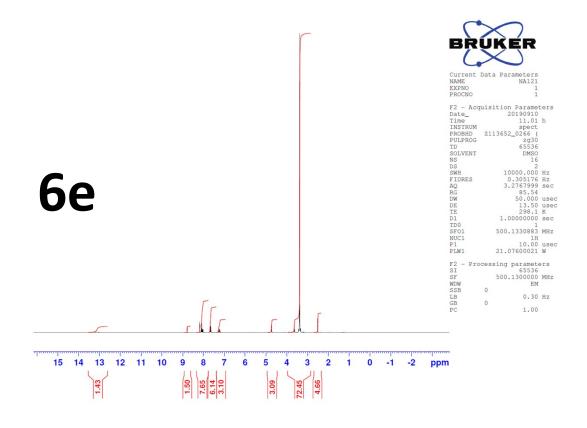


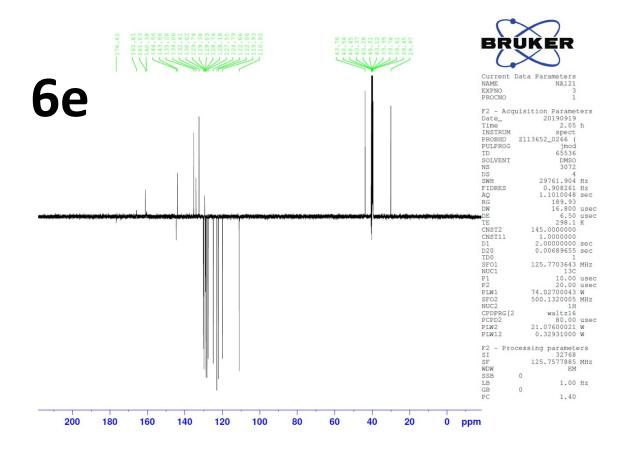
S13

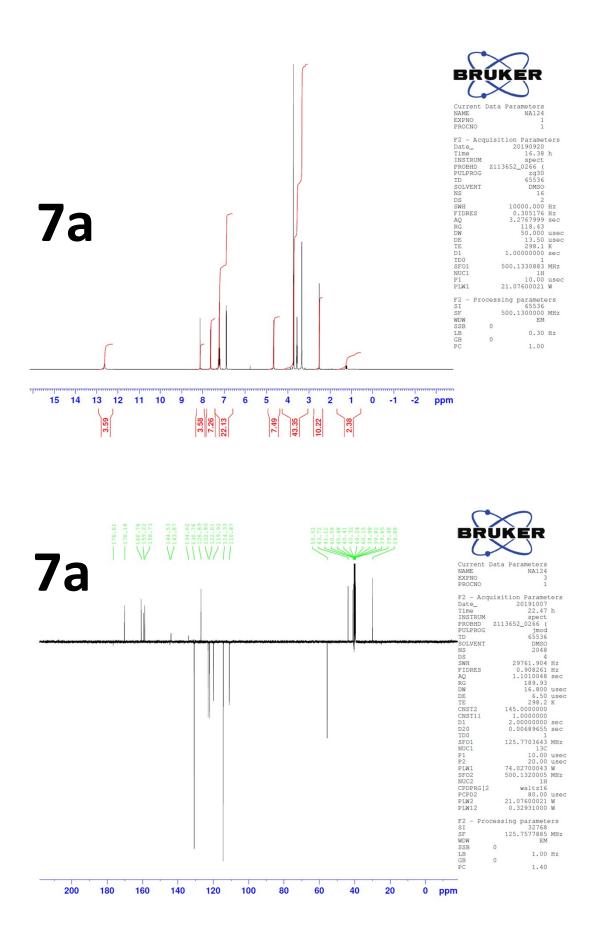


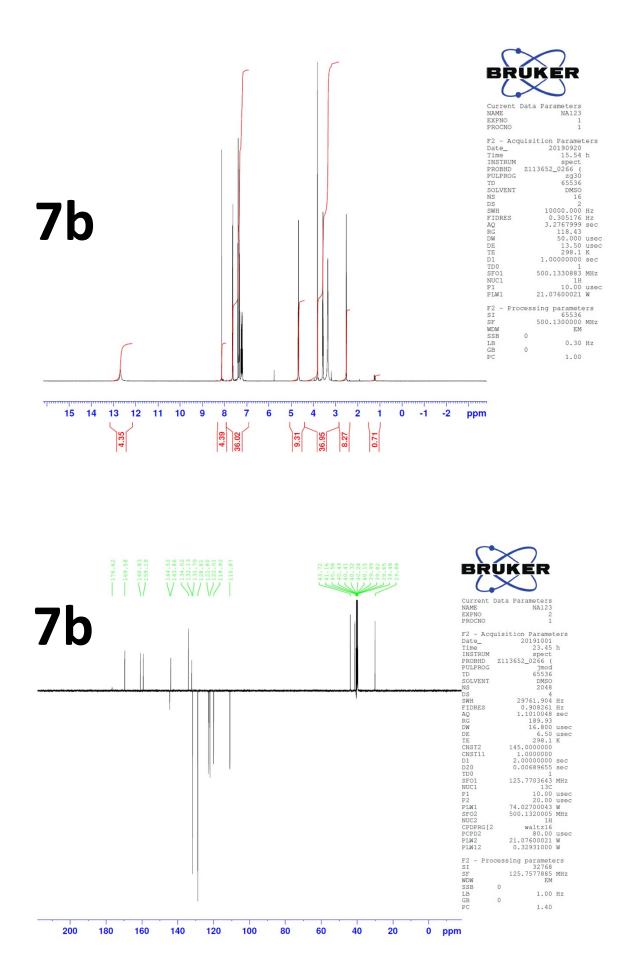


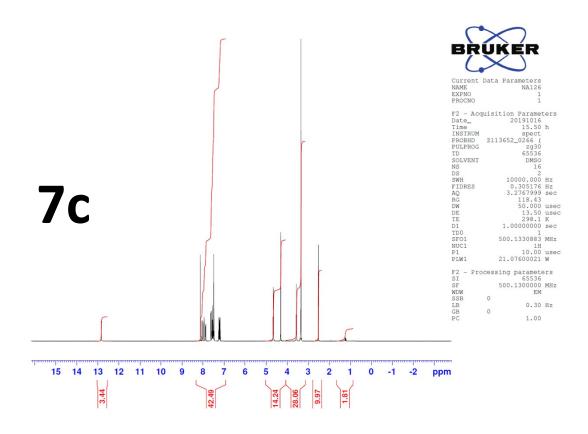


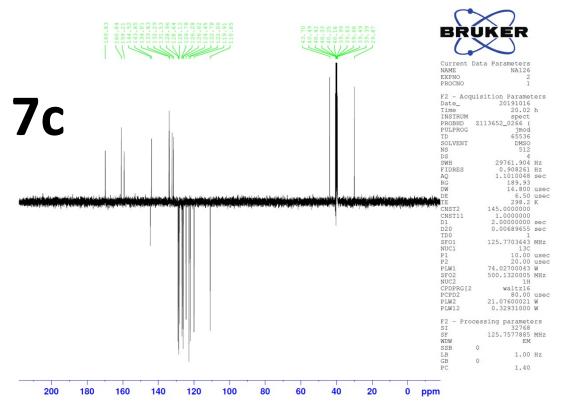


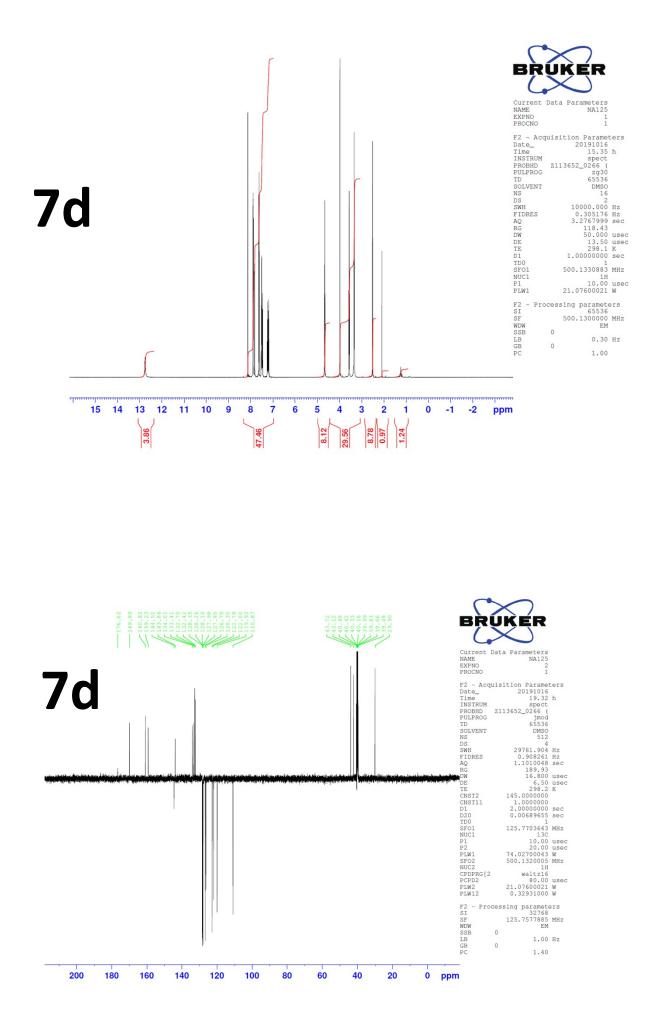


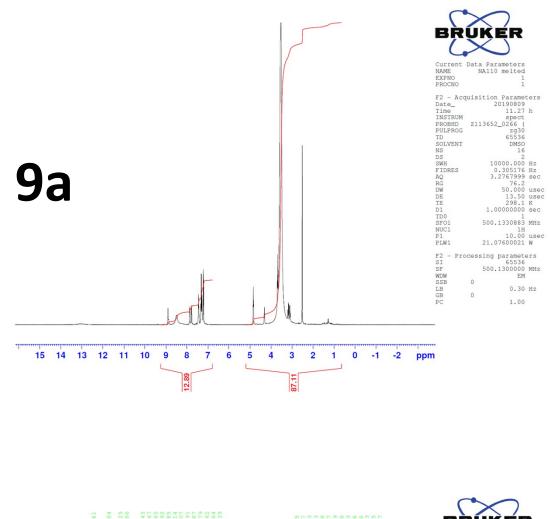


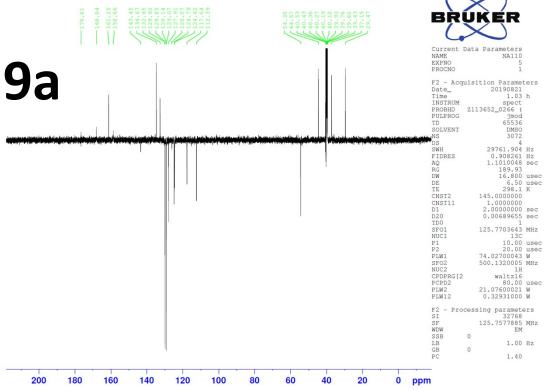


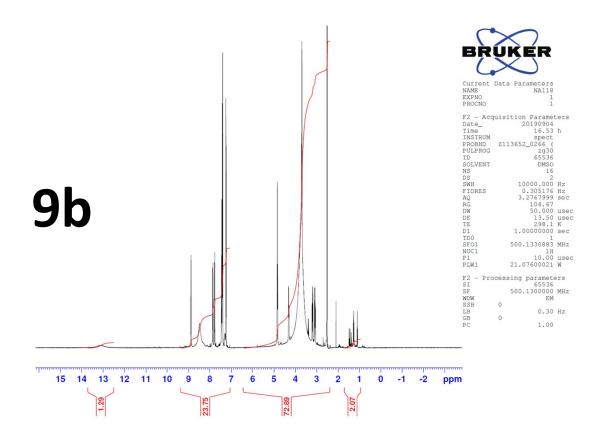


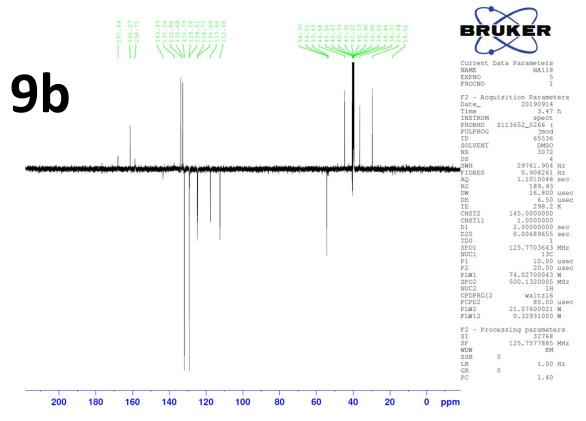


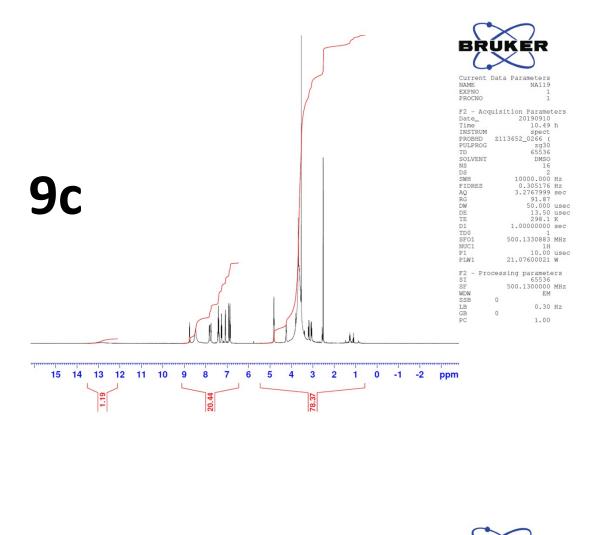


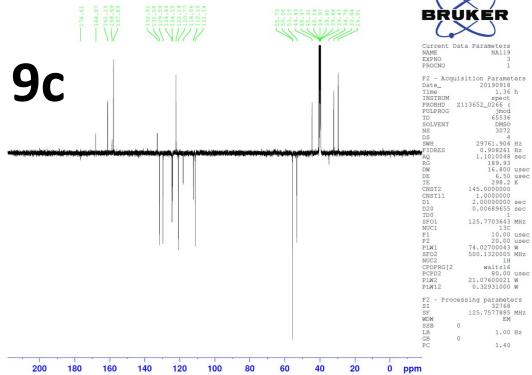


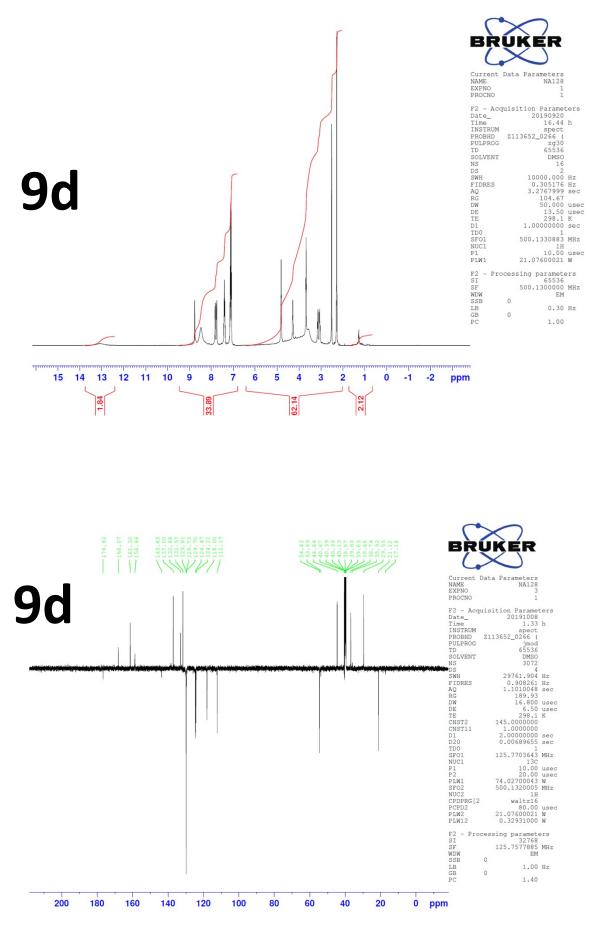


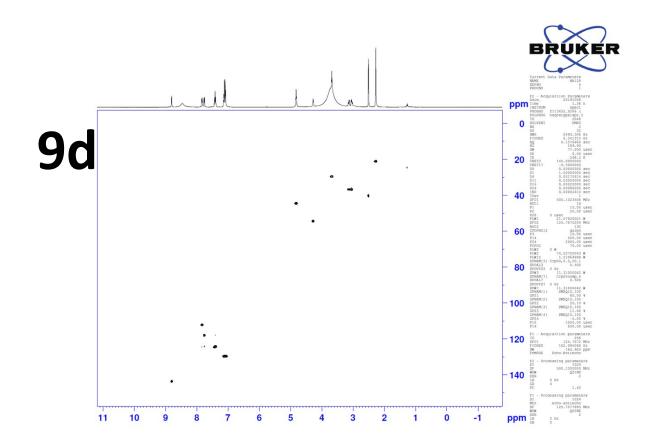


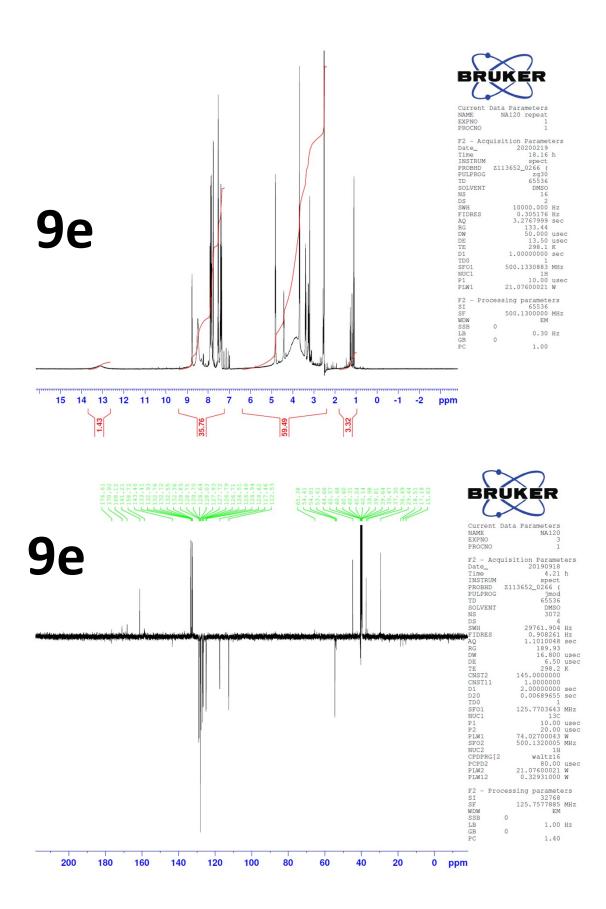


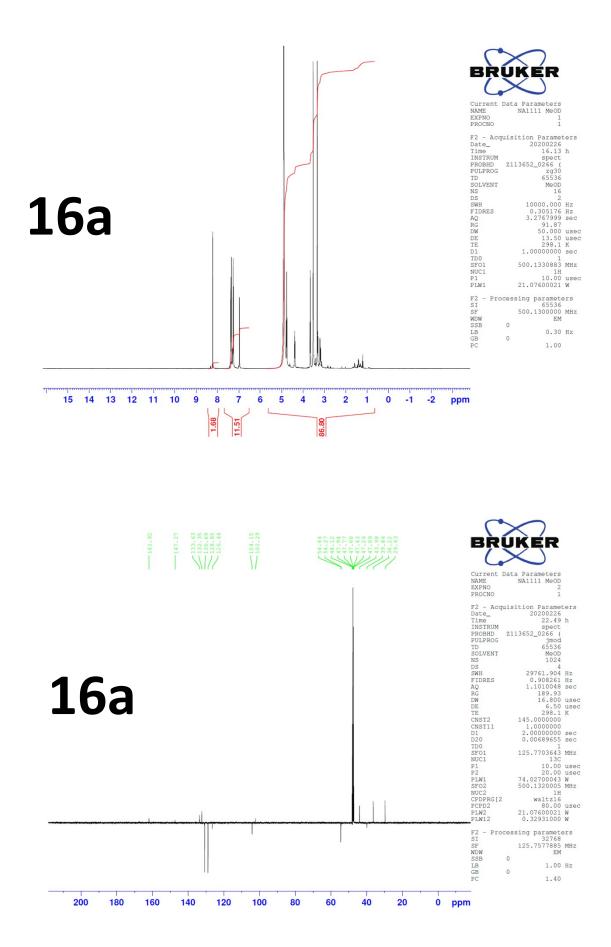


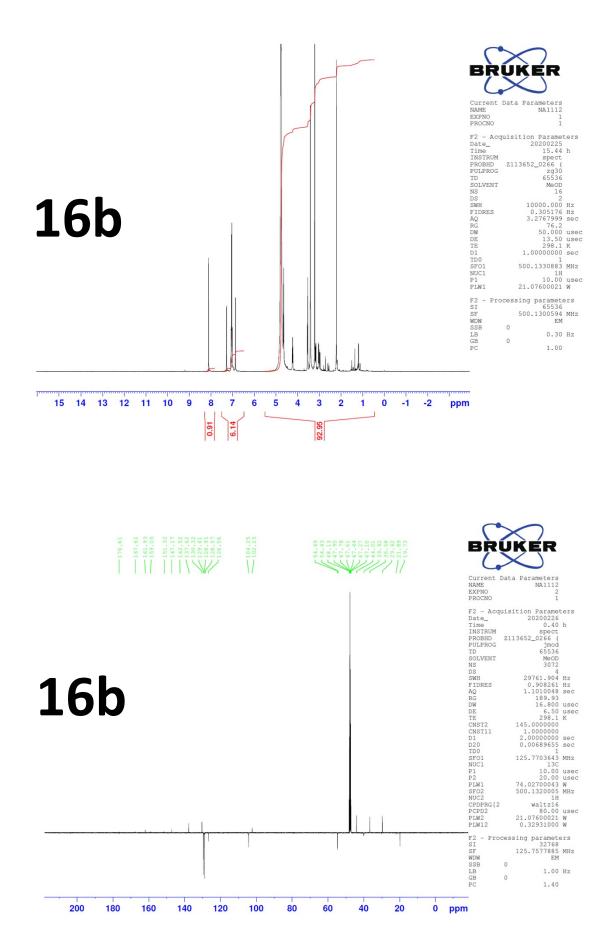


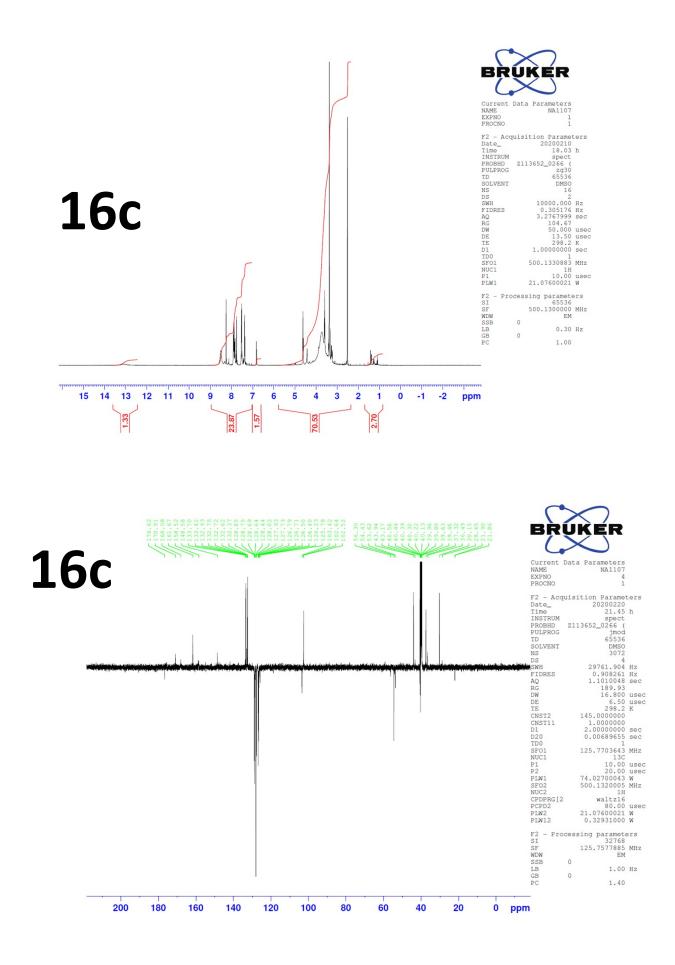












3.3. HPLC Charts:

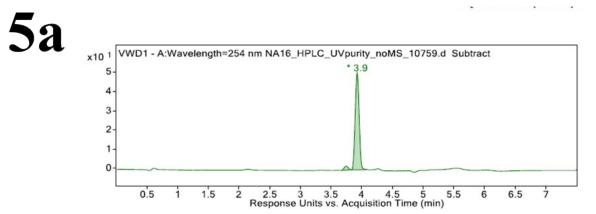


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

User Chromatogram Peak List							
RT (min)	Area	Area %	Area Sum (%)	Symmetry	Width (min)		
3.70	9.62	4.21	4.04	1.25	0.200		
3.90	228.55	100.00	95.96	1.02	0.200		

5b

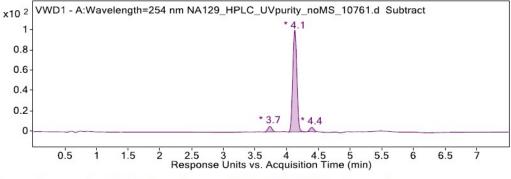
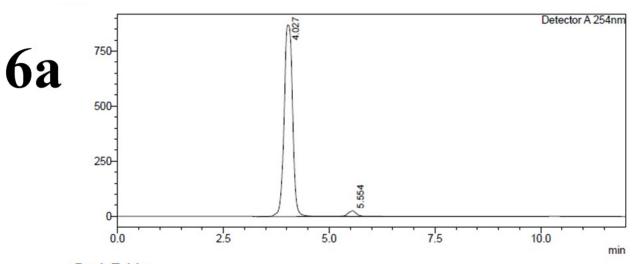


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

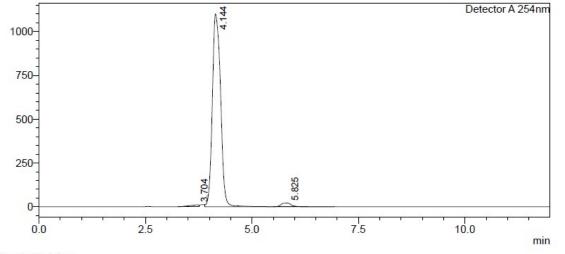
RT (min)	Area	Area %	Area Sum (%)	Symmetry	Width (min)
3.70	23.51	5.17	4.73	1.19	0.200
4.10	454.63	100.00	91.53	1.14	0.300
4.40	18.54	4.08	3.73	1.23	0.100



<Peak Table>

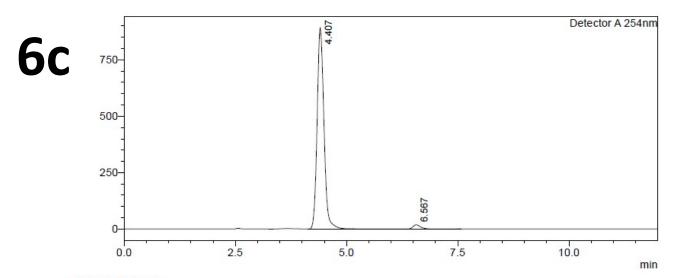
Detect	or A 254nm		N	o			
Peak#	Ret. Time	Area	Height	Conc.	Area%	Unit	Mark
1	4.027	12067622	869684	97.334	97.334		S
2	5.554	330540	24671	2.666	2.666		TV
Total		12398162	894354		100.000		





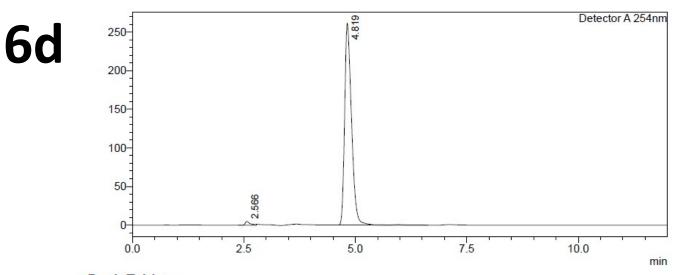
<Peak Table>

Detect	or A 254nm						
Peak#	Ret. Time	Area	Height	Conc.	Area%	Unit	Mark
1	3.704	172783	10747	1.123	1.123		
2	4.144	14873977	1100700	96.664	96.664		SV
3	5.825	340571	21063	2.213	2.213		Т
Total		15387331	1132510		100.000		



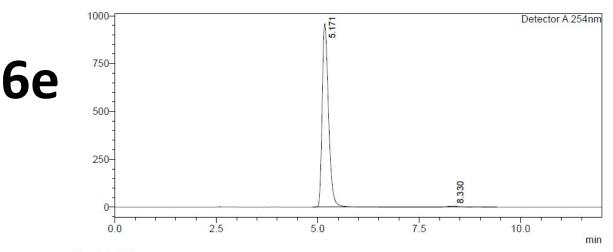
<Peak Table>

Detect	or A 254nm	6					
Peak#	Ret. Time	Area	Height	Conc.	Area%	Unit	Mark
1	4.407	9583264	891048	97.632	97.632		SV
2	6.567	232408	18956	2.368	2.368		Т
Tota		9815672	910004		100.000		



<Peak Table>

Detecto	or A 254nm	(2010A 8000012000 80		0		20
Peak#	Ret. Time	Area	Height	Conc.	Area%	Unit	Mark
1	2.566	38346	4863	1.353	1.353		
2	4.819	2794845	261036	98.647	98.647		S
Total		2833191	265899		100.000		



<Peak Table>

Detector A 254nm							
Peak#	Ret. Time	Area	Height	Conc.	Area%	Unit	Mark
1	5.171	10438829	958772	98.975	98.975		S
2	8.330	108114	5998	1.025	1.025		Т
Total		10546943	964770		100.000	2	

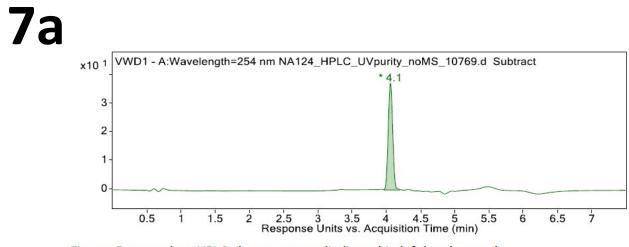


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

R (r		Area	Area %	Area Sum (%)	Symmetry	Width (min)
	4.10	169.54	100.00	100.00	1.09	0.200

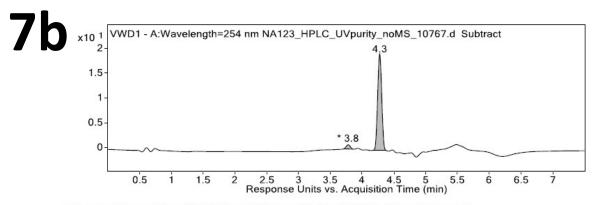


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

	0	And the second second		
User	Chromat	togram	Peak	LIST

RT (min)	Area	Area %	Area Sum (%)	Symmetry	Width (min)
3.80	3.72	4.30	4.12	1.06	0.100
4.30	86.49	100.00	95.88	0.96	0.200

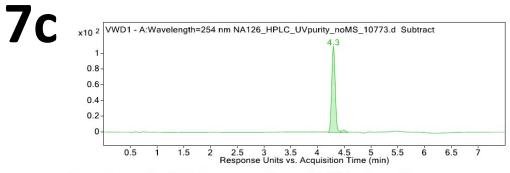


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

RT (min)	Area	Area %	Area Sum (%)	Symmetry	Width (min)
4.30	500.35	100.00	96.79	1.03	0.200
4.50	16.59	3.32	3.21	1.22	0.100

7d

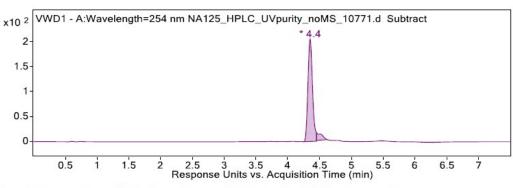
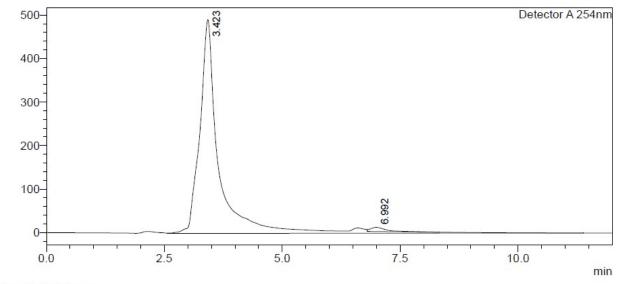


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

User Chromatogram Peak List

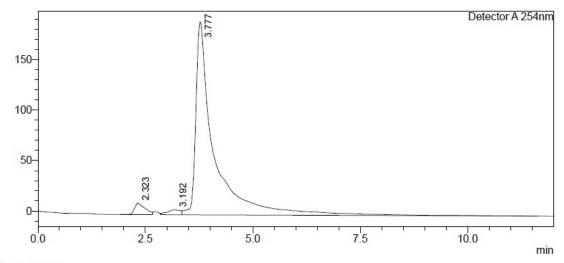
RT (min)	Area	Area %	Area Sum (%)	Symmetry	Width (min)
4.40	1031.23	100.00	92.43	1.36	0.200
4.50	84.51	8.20	7.57	Infinity	0.200



<Peak Table>

Detect	Detector A 254nm										
Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name				
1	3.423	14074909	491002	98.420	64 55 # 55 # 5 5	SV					
2	6.992	225940	9393	1.580		TV					
Total		14300849	500395								

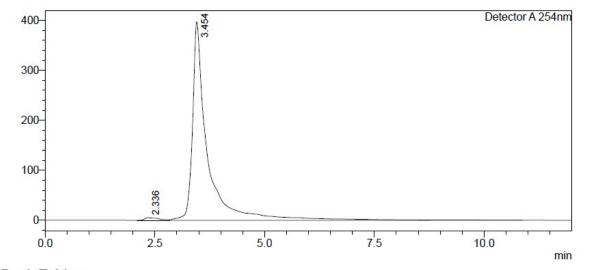
9b



<Peak Table>

Detecte Peak#	Ret. Time	Area	Height	Conc.	Area%	Unit	Mark
1	2.323	187965	11295	3.135	3.135		
2	3.192	102866	5150	1.716	1.716		V
3	3.777	5705411	190986	95.150	95.150	4	SV
Total	10.12 A. 10.10	5996242	207430		100.000		

9c



<Peak Table> Detector A 254nm

Delect	01 A 234000						
Peak#	Ret. Time	Area	Height	Conc.	Area%	Unit	Mark
1	2.336	115196	5326	1.199	1.199		
2	3.454	9492896	397489	98.801	98.801		SV
Total		9608093	402815		100.000		

9d

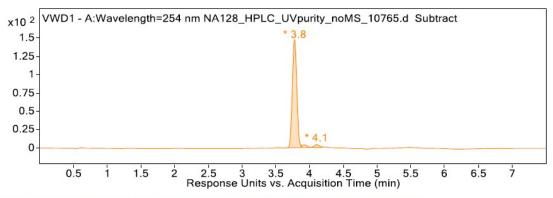


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

User Ch	nromatogra	m Peak L	ist	
RT	2 - 1	- I - And a start of the later in the		

RI (min)	Area	Area %	Area Sum (%)	Symmetry	Width (min)
3.80	684.74	100.00	94.81	0.98	0.200
3.90	17.91	2.62	2.48	1.79	0.100
4.10	19.57	2.86	2.71	1.29	0.200

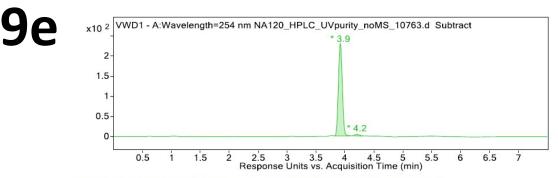


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

RT (min)	Area	Area %	Area Sum (%)	Symmetry	Width (min)
3.90	1072.78	100.00	97.96	1.09	0.300
4.20	22.36	2.08	2.04	0.95	0.200

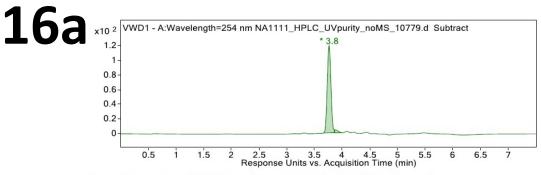


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

User Chromatogram Peak List

RT (min)	Area	Area %	Area Sum (%)	Symmetry	Width (min)
3.80	552.97	100.00	97.05	1.03	0.200
3.90	16.78	3.04	2.95	Infinity	0.100

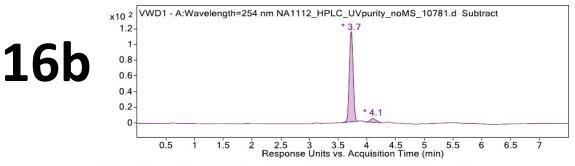


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

RT (min)	Area	Area %	Area Sum (%)	Symmetry	Width (min)
3.70	532.75	100.00	94.03	1.15	0.300
4.10	33.81	6.35	5.97	1.15	0.300

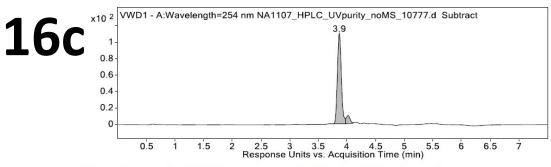


Figure: Base peak or HPLC chromatogram (indicated in left hand corner)

RT (min)	Area	Area %	Area Sum (%)	Symmetry	Width (min)
3.90	528.3	100.00	91.73	1.19	0.200
4.00	47.6	9.01	8.27	1.71	0.100

References:

- 1. P. Labute, C. Williams, M. Feher, E. Sourial and J. M. Schmidt, *J. Med. Chem.*, 2001, **44**, 1483-1490.
- F. Madeira, Y. M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R. N. Tivey, S. C. Potter, R. D. Finn and R. Lopez, *Nucleic Acids Res.*, 2019, 47, 636-641, Accessible from: [https://www.ebi.ac.uk/Tools/msa/clustalo/].
- 3. C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze and V. B. Chen, *Protein Sci.*, 2018, **27**, 293-315.
- 4. M. Wiederstein and M. J. Sippl, *Nucleic Acids Res.*, 2007, **35**, W407-W410, Accessible from: [https://prosa.services.came.sbg.ac.at/prosa.php].
- 5. M. J. Sippl, *Proteins: Struct., Funct., Bioinf.*, 1993, **17**, 355-362.
- 6. D. Eisenberg, R. Lüthy and J. U. Bowie, *Meth. Enzymol*, 1997, **277**, 396-404.
- 7. S. S. Elbaramawi, C. Hughes, J. Richards, A. Gupta, S. M. Ibrahim, E.-S. M. Lashine, M. E. El-Sadek, A. J. O'Neill, M. Wootton and J. M. Bullard, *Egypt. J. Chem.*, 2018, **61**, 9-25.
- 8. Ş. Demirayak, K. Benkli and K. Güven, *Eur. J. Med. Chem.*, 2000, **35**, 1037-1040.
- 9. G. H. Hitchings, K. W. Ledig and R. A. West, US Pat., US3037980A, 1962.