

Pyrovobasine, Hybrid Alkylated Pyrrolidine Monoterpene Indole Alkaloid Pseudodimer Discovered Using a Combination of Mass Spectral and NMR- Based Machine Learning Annotations

Hugues Fouotsa,^{a, b, c} Pierre Mkounga,^c Alain Meli Lannang,^d Jérôme Vanheuverzwijn,^b Zhiyu Zhou,^b
Karine Leblanc,^a Somia Rharrabti,^a Augustin Ephrem Nkengfack,^c Jean-François Gallard,^e Véronique
Fontaine,^b Franck Meyer,^b Erwan Poupon,^a Pierre Le Pogam^{*,a} and Mehdi A. Beniddir^{*,a}

^aÉquipe "Chimie des Substances Naturelles" Université Paris-Saclay, CNRS, BioCIS, 5 rue J.-B. Clément, 92290 Châtenay-Malabry, France. E-mail: pierre.le-pogam-alluard@universite-paris-saclay.fr ; mehdi.beniddir@universite-paris-saclay.fr

^bFaculty of Pharmacy, Microbiology, Bioorganic and Macromolecular Chemistry Unit, Université Libre de Bruxelles, Campus de la Plaine-CP 206/04, Boulevard du Triomphe, ACC.2, Po Box 1050 Belgium.

^cDepartment of Organic Chemistry, Faculty of Science, University of Yaoundé I, P.O. Box 812, Yaoundé, Cameroon.

^dDepartment of Chemistry, Higher Teachers Training College, University of Maroua, P.O. Box 55, Maroua, Cameroon.

Institut de Chimie des Substances Naturelles, CNRS, ICSN UPR 2301, Université Paris-Saclay, 91198 Gif-sur-Yvette, France.

pierre.le-pogam-alluard@universite-paris-saclay.fr
mehdi.beniddir@universite-paris-saclay.fr

Summary

S1.	Global molecular network	3
S2.	Molecular family A	5
S3.	Cartesian coordinates of 1	5
S4.	HRESIMS spectrum of pyrrovobasine (1)	8
S5.	¹ H NMR spectrum (600 MHz, CDCl ₃) of pyrrovobasine (1)	9
S6.	¹³ C NMR spectrum (150 MHz, CDCl ₃) of pyrrovobasine (1)	10
S7.	COSY spectrum (600 MHz, CDCl ₃) of pyrrovobasine (1)	11
S8.	HSQC spectrum (600 MHz, CDCl ₃) of pyrrovobasine (1)	12
S9.	HMBC spectrum (600 MHz, CDCl ₃) of pyrrovobasine (1)	13
S10.	ROESY spectrum (600 MHz, CDCl ₃) of pyrrovobasine (1)	14

S1. Global molecular network

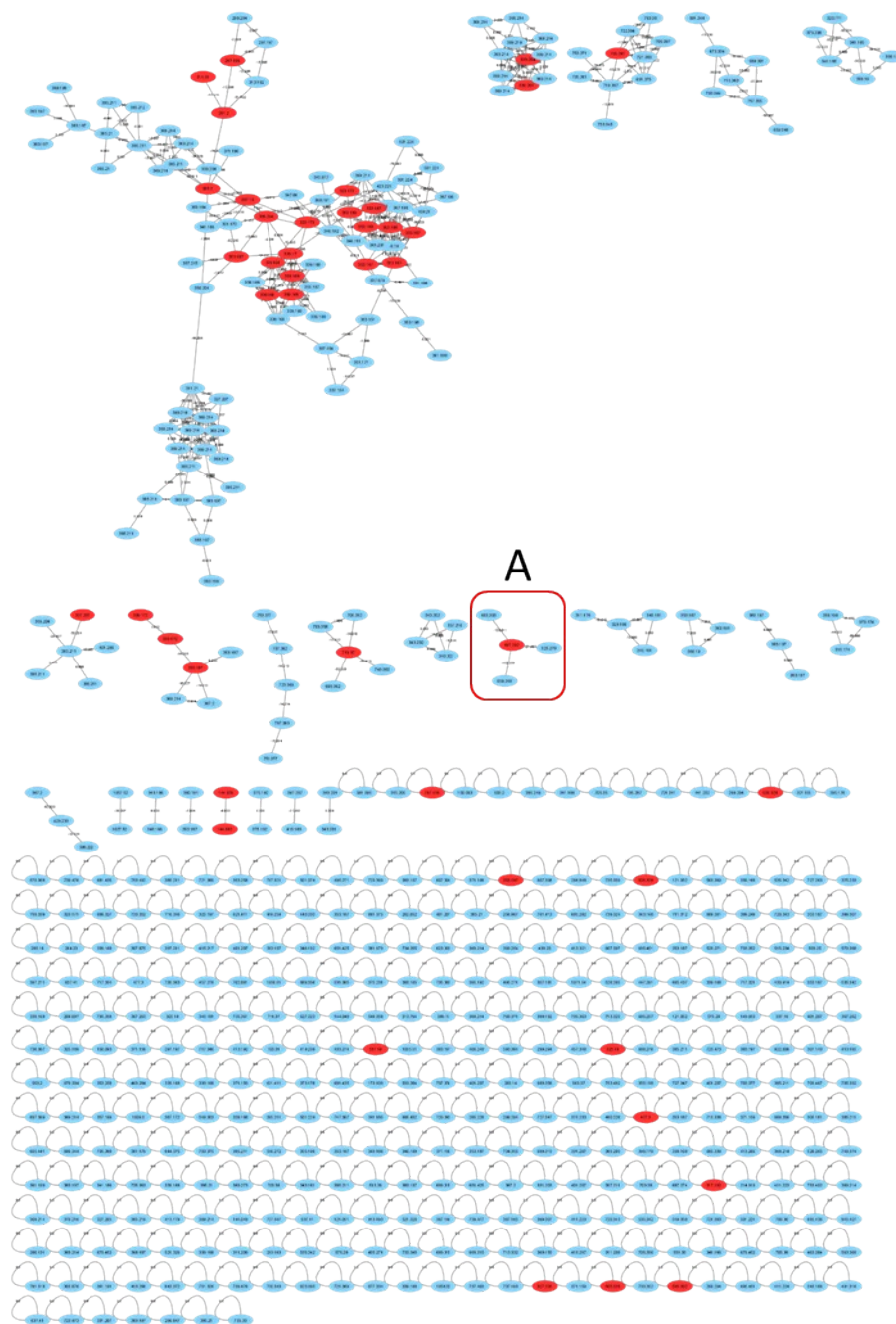


Figure S1: Molecular network of the stem bark alkaloid extract of *Voacanga africana*. Red nodes are dereplicated natural products against the GNPS library. A: cluster with targeted nodes for isolation. (this network is accessible at the following address: <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f050220a3dfb49389b1a176306c2eaa2>)

Table S1: Dereplicated MIA compounds from the GNPS library

Compound	Library	Matching score	Comment
akuammigine	GNPS	0.82	not described in <i>Voacanga</i> genus
vobtusine	MIADB	0.82	described in <i>Voacanga</i> genus
ibogamine	MIADB	0.82	not described in <i>Voacanga</i> genus
ajmalicine	GNPS	0.81	not described in <i>Voacanga</i> genus
tubotaiwine	MIADB	0.80	not described in <i>Voacanga</i> genus
pericyclivine	MIADB	0.78	not described in <i>Voacanga</i> genus
ibogaine	MIADB	0.76	described in <i>Voacanga</i> genus
vobasine	MIADB	0.75	described in <i>Voacanga</i> genus
perivine	MIADB	0.75	not described in <i>Voacanga</i> genus
pleiocarpamine	MIADB	0.71	not described in <i>Voacanga</i> genus
ibogamine-18-carboxylic acid, methyl ester	GNPS	0.71	not described in <i>Voacanga</i> genus
iboxygaine	MIADB	0.71	not described in <i>Voacanga</i> genus
tabersonine	MIADB	0.71	described in <i>Voacanga</i> genus
ceridimine	MIADB	0.70	not described in <i>Voacanga</i> genus
retuline	MIADB	0.69	not described in <i>Voacanga</i> genus
voacamine	MIADB	0.66	described in <i>Voacanga</i> genus
alloyohimbine	MIADB	0.65	not described in <i>Voacanga</i> genus
corynantheidal	MIADB	0.65	not described in <i>Voacanga</i> genus

(this table is accessible at the following address:

https://gnps.ucsd.edu/ProteoSAFe/result.jsp?task=f050220a3dfb49389b1a176306c2eaa2&view=view_all_annotations_DB)

S2. Molecular family A

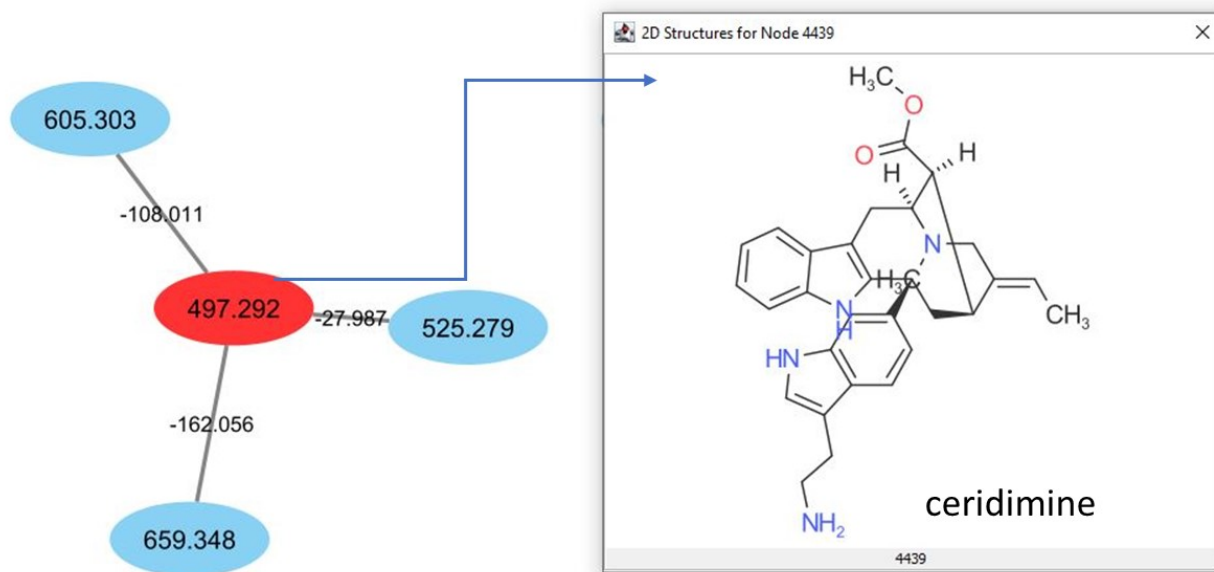


Figure S2: molecular family A (red node = ceridimine as it was annotated by GNPS, m/z 605.303 was herein targeted for isolation and structure elucidation)

S3. Cartesian coordinates of 1

Electronic energy (B3LYP): -1953.9174848122968 Ha.

Lowest frequency: 11.6382 cm⁻¹.

Free enthalpy: -1953.275435 Ha.

Coordinates of the lowest energy conformer found at the B3LYP/6-31G* level

1 C -4.823921 3.184522 -2.454515

2 C -3.753586 3.584850 -3.279894

3 C -2.452749 3.162931 -3.026361

4 C -2.249151 2.327662 -1.922815

5 N -1.099544 1.748387 -1.434151

6 C -3.309781 1.916864 -1.071639

7 C -4.613489 2.356940 -1.357758

8 C -2.740083 1.081727 -0.037408

9 C -3.533932 0.408775 1.051407

10 C -1.382387 1.001633 -0.296092

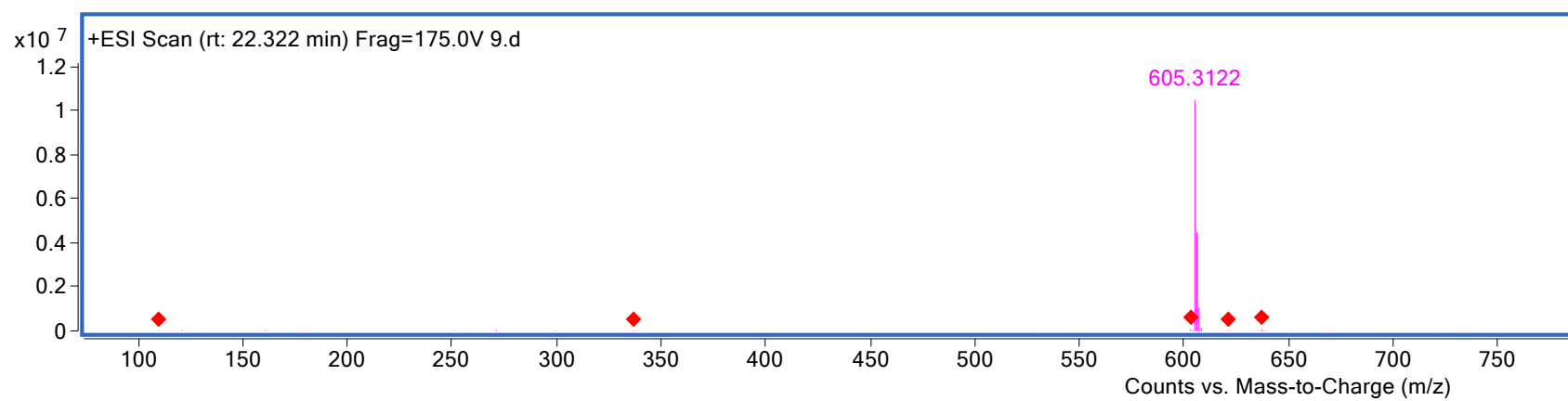
11 C -0.231664 0.283568 0.372519

12 C -0.639758 -0.665785 1.525665

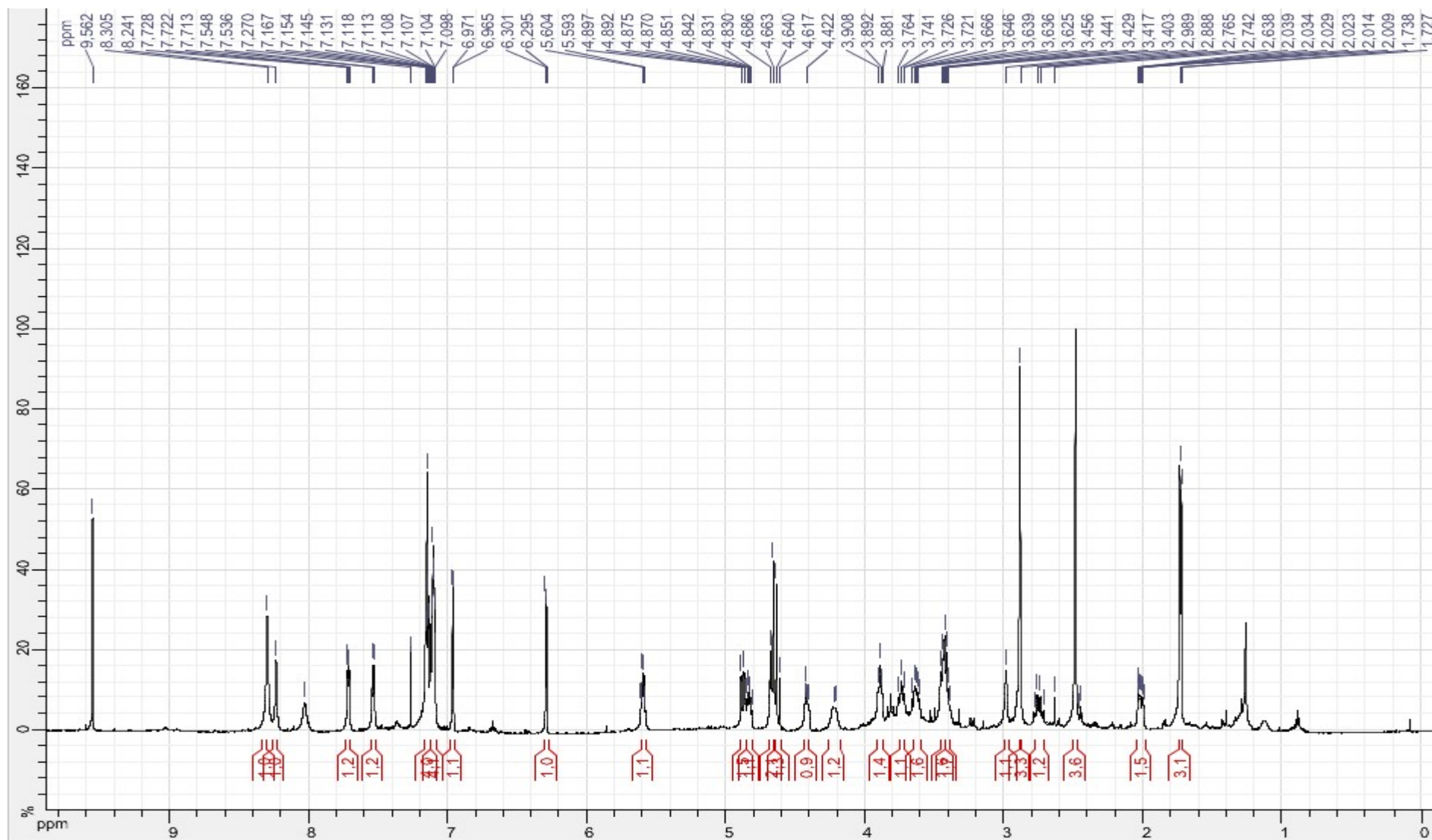
13 C -1.393591 -1.981096 1.167067
14 C -1.870530 -2.636269 2.455936
15 C -1.267454 -3.673266 3.053615
16 C -0.008853 -4.385033 2.640821
17 C -3.112394 -2.006646 3.061500
18 N -4.196803 -1.832354 2.085784
19 C -5.410018 -1.325400 2.705745
20 C -3.810198 -1.120975 0.855366
21 C -2.640814 -1.921426 0.244987
22 C -2.352987 -1.673903 -1.226464
23 O -1.276262 -1.846812 -1.781624
24 O -3.465791 -1.386270 -1.916065
25 C -3.312809 -1.213840 -3.337334
26 C 0.817272 1.274455 0.836386
27 N 0.460189 2.513806 1.348754
28 C 2.190796 1.143519 0.880942
29 C 2.708207 2.378704 1.418991
30 C 4.005173 2.855783 1.682179
31 C 1.594701 3.216175 1.700843
32 C 1.741005 4.500916 2.234275
33 C 3.034234 4.944680 2.487843
34 C 4.153282 4.131184 2.213461
35 C 2.994140 -0.064582 0.484386
36 C 3.426808 -0.042625 -0.994382
37 N 4.165218 -1.251582 -1.378144
38 C 5.527407 -1.460159 -1.190339
39 C 6.459258 -0.523485 -0.606312
40 O 6.210801 0.590154 -0.153516
41 H 7.499739 -0.913889 -0.590811
42 C 3.620969 -2.382288 -1.928956
43 C 2.170520 -2.519409 -2.272120
44 O 1.374151 -2.497919 -1.090870
45 H 0.439785 -2.378998 -1.356542
46 C 4.637086 -3.319042 -2.111679
47 C 5.830117 -2.742501 -1.654404
48 H -5.828439 3.531084 -2.681429
49 H -3.947079 4.235262 -4.128472
50 H -1.626552 3.471842 -3.661502
51 H -0.163852 1.916992 -1.773117
52 H -5.449369 2.056469 -0.730852
53 H -4.504731 0.911071 1.127052
54 H -3.057706 0.563333 2.023579

55 H 0.250154 -0.341109 -0.388381
56 H 0.288117 -0.984550 2.011426
57 H -1.177783 -0.096431 2.287642
58 H -0.669857 -2.632857 0.672315
59 H -1.726401 -4.058882 3.965997
60 H -0.205718 -5.452207 2.469567
61 H 0.740630 -4.333756 3.442197
62 H 0.448431 -3.978708 1.734560
63 H -3.492415 -2.647641 3.865096
64 H -2.844733 -1.043302 3.545044
65 H -6.200161 -1.243026 1.950709
66 H -5.748249 -2.032451 3.472420
67 H -5.295265 -0.338132 3.192065
68 H -4.658575 -1.215599 0.170518
69 H -3.023092 -2.955263 0.231566
70 H -2.910098 -2.122805 -3.791021
71 H -4.315249 -1.005493 -3.710249
72 H -2.647515 -0.373335 -3.544965
73 H -0.483983 2.870573 1.356689
74 H 4.873739 2.244122 1.453163
75 H 0.878338 5.128772 2.443052
76 H 3.184375 5.937902 2.902720
77 H 5.149680 4.513600 2.418167
78 H 2.427020 -0.985156 0.650251
79 H 3.898045 -0.119224 1.099953
80 H 4.068029 0.817584 -1.185391
81 H 2.548544 0.024525 -1.641434
82 H 2.049659 -3.466205 -2.818859
83 H 1.863452 -1.713246 -2.958520
84 H 4.501120 -4.306599 -2.531853
85 H 6.816394 -3.188239 -1.644206

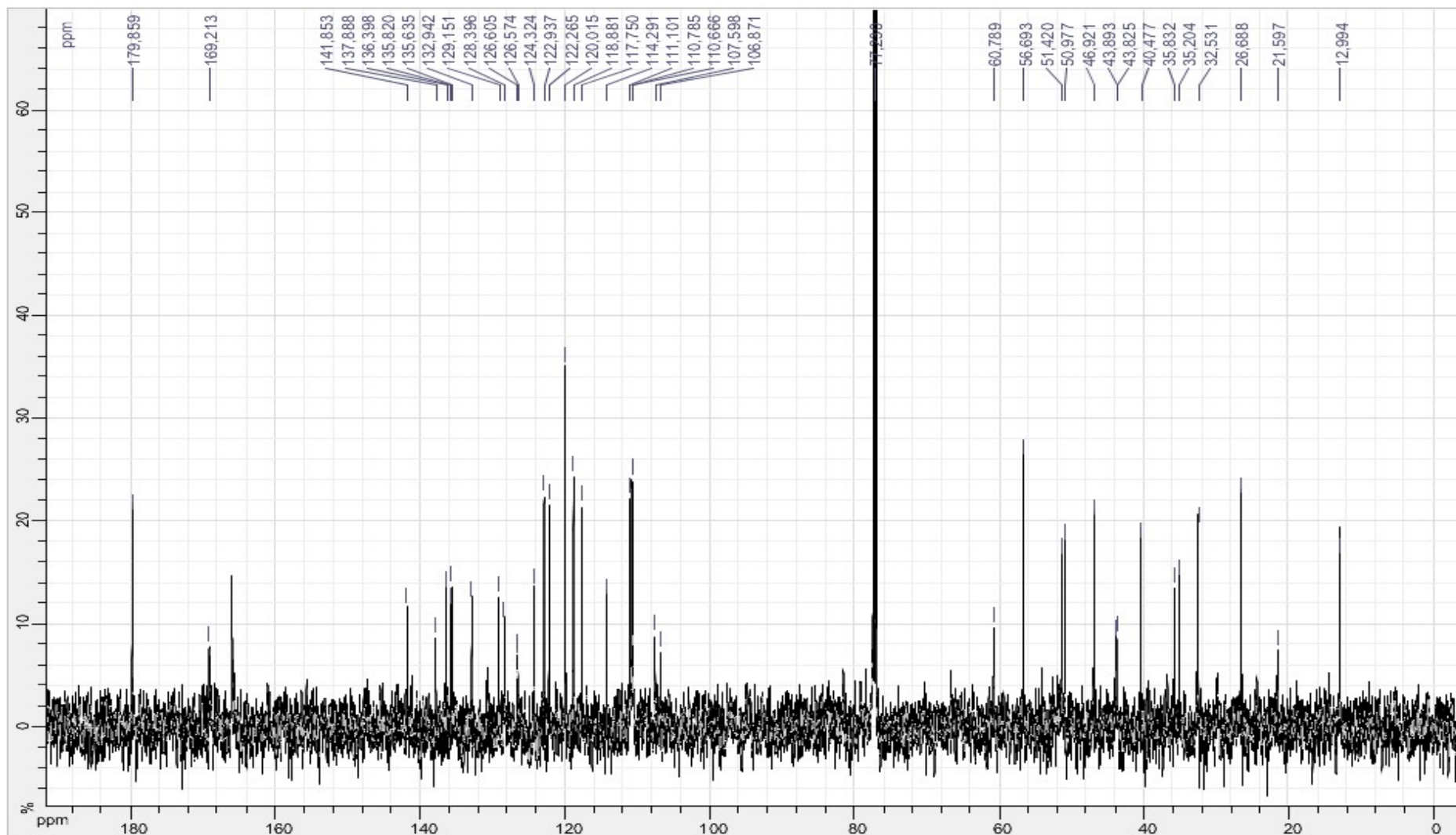
S4. HRESIMS spectrum of pyrrovobasine (1)



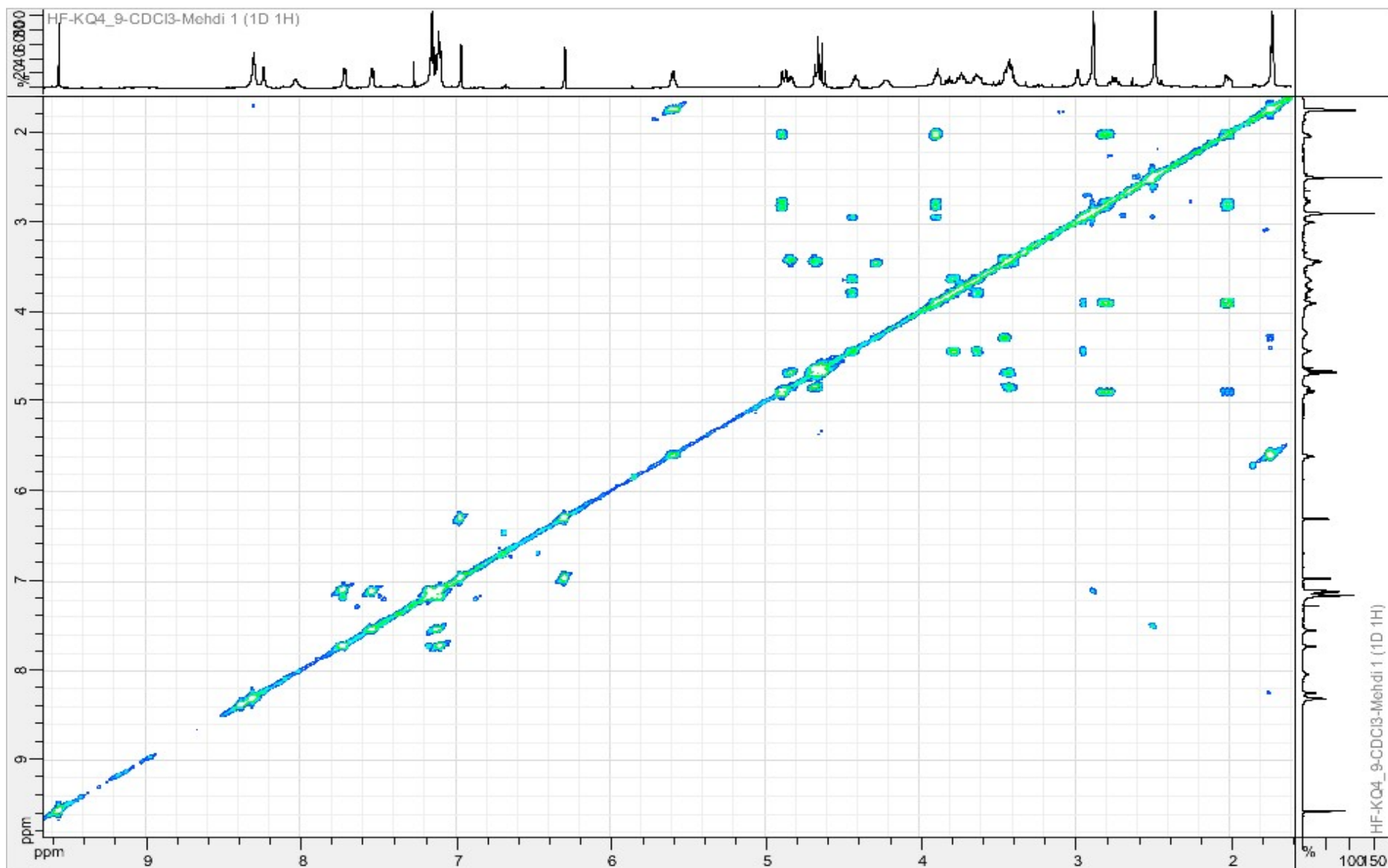
S5. ¹H NMR spectrum (600 MHz, CDCl₃) of pyrrovobasine (1)



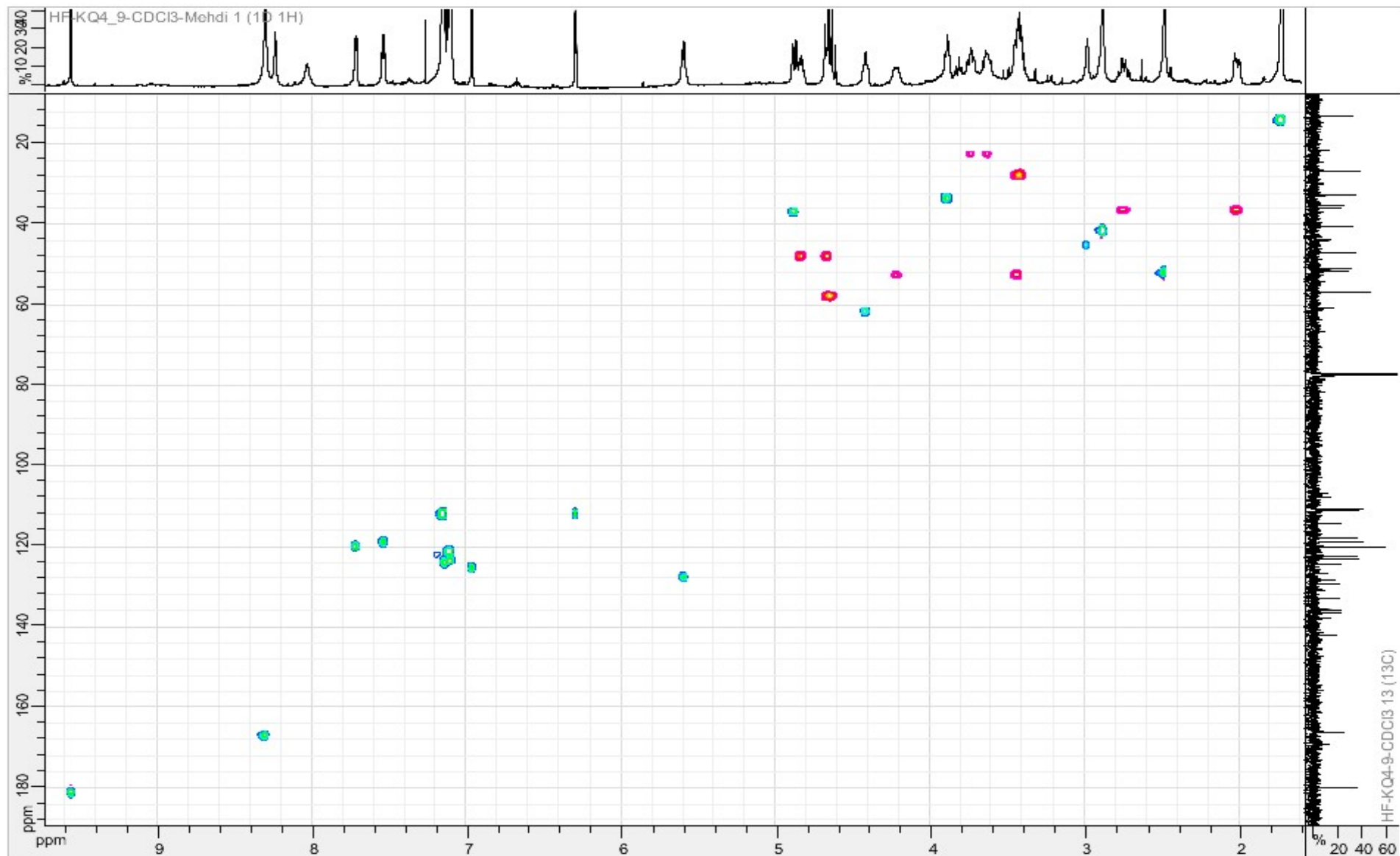
S6. ^{13}C NMR spectrum (150 MHz, CDCl_3) of pyrrovobasine (1)



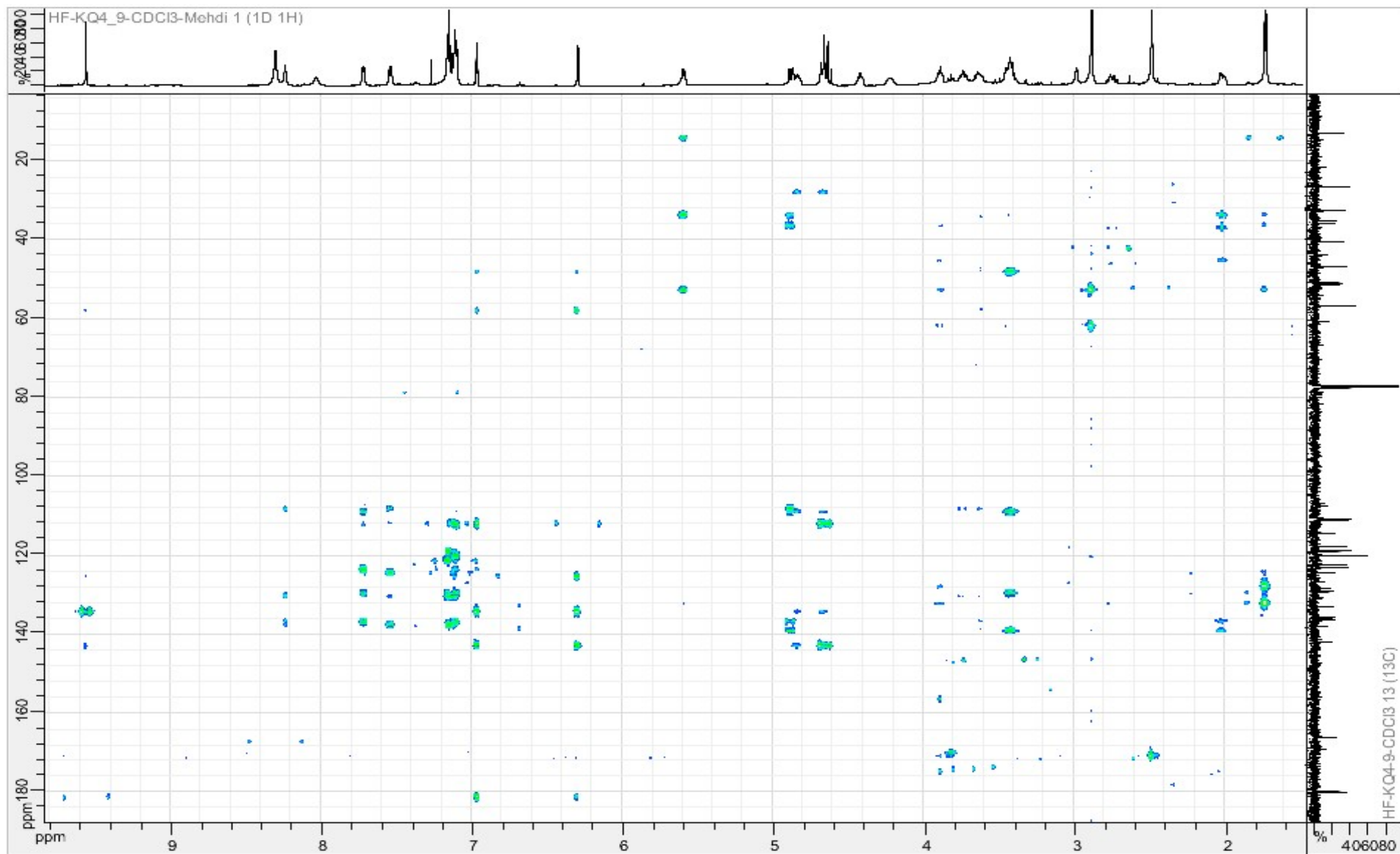
S7. COSY spectrum (600 MHz, CDCl₃) of pyrrovobasine (1)



S8. HSQC spectrum (600 MHz, CDCl₃) of pyrrovobasine (1)



S9. HMBC spectrum (600 MHz, CDCl₃) of pyrrovobasine (1)



S10. ROESY spectrum (600 MHz, CDCl₃) of pyrrovobasine (1)

