Untargeted Proteomics and Systems-Based Mechanistic Investigation of Artesunate in Human Bronchial Epithelial Cells

Kodihalli C. Ravindra[1], Wanxing Eugene Ho[2,3], Chang Cheng[4,5], Luiz C. Godoy[1], John S. Wishnok[1], Choon Nam Ong[2], W.S. Fred Wong[5], Gerald N. Wogan[1], Steven R. Tannenbaum[1,3]*

Author Affiliation:

[1] Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2] Saw Swee Hock School of Public Health, National University of Singapore, Singapore 119228

[3] Singapore-MIT Alliance for Research and Technology (SMART), Singapore 138602

[4] Department of Gastroenterology & Hepatology, Singapore General Hospital, Singapore 169608

[5] Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119228

* Corresponding author

Email: srt@mit.edu

Supporting information

Table of Contents

Table S1: Protein list identified by Pseudo-shotgun proteomics and In-Gel digestion method.

	Protein name (pseudo-shotgun proteomics)	Spectral counts
1	Ribosomal protein S18, isoform CRA_c	26
2	Serum albumin	21
3	Triosephosphate isomerase isoform 2	18
4	Alpha-ketoglutarate-dependent dioxygenase FTO	17
5	KH-type splicing regulatory protein (FUSE binding protein 2), isoform CRA_b	16
6	60S ribosomal protein L14	15
7	Glyceraldehyde-3-phosphate dehydrogenase	15
8	Peroxiredoxin-1	14
9	**Enolase 1 variant**	13
10	40S ribosomal protein S17	12
11	Microtubule-associated protein 4 isoform 1 variant	12
12	Heat shock protein beta-1	11
13	Splicing factor proline/glutamine-rich (polypyrimidine tract binding protein associated)	11
14	40S ribosomal protein S9	10
15	**Actin, cytoplasmic 2**	10
16	Alpha-tubulin	10
17	Ribosomal protein S19, partial	10
18	Small nuclear ribonucleoprotein Sm D2 isoform 1	9
19	Vimentin variant	9
20	Cytochrome C	8
21	Eukaryotic translation elongation factor 1 alpha 1 variant	8
22	Profilin-1	8
23	Ribosomal protein S5 variant	8
24	Heterogeneous nuclear ribonucleoprotein K	7
25	**L-lactate dehydrogenase A chain isoform 3**	7
26	Calpastatin isoform a variant	6
27	Chromosome 11 open reading frame 48, isoform CRA_c	6
28	Ribosomal protein S6	6
29	Splicing factor 1 isoform 6	6
30	**Tubulin, beta**	6
31	40S ribosomal protein S13	5
32	60S ribosomal protein L35	5
33	Aldolase	5
34	Cofilin-1	5
35	Ribosomal protein L23a, isoform CRA_a	5
36	Ribosomal protein L4, isoform CRA_b	5

37	ATP synthase subunit g, mitochondrial	4
38	**ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit (oligomycin sensitivity conferring protein)**	4
39	Desmoplakin I	4
40	Filamin-A isoform 1	4
41	**Phosphoglycerate kinase 1**	4
42	Pyruvate kinase isozymes M1/M2 isoform c	4
43	Ribosomal protein L10	4
44	Ribosomal protein L18	4
45	Transgelin 2	4
46	Transgelin, isoform CRA_c	4
47	Tubulin, beta 4, isoform CRA_b	4
48	Zyxin, isoform CRA_b	4
49	40S ribosomal protein S28	3
50	Galectin-1	3
51	Heterochromatin protein 1-binding protein 3	3
52	Heterogeneous nuclear ribonucleoproteins A2/B1 isoform B1	3
53	Nucleophosmin	3
54	Peptidyl-prolyl cis-trans isomerase A precursor	3
55	Serine/arginine-rich splicing factor 4	3
56	Tropomyosin 1 (alpha)	3
57	Annexin A2	2
58	Chloride intracellular channel 1 variant	2
59	Heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA binding protein 1, 37kDa), isoform CRA_a	2
60	**Malate dehydrogenase, cytoplasmic isoform 1**	2
61	Ribosomal protein S8, isoform CRA_a	2
62	Signal recognition particle 14 kDa protein	2
63	Cysteine and glycine-rich protein 1 isoform 1	2
64	Phosphoglycerate mutase	2

	In-gel digestion method	**Spectral counts**
1	Myosin-9	66
2	**Actin, cytoplasmic 2**	56
3	**Phosphoglycerate kinase 1**	38
4	Tubulin alpha-1A chain isoform 2	27
5	**Tubulin, beta**	27
6	Pyruvate kinase isoenzymes M1/M2 isoform a	26
7	**Enolase 1 variant**	22
8	Heat shock cognate 71 kDa protein isoform 1	22
9	Elongation factor 1	19
10	Glyceraldehyde-3-phosphate dehydrogenase isoform 1	17
11	Aldehyde dehydrogenase 7 family, member A1	17

12	**ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit 1, cardiac muscle**	13
13	Protein disulfide-isomerase precursor	12
14	**Malate dehydrogenase, cytoplasmic isoform 1**	9
15	**L-lactate dehydrogenase A chain isoform 3**	8
16	Glucose-6-phosphate isomerase isoform 1	8
17	Glucose-6-phosphate 1-dehydrogenase isoform b	8
18	Probable ATP-dependent RNA helicase DDX5	6
19	Chaperonin containing TCP1, subunit 3 (gamma) variant	6
20	L-lactate dehydrogenase B chain	5
21	Dihydropyrimidinase-related protein 2 isoform 1	5
22	Puromycin-sensitive aminopeptidase	4
23	IQ motif containing GTPase activating protein 1	4
24	Serine hydroxymethyltransferase, mitochondrial isoform 1 precursor	4
25	BAT1 protein	4

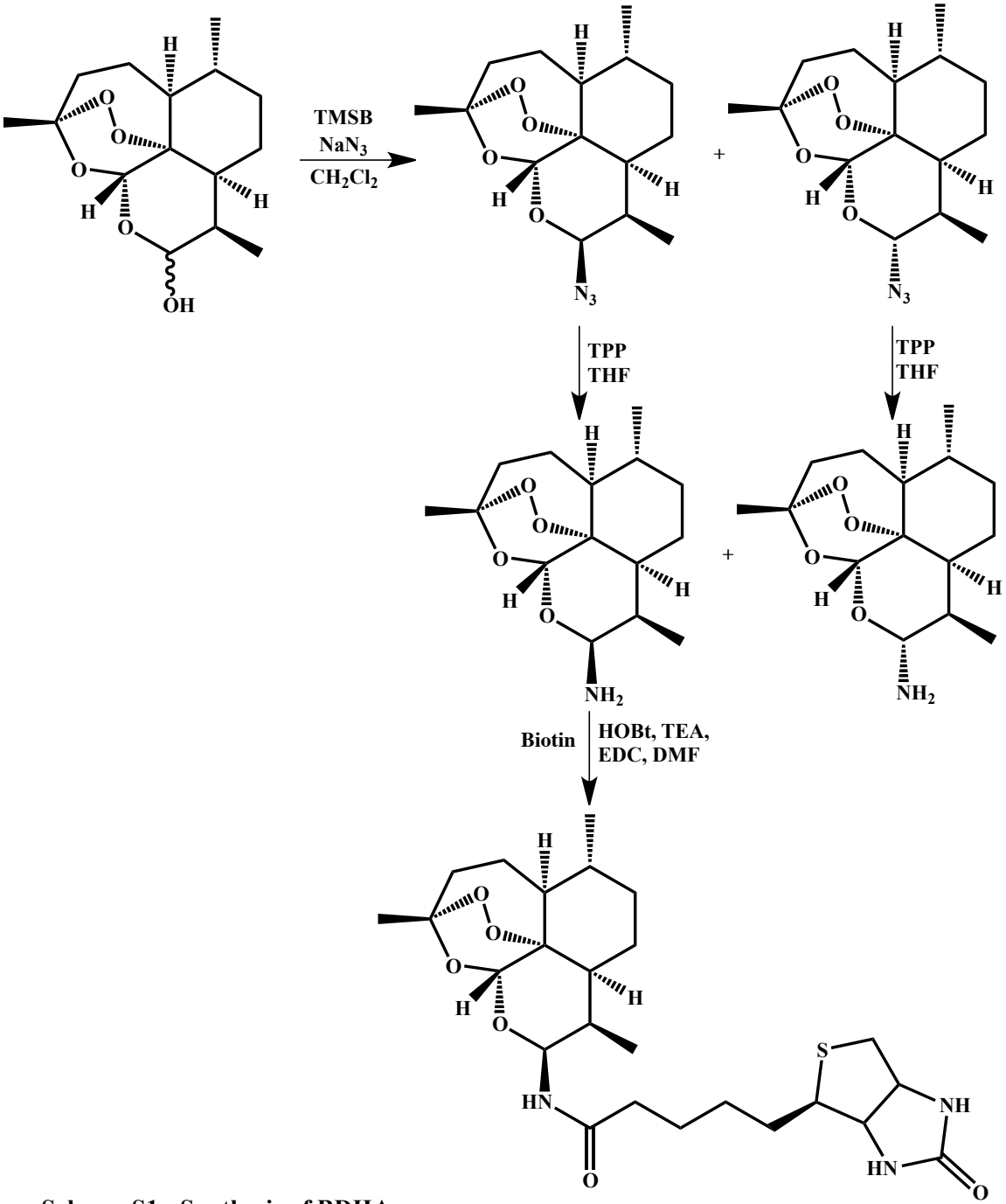
***** highlighted proteins are identified by in-gel as well as pseudo-shotgun proteomics methods.

* Spectral counts in Table 1 are the average spectral counts of triplicate experiments.

Note: The term 'Pseudo-shotgun proteomics' reflects that the proteins were enriched prior to digestion by the elution of biotinylated proteins (BDHA targets) and pre-fractionated by off-gel electrophoresis.

Synthesis of chemical probes

10β-Aminodihydroartemisinin (a-DHA) is synthesized based on earlier published protocol[1] and summarized below.



Scheme S1: Synthesis of BDHA.

Synthesis of biotinylated dihydroartemisinin (BDHA):

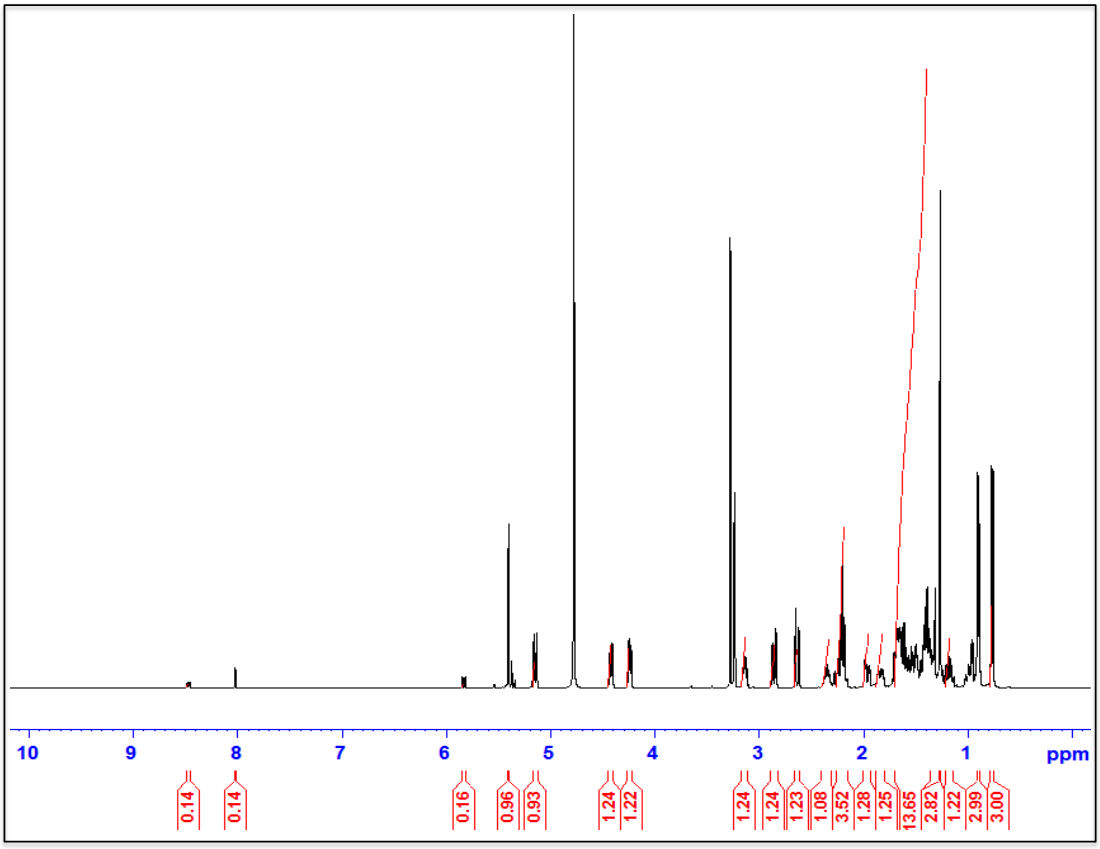
To the suspension of biotin (100 mg, 0.41 mol) in dry DMF (2 mL), triethylamine (124mg, 1.22 mol) was added, resulting in a clear solution. Subsequently, 1-hydroxybenzotriazole (55.3 mg, 0.41 mol), 10β-aminodihydroartemisinin (229.47 mg, 0.45 mol) and 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide (235.4 mg, 1.22 mol) were added. The reaction mixture was stirred at room temperature for 2 hours and then the mixture was diluted with ethyl acetate and washed with 1N hydrochloric acid solution. The solvent was distilled under reduced pressure, and the residue was extracted with ethylacetate, washed with water, dried over anhydrous sodium sulfate, concentrated under reduced pressure, and purified by HPLC to obtain biotinylated DHA (BDHA). The analytical data was as follows: high-resolution mass spectrum (ESI) m/z: 510.2633 $[M+H^+]^+$ (calculated: 510.2648); ^{1}H-NMR (400MHz, MeOD) δ 5.41 (1H, s, 12-H), 5.16-5.14 (1H, d), 4.43-4.40 (1H, dd), 4.25-4.22 (1H, dd), 3.16-3.12 (1H, m), 2.88-2.83 (1H, dd), 2.65-2.62 (1H, d, *J*=12.92 Hz), 2.39-2.32 (1H, m), 2.25-2.16 (3H, m), 1.99-1.94 (1H, td), 1.88-1.81 (1H, m), 1.71-1.38 (13H, m), 1.27 (3H, s, CH_3), 1.22-1.14 (1H, m), 0.91-0.89 (3H, d, *J*=6.32 Hz, 6-CH_3), 0.77-0.76 (3H, d, *J*=7.16 Hz, 9-CH_3).

Reference:

1. Xie, L., Zhai, X., Ren, L., Meng, H., Liu, C., Zh, W., and Zhao, Y. (2011) Design, Synthesis and Antitumor Activity of Novel Artemisinin Derivatives Using Hybrid Approach. *Chem. Pharm. Bull. 59*, 984-990.

NMR analysis: NMR spectra of BDHA were determined with a Bruker 400 MHz spectrometer. The solvent was deuterated methanol. Chemical shifts are (δ) reported in parts per million (ppm) and the residual proton signals in the solvent were used as internal references. The mass spectrum was recorded on an Agilent 6530 Quadrupole time-of-flight instrument under the conditions described in the body of the manuscript.

A.



B.

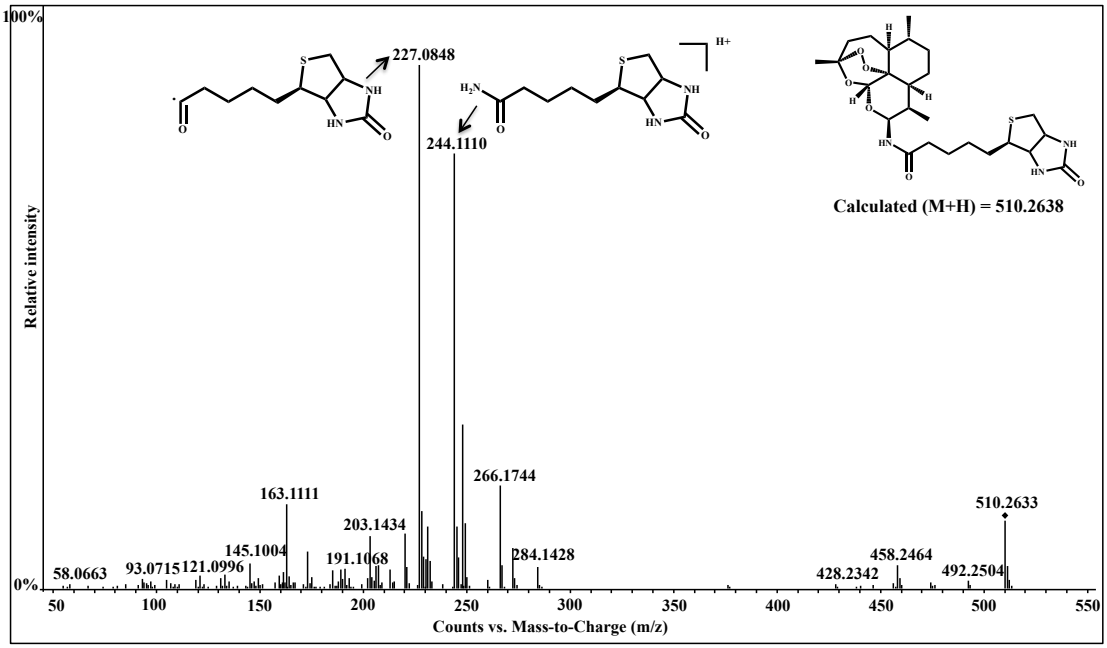


Figure S1: A. ^{1}H NMR, B. HRMS spectrum of BDHA.

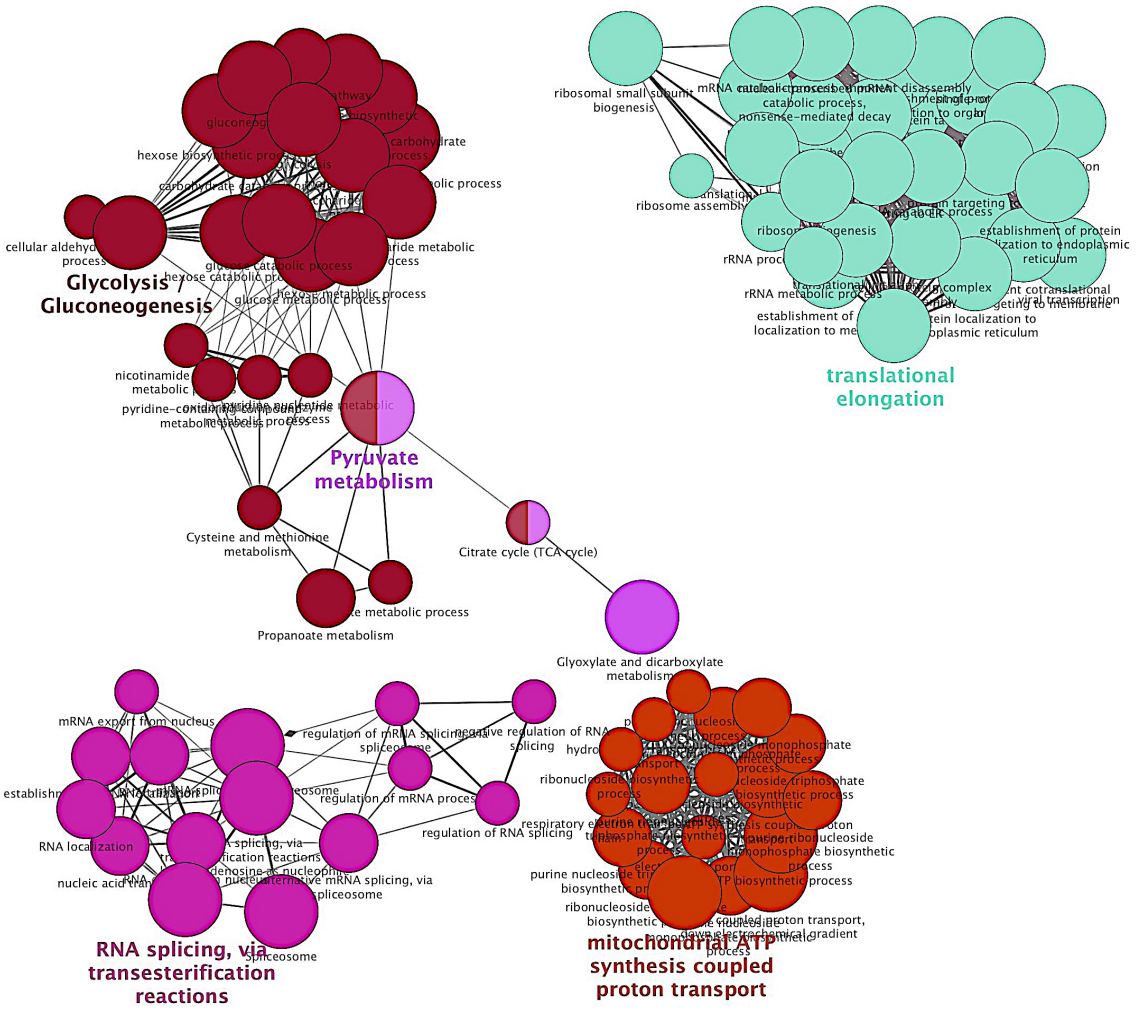


Figure S2: ClueGO analysis of BDHA binding proteins. GO biological process/KEGG pathway terms specific for BDHA target proteins from the BEAS2B cells. The network was sorted according to GO BP criteria. The open source network-visualization package Cytoscape and the Enrichment Map plug-in were used for visualization of the Gene Set Enrichment Analysis results. Each node represents a gene ontology biological process. The node size represents the enrichment significance. GO biological processes were assigned by the ClueGo Cytoscape plug-in.

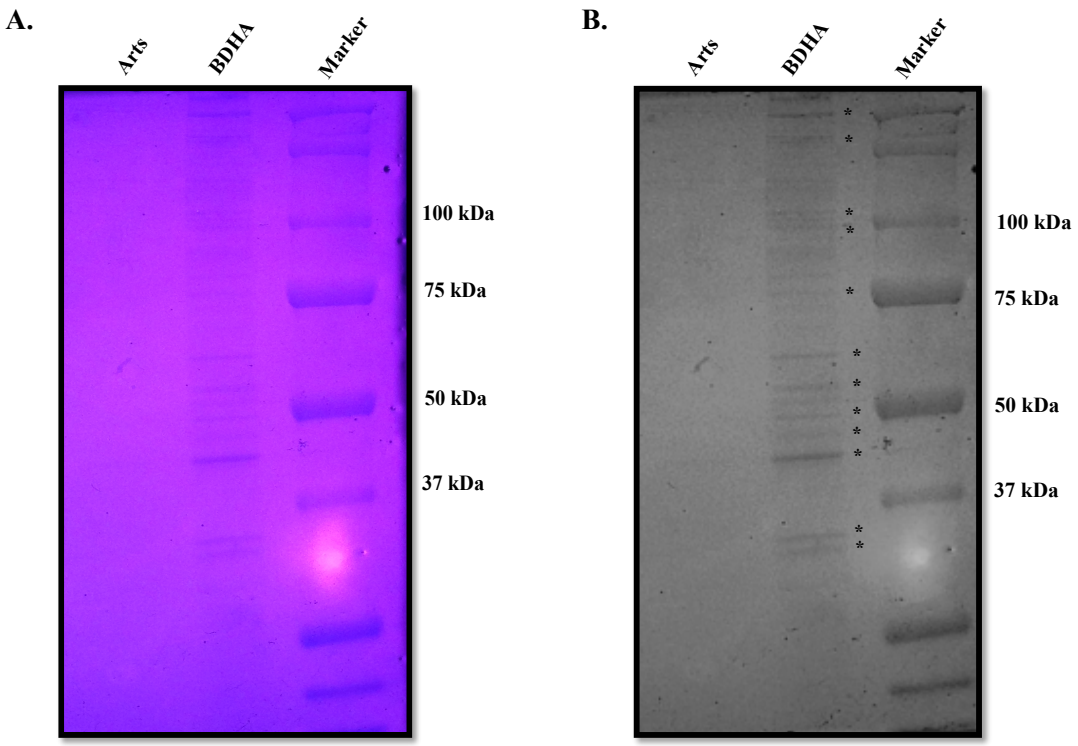


Figure S3: Coomassie stained gel showing proteins captured by BDHA. A total of 12 bands from this gel were processed for In-gel digestion. The identified proteins are listed in supplementary Table S1. A. Original image B. Gray-scale image. * - in the gray-scale image indicates the bands cut for in-gel digestion.