Data for comparative proteomics analysis of the antitumor effect of CIGB-552 peptide in HT-29 colon adenocarcinoma cells

Teresa Núñez de Villavicencio-Díaz [a], Yassel Ramos Gómez [a],
Brizaida Oliva Argüelles [b], Julio R. Fernández Masso [a],
Arielis Rodríguez-Ulloa [a], Yiliam Cruz García [c],
Osmany Guirola-Cruz [a], Yasset Perez-Riverol [d], Luis Javier
González [a], Inés Tiscornia [e], Sabina Victoria [e],
Mariela Bollati-Fogolín [e], Vladimir Besada Pérez [a],
Maribel Guerra Vallespi [b]

[a] Department of Systems Biology, Center for Genetic Engineering and Biotechnology, Cuba
[b] Pharmaceuticals Department, Center for Genetic Engineering and Biotechnology, Cuba
[c] Department of Preclinical Studies, National Institute of Oncology and Radiobiology of Cuba, Cuba
[d] European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust
Genome Campus, Hinxton, Cambridge, UK
[e] Cell Biology Unit, Institut Pasteur of Montevideo, Uruguay

ARTICLE INFO

ABSTRACT

CIGB-552 is a second generation antitumor peptide that displays potent cytotoxicity in lung and colon cancer cells. The nuclear subproteome of HT-29 colon adenocarcinoma cells treated with CIGB-552 peptide was identified and analyzed [1]. This data article provides supporting evidence for the above analysis.

Specifications Table

Subject area	Biology
More specific subject area	Pharmacology, Proteomics
Type of data	Figure, table, methods
How data was acquired	Mass spectrometry: hybrid quadrupole orthogonal acceleration tandem mass spectrometer QTof-2 (Micromass, Manchester, U.K.)
Data format	Analyzed
Experimental factors	Isolation of the Nuclear Proteins Enriched Fraction, trypsin digestion and isotope labeling of peptides
Experimental features	Subcellular fractionation, protein and peptide fractionation by DF-PAGE and LC-MS/MS peptide identification
Data source location	Havana, Cuba
Data accessibility	The data are provided in this article.

Value of the data

- The data details the DF-PAGE separation method used in the proteomics analysis of GIGB-552 peptide effect.
- The data details the bioinformatics-driven approach used for the functional classification of the identified and differentially modulated proteins.
- The data provides an overview of the nuclear proteins differentially modulated in HT-29 colon adenocarcinoma cells treated with the antitumor peptide CIGB-552 and their functional classification.

1. Data, experimental design, materials and methods

We performed a comparative proteomics experiment in duplicate focusing on the quantification of the nuclear subproteome of the human HT-29 colon adenocarcinoma cells via a dual-fractionation by polyacrylamide gel electrophoresis (DF-PAGE) approach. The differentially modulated proteins were functionally analyzed using a systems biology workflow that integrates the information obtained from two main groups of bioinformatics tools [1].

2. Dual fractionation by polyacrylamide gel electrophoresis (DF-PAGE)

The DF-PAGE method combines sequentially, protein fractionation by SDS-PAGE, in-gel tryptic digestion and peptide fractionation by SDS-free PAGE [2,3]. In the first fractionation step (SDS-PAGE), proteins are solubilized in a SDS containing solution and separated according to their size. The presence of SDS ensures the solubilization of virtually all the proteins including highly hydrophobic proteins [2]. After the in-gel tryptic digestion, the peptide mixture is transferred to a second SDS-free gel. In the absence of SDS, peptides migrate according to their charge and size which is orthogonal to the peptide separation in RP-C18 during the LC-MS/MS analysis [2]. For quantitative DF-PAGE, isotope labeling of peptides is introduced with normal- or deuterated N-acetoxysuccinimide just before peptide fractionation by SDS-free PAGE. Fig. 1 shows a SDS-PAGE analysis of soluble and nuclear fractions for CIGB-552 peptide treated and control samples of two independent experiments. Fig. 2 represents a schematic representation of the DF-PAGE method [3]. Finally, Fig. 3 shows same data as Fig. 1 but for the nuclear fractions obtained from control and treated samples.
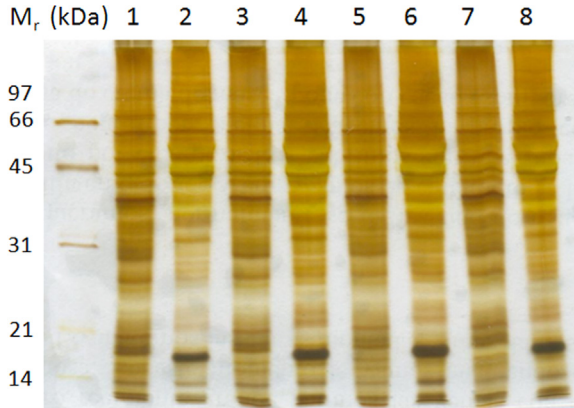
Fig. 1. SDS-PAGE analysis of nuclear and PBS-Triton X-100 soluble fractions for both control and CIGB552-treated samples. Lanes 1,2: PBS-Triton X-100 soluble and nuclear fractions of control sample, experiment1; Lanes 3,4: PBS-Triton X-100 soluble and nuclear fractions of CIGB552 peptide-treated, experiment1; Lanes 5,6: PBS-Triton X-100 soluble and nuclear fractions of control sample, experiment2; Lanes 7,8: PBS-Triton X-100 soluble and nuclear fractions of CIGB552 peptide-treated, experiment2. The 12.5% T gel was silver stained according to standard procedures [12].
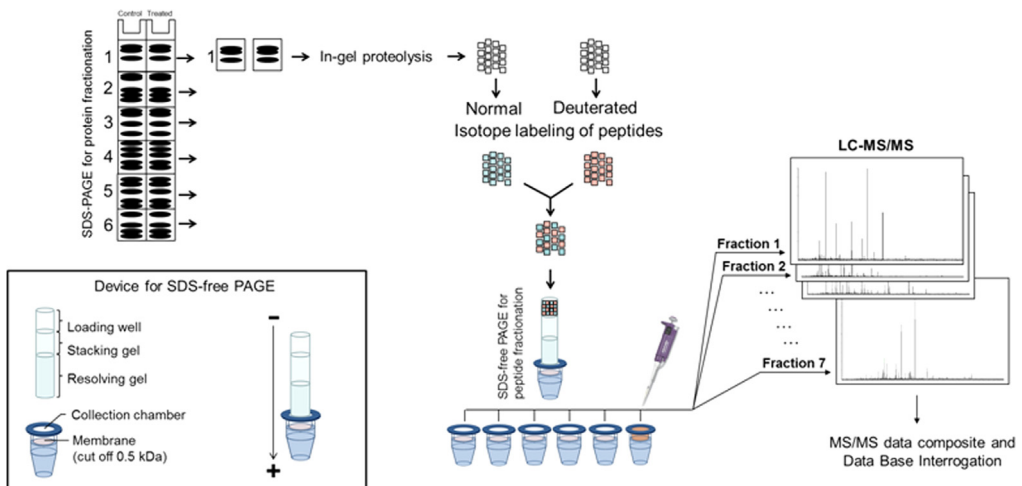

Fig. 2. Schematic representation of DF-PAGE method for quantitative proteomics. Nuclear proteins from both, control and CIGB-552 peptide treated cells were fractionated by SDS-PAGE. Homologous fractions are in-gel digested with trypsin and further isotopically labeled with normal or deuterated N-acetoxy-succinimide. Samples from control and CIGB-552 peptide treated cells are mixed and fractionated by SDS-free PAGE. Collected fractions are then analyzed by LC-MS/MS.

3. Bioinformatics-driven functional interpretation

The functional interpretation of the comparative proteomics data was performed based on two main groups of bioinformatics tools: enrichment analysis and information retrieval and text mining [4]. Since we studied the nuclear subproteome, the identified proteins (Supplementary Table 1) were classified according to subcellular location using information from UniProtKB [5], NextProt [6] and HPA [7] databases and the literature (Supplementary Table 2). Fig. 4 summarizes the procedure used for nuclear protein classification and the results obtained. We considered UniProtKB as the primary resource and searched the literature and/or NextProt and HPA databases when information of nuclear
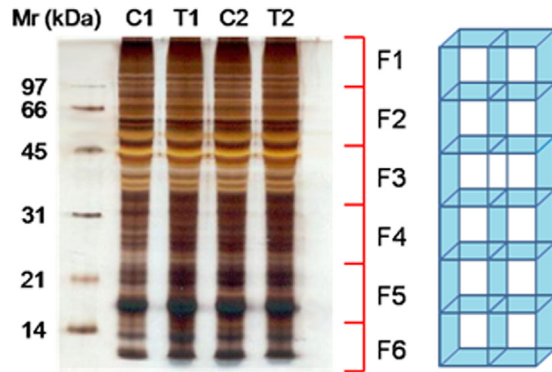
Fig. 3. SDS-PAGE analysis of nuclear fractions obtained from control and CIGB-552 peptide treated cells of two independent experiments. At the right side is shown a schematic representation of the device used for cutting the gel and obtaining the SDS-PAGE fractions. The 12.5% T gel was silver stained according to standard procedures [12]. C1, T1: control and CIGB-552 peptide treated samples from the experiment1. C2, T2: control and CIGB-552 peptide treated samples from the experiment2. The number and protein mass range of the fractions F1-6 are indicated at the right side of the figure.
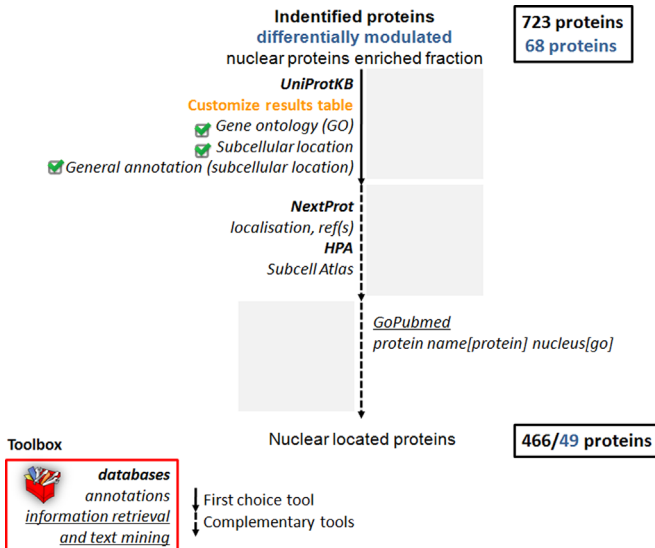

Fig. 4. Schematic representation of the subcellular location classification procedure that shows the number of identified/differentially modulated proteins classified as nuclear located.

localization was missing in the former. Literature information was obtained using GoPubMed [8] text mining tool defining the query as the protein name and nucleus as a GO annotation (Fig. 3).

The differentially modulated proteins were functionally analyzed using function information extracted from UniProtKB, GO biological process and molecular function enrichment analyses (Supplementary Table 3) and the literature. Fig. 5 shows a schematic representation of the followed procedure. Protein function and literature information on protein-disease/biological process relationships helped us to understand the relevance of the differentially modulated proteins. This information was combined with the results of the enrichment with DAVID [9] and GeneCoDis [10] tools to study the proteins not covered by this type of analysis. In addition, Fig. 6 illustrates the use of Chilibot [11] text mining tool to represent in a biological network format protein-biological process and protein–protein functional relationships.
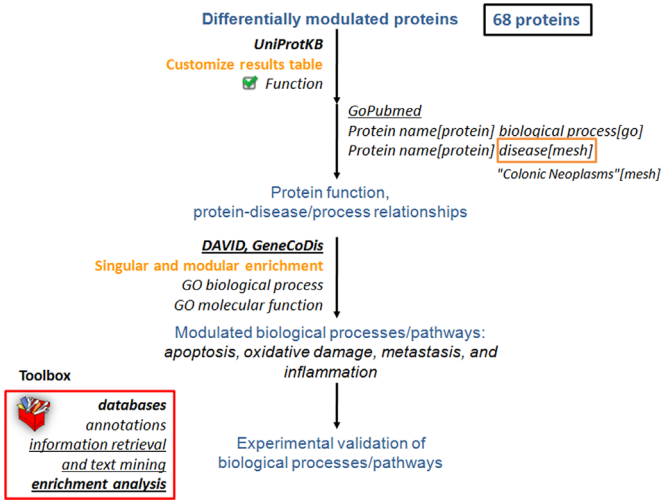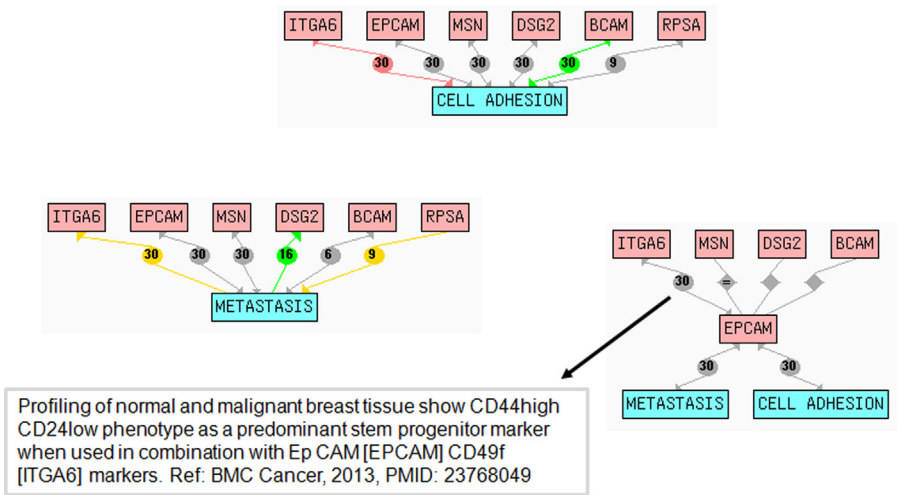
Fig. 5. Schematic representation of the bioinformatics-driven functional interpretation procedure to study the function of the differentially modulated proteins and the results obtained which derives in the experimental validation of specific biological processes and pathways.
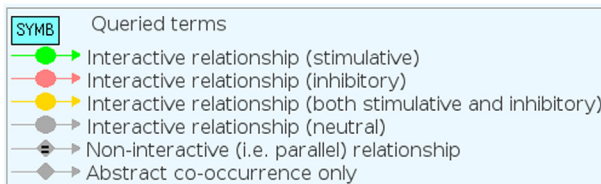

Fig. 6. Chilibot text mining analysis for identifying functional relationships between a set of differentially modulated proteins and between these proteins and cell adhesion and metastasis.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2015.06.024.

References

[1] T. Núñez de Villavicencio-Díaz, Y. Ramos Gómez, B. Oliva Argüelles, J.R. Fernández Masso, A. Rodríguez-Ulloa, Y. Cruz García, et al., Comparative proteomics analysis of the antitumor effect of CIGB-552 peptide in HT-29 colon adenocarcinoma cells, J. Prot. 126 (2015) 163–171, http://dx.doi.org/10.1016/j.jprot.2015.05.024.
[2] Y. Ramos, E. Gutierrez, Y. Machado, A. Sánchez, L. Castellanos-Serra, L.J. González, et al., Proteomics based on peptide fractionation by SDS-free PAGE, J. Prot. Res. 7 (2008) 2427–2434, http://dx.doi.org/10.1021/pr700840y.
[3] Y. Ramos, V. Besada, L. Castellanos-Serra, Peptide fractionation by SDS-free polyacrylamide gel electrophoresis for proteomic analysis via DF-PAGE, Methods Mol. Biol. 869 (2012) 197–204, http://dx.doi.org/10.1007/978-1-61779-821-4_16.
[4] T.N. Villavicencio-Diaz, A. Rodriguez-Ulloa, O. Guirola-Cruz, Y. Perez-Riverol, Bioinformatics tools for the functional interpretation of quantitative proteomics results, Curr. Top. Med. Chem. 14 (2014) 435–449.
[5] E. Boutet, D. Lieberherr, M. Tognolli, M. Schneider, A. Bairoch, UniProtKB/Swiss-Prot, Methods Mol. Biol. 406 (2007) 89–112.
[6] P. Gaudet, G. Argoud-Puy, I. Cusin, P. Duek, O. Evalet, A. Gateau, et al., NeXtProt: organizing protein knowledge in the context of human proteome projects, J. Prot. Res. 12 (2013) 293–298, http://dx.doi.org/10.1021/pr300830v.
[7] F. Pontén, J.M. Schwenk, A. Asplund, P.-HD Edqvist, The Human Protein Atlas as a proteomic resource for biomarker discovery, J. Intern. Med. 270 (2011) 428–446, http://dx.doi.org/10.1111/j.1365-2796.2011.02427.x.
[8] A. Doms, M. Schroeder, GoPubMed: exploring PubMed with the gene ontology, Nucleic Acids Res. 33 (2005) W783–W786, http://dx.doi.org/10.1093/nar/gki470.
[9] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Prot. 4 (2009) 44–57, http://dx.doi.org/10.1038/nprot.2008.211.
[10] D. Tabas-Madrid, R. Nogales-Cadenas, A. Pascual-Montano, GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics, Nucleic Acids Res. 40 (2012) W478–W483, http://dx.doi.org/10.1093/nar/gks402.
[11] H. Chen, B.M. Sharp, Content-rich biological network constructed by mining PubMed abstracts, BMC Bioinform. 5 (2004) 147, http://dx.doi.org/10.1186/1471-2105-5-147.
[12] J. Heukeshoven, R. Dernick, Improved silver staining procedure for fast staining in phastsystem development unit. I. Staining of sodium dodecyl sulfate gels, Electrophoresis 9 (1988) 28–32, http://dx.doi.org/10.1002/elps.1150090106.