Analysis of residue interaction networks to improve prediction of protein succination

Arianna Carolina Rosa¹, Gianluca Miglio¹,²
¹Department of Drug Science and Technology, and ²Scientific Computing Competence Centre (C²S), University of Turin, Italy
E-mail: gianluca.miglio@unito.it

Introduction

Protein succination (Figure 1) results from a specific addition of fumarate to certain cysteine (Cys) residues in proteins. In the last years, the interest toward this type of post-translational Cys modification is markedly increased following the observation of an aberrant level of modified proteins in tumours found in patients affected by hereditary leiomyomatosis and renal cell cancer (Yang et al., Metabolites 2014; 4: 640-54), in tissues of obese and diabetic mice (Thomas et al., Obesity 2012, 20: 263-9), as well as in cells exposed to dimethyl fumarate (Piroli et al., Biochim J 2016; 482: 231-45), which is a drug used to treat patients with multiple sclerosis or psoriasis. Despite the compelling findings, however, further efforts remain to be accomplished to obtain a comprehensive examination and characterization of this type of Cys modification. The aim of this study was to elucidate the role of microenvironment-related factors in governing the specificity of protein succination.

Methods

The analytical procedure is shown in Figure 2. A dataset of fumarate-sensitive proteins and sites was built by collecting data as previously described (Miglio et al., Biochim Biophys Acta 2016; 1864: 211-8). Data were generated/collected from two workflows leading to the: (a) generation of topological data by the analysis of the residue interaction networks (RINs), using together the UCSF Chimera (1.11.2) software (http://www.cgl.ucsf.edu/chimera/) and RINalyzer (http://www.rinalyzer.de), a Cytoscape plug-in for protein structure network assessment; (b) collection of biochemical and biological data from web sources/tools (PropKa (http://nbor.222.ucsd.edu/pdb2pqr 2.0.0)), and DSSP web tool (http://swift.cmbi.ru.nl/cg/dssp/)). Finally, the collected data were analysed and visualized using the R software (The R Project for Statistical Computing; https://www.r-project.org/).

Results

![Figure 2. Overview of the analytical procedure.](image)

![Figure 3. Overview of the included proteins and sites. A total of 278 Cys residues were found in 42 RINs generated from the 41 proteins included in this study. According to their reactivity toward fumarate, 51 and 227 sites were judged as modifiable cysteine (MC) and non-modifiable Cys (NMC) residues, respectively.](image)

![Figure 4. Comparisons between MC and NMC sites. (A) Number of Cys-interacting amino acids (contacts; p = 5.76 × 10⁻¹, Pearson’s χ² test). (B) Accessible surface area of the sulphur atom (p = 8.84 × 10⁻⁶, Wilcoxon rank sum test). (C) Secondary structure of the Cys-bearing peptide (p = 6.51×10⁻³, Pearson’s χ² test). (D) Cysteine-interacting amino acids (p = 2.72 × 10⁻⁸, Pearson’s χ² test).](image)

Conclusions

- Significant differences between MC and NMC sites were determined when their microenvironments were compared.
- The reactivity of a Cys site toward fumarate was accurately predicted when the data for 8 RIN-based topological features were analysed.
- The adoption of concepts of network theory and machine learning could provide helpful strategies to profile a Cys site and quantify its likelihood to be modified by fumarate.

Acknowledgements

Dr Francesca Damiano and Dr Carlotta Corigliano Liso (Department of Drug Science and Technology, University of Turin) are gratefully acknowledged for their support in the method implementation. This work was supported by funding of the "Università degli Studi di Torino, Ricerca Locale Ex-60", 2015 e 2016:17" to GM.