Biochimica et Biophysica Acta xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Intercriteria analysis of calorimetric data of blood serum proteome

Sashka Krumova [a], Svetla Todinova [a], Deyan Mavrov [b], Pencho Marinov [c], Vassia Atanassova [a], Krassimir Atanassov [a], Stefka G. Taneva [a,*]

[a] Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, Acad. "G. Bonchev" Str. Bl. 21 and Bl. 105, 1113 Sofia, Bulgaria
[b] Computer Systems and Technologies Department, "Prof. Dr. Asen Zlatarov" University, 1 "Prof. Yakimov" Blvd., 8010 Burgas, Bulgaria
[c] Institute of Information and Communication Technologies, Acad. G. Bonchev Str., Bl. 25A, 1113 Sofia, Bulgaria

ARTICLE INFO

Article history:
Received 27 April 2016
Received in revised form 19 September 2016
Accepted 13 October 2016
Available online xxxx

Keywords:
Multiple myeloma
Differential scanning calorimetry
Serum protein electrophoresis
Thermodynamic parameters
Intercriteria analysis
Correlation analysis

ABSTRACT

Background: Biological microcalorimetry has entered into a phase where its potential for disease diagnostics is readily recognized. A wide variety of oncological and immunological disorders have been characterized by differential scanning calorimetry (DSC) and characteristic thermodynamic profiles were reported. Now the challenge before DSC is not the experimental data collection but the development of analysis protocols for reliable data stratification/classification and discrimination of disease specific features (calorimetric markers).
Methods: In this work we apply InterCriteria Analysis (ICA) approach combined with Pearson's and Spearman's correlation analysis to a large dataset of calorimetric and biochemical parameters derived for the serum proteome of patients diagnosed with multiple myeloma (MM).
Results: We have identified intercriteria dependences that are general for the various types of MM and thus can be regarded as a characteristic of this largely heterogeneous disease: strong contribution of the monoclonal (M) protein concentration to the excess heat capacity of the immunoglobulins-assigned thermal transition; shift of the albumin assigned calorimetric transition to allocation where it overlaps with the globulins assigned transition and strong shift of the globulins assigned transition temperature attributable to M proteins conformational changes.
Conclusions: Our data justify the applicability of ICA for deciphering of the complex thermodynamic behavior of the MM blood serum proteome.
General significance: The applied approach is suitable for more general application in the analysis of biocalorimetric data since it can help identify the biological relevance of the distinguished thermodynamic features observed for variety of diseases.

1. Introduction

The recent advances in the application of differential scanning calorimetry (DSC) in the studies of the thermally induced changes in the properties of the major plasma proteins in variety of diseases provide solid evidence for its strong potential for disease diagnostics and monitoring. In the last decade it became evident that the plasma/serum DSC profiles (thermograms) of disease individuals are affected not only by the disease-related protein concentration alterations but also by the emergence of new/modified molecular interactions and complexes formation. The complexity of the plasma/serum calorimetric profiles (hundreds of serum proteins giving rise to a multitude of overlapping calorimetric transitions) however obscures their straightforward interpretation. A major hurdle for the direct clinical application of DSC for disease diagnostics is the strict definition of disease-specific thermodynamic features that can only be obtained by the global analysis of extended arrays of data obtained for large variety of diseases. Multi-sample screening DSC instruments that are available nowadays allow for the generation of vast calorimetric databases for different diseases, however at present the development of analytical approaches that allow for the extraction of disease-specific thermodynamic markers characteristic for complex body fluids such as plasma and serum, the prediction of patients survival, and the monitoring of the effect of treatment from experimental databases are not yet fully elaborated. To this point three mathematical routines are introduced for in-depth analysis and stratification of serum/plasma calorimetric profiles: (i) algorithm based on similarity in shape and in distance of a test and a reference thermograms [9]; (ii) classification routine relaying on localized thermogram features and principal components [12] and (iii) thermograms deconvolution analysis for patients classification [24].

In this work, we deal with a database that contains calorimetric (thermodynamic parameters) and electrophoretic (protein concentrations of the major serum proteins) data recorded for 355 patients

* Corresponding author.
E-mail address: sgtaneva@gmail.com (S.G. Taneva).

http://dx.doi.org/10.1016/j.bbagen.2016.10.012
0304-4165/

diagnosed with multiple myeloma (MM). MM is a plasma cell dyscrasia that is either characterized with the secretion of specific serum biomarkers - monoclonal intact immunoglobulins (M proteins), monoclonal immunoglobulin light chains (Bence Jones protein, BJ) or proceeds without the secretion of biomarkers in the blood stream (non-secretory myeloma, NSMM). On genetic and molecular level MM exhibits heterogeneous features and a molecular classification distinguishing 7 groups of MM was proposed by Gutiérrez et al. [14]. Our in-depth DSC studies on MM revealed a large heterogeneity of the calorimetric profiles that posed the necessity of categorizing them on the basis of certain criteria. For this purpose we developed a calorimetric classification based on the isotype of secreted monoclonal immunoglobulins (IgG, IgA, IgM), presence or lack of monoclonal free light chains, number of the successive thermal transitions, melting temperature (T_m) and excess heat capacity of the thermal transitions of albumin (c_P^{HSA}) and immunoglobulins (c_P^{Igs}), weighted average center of the thermogram and shape similarity criterion (details are given in [16,19,21]). For each MM isotype several sets/subsets were determined that all differed (although to different extent) from the control healthy calorimetric set. The observed MM-specific calorimetric changes were interpreted in terms of specific binding of biomarkers (disease related ligands) resulting in stabilization of the globulins and of fraction of albumin to a different degree in the secretory (IgG, IgA, IgM, BJ) and non-secretory MM cases [16,20,21]. It is to be noted that our data also strongly suggested that DSC can be utilized to diagnose and monitor NSMM patients since these cases are characterized with altered DSC profiles even when the clinically utilized biochemical and immunological markers are within the reference limits and diagnosis heavily relies on invasive methods.

To better understand the nature of the observed changes in the calorimetric profiles of MM patients and their relation to the immunological markers that are generally used in the clinical practice (levels of α-, β-, γ-globulins, human serum albumin (HSA), M-proteins and monoclonal free light chains) in this work we generate and explore a database that contains a panel of calorimetric and biochemical data. This database was subjected to InterCriteria Analysis (ICA) and compared with both Pearson's and Spearman's correlation analysis in order to-identify new interdependences between the studied calorimetric and biochemical data. This approach helped us deduce to which extent the different protein fractions affect the different thermal transitions. The presented methodology can well be applied to databases generated for other diseases or containing other types of biological data and thus help to decipher the complex thermodynamic behavior of blood plasma/serum in variety of health conditions.

2. Materials and methods

2.1. Input data

The input dataset is created based on calorimetric and serum protein electrophoresis (SPE) data determined for the multiple myeloma blood proteome of 355 patients, diagnosed on basic clinical criteria with secretory (with secretion of IgG, IgA, IgM and BJ protein) and nonsecretory myeloma in the National Specialized Hospital for Active Treating of Hematological Diseases, Sofia. The experiments were approved by the ethics committee of the National Specialized Hospital for Active Treating of Hematological Diseases, Sofia and informed consent was signed by all patients involved.

2.2. Calorimetric inputs

The ICA algorithm was run over datasets containing either inputs for individual patients' or for patients grouped on the basis of similarities in their calorimetric profiles. The calorimetric scans are categorized in calorimetric sets based on several criteria – type of secreted paraprotein, number of the successive thermal transitions, melting temperatures (T_m) and excess heat capacities of the thermal transitions of albumin (c_P^{HSA}) and immunoglobulins (c_P^{Igs}), weighted average center of the thermogram (T_{FM}) and shape similarity criterion. This yielded calorimetry-based groups of thermograms for IgG, IgA, IgM, BJ MM and NS MM, their mean thermograms and other details are presented elsewhere [16,20,21]. An example of representative median thermograms for different calorimetric sets is presented in Fig. 1.

The following calorimetric input data were used: transition temperature (T_m^{HSA}) and excess heat capacity (c_P^{HSA}) of the albumin assigned thermal transition; transition temperature (T_m^{Igs}) and excess heat capacity (c_P^{Igs}) of the immunoglobulins assigned thermal transition; the ratio of the heat capacities corresponding to the most abundant serum proteins, c_P^{HSA}/c_P^{Igs}; the total enthalpy (ΔH_{cal}) of the calorimetric profile (integrated area under the heat capacity curve); the weighted average center of the thermogram, T_{FM}:

$$T_{FM} = \int_{T_1}^{T_2} T c_P^{ex}(T) dT / \int_{T_1}^{T_2} c_P^{ex}(T) dT \quad (1)$$

where T_1 and T_2 are the initial and final temperatures of the thermogram, respectively [10], c_P^{ex} is the excess heat capacity and T is the temperature; similarity metric parameter ρ [9], which combines two factors: similarities in shape (Pearson's correlation coefficient) and in space (spatial distance metric).

The number, location (T_m) and amplitude (c_P^{ex}) of the successive thermal transitions in the thermograms are all factors that influence both the weighted average center of the thermograms (T_{FM}) and the similarity metric (ρ).

2.3. Serum protein electrophoresis inputs

The serum concentrations of α-, β-, γ- globulins, HSA and M proteins were determined by means of serum protein electrophoresis (SPE) (MINICAP, Sebia, USA). The technique allows for the separation of serum proteins by their size to charge ratio in a capillary filled with an electrolyte and the precise determination of the concentrations of the abovementioned serum proteins [17].

2.4. InterCriteria Analysis

The intercriteria analysis is based on two mathematical formalisms — intuitionistic fuzzy sets (IFSs) [25] and index matrices, proposed and elaborated by Atanassov et al. [3–7] and aims at the

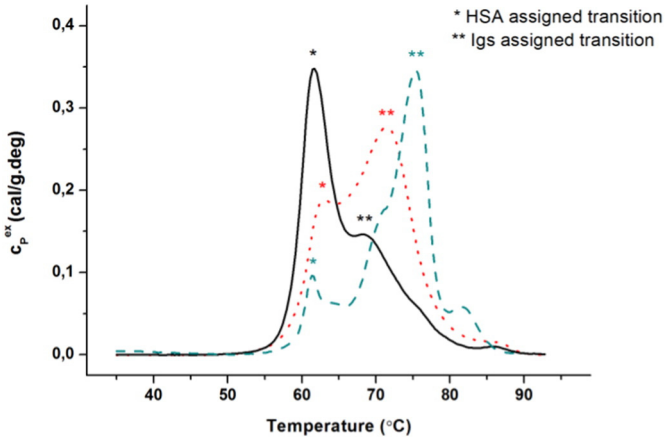

Fig. 1. Examples of mean serum thermograms corresponding to healthy controls (solid line) and calorimetry-based secretory sets of thermograms (BJ protein (dotted line) and IgG (dashed line) multiple myeloma). The thermograms were recorded on highly sensitive DASM 4 calorimeter (Privalov, BioPribor) at scanning rate 1 °C/min. The sera were diluted 8 times in PBS buffer, pH 7.2. For clarity the main transitions of interest (HSA- and Igs-assigned ones) are also denoted.

Please cite this article as: S. Krumova, et al., Intercriteria analysis of calorimetric data of blood serum proteome, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbagen.2016.10.012