Analytica Chimica Acta 916 (2016) 8–16

Contents lists available at ScienceDirect

Analytica Chimica Acta

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

Prediction of retention time in reversed-phase liquid chromatography as a tool for steroid identification

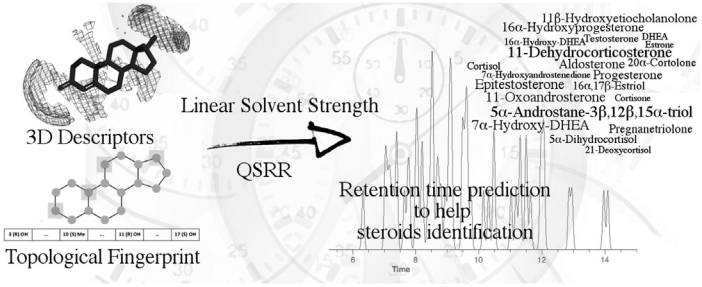
Giuseppe Marco Randazzo [a], David Tonoli [a, b, c], Stephanie Hambye [a], Davy Guillarme [a], Fabienne Jeanneret [a, b, c], Alessandra Nurisso [a], Laura Goracci [d], Julien Boccard [a], Serge Rudaz [a, b, *]

[a] School of Pharmaceutical Sciences, University of Geneva and University of Lausanne, Geneva, Switzerland
[b] Swiss Centre for Applied Human Toxicology (SCAHT), Universities of Basel and Geneva, Basel, Switzerland
[c] Human Protein Sciences Department, University of Geneva, Geneva, Switzerland
[d] Department of Chemistry, Biology and Biotechnology, University of Perugia, Perugia, Italy

HIGHLIGHTS

- Difficulties regarding steroid identification are encountered when considering isotopomeric steroids.
- Quantitative structure retention relationship (QSRR) models were developed from the linear solvent strength theory.
- A dataset composed of 91 steroids and VolSurf + descriptors combined with a new dedicated molecular fingerprint, were used.
- The list of candidates associated with identical monoisotopic masses was reduced, facilitating steroid identification.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 15 December 2015
Received in revised form
14 February 2016
Accepted 16 February 2016
Available online 19 February 2016

Keywords:
Retention time prediction
Steroids
Isotopomers identification
LSS theory
Reversed-phase liquid chromatography
Quantitative structure–retention relationships

ABSTRACT

The untargeted profiling of steroids constitutes a growing research field because of their importance as biomarkers of endocrine disruption. New technologies in analytical chemistry, such as ultra high-pressure liquid chromatography coupled with mass spectrometry (MS), offer the possibility of a fast and sensitive analysis. Nevertheless, difficulties regarding steroid identification are encountered when considering isotopomeric steroids. Thus, the use of retention times is of great help for the unambiguous identification of steroids. In this context, starting from the linear solvent strength (LSS) theory, quantitative structure retention relationship (QSRR) models, based on a dataset composed of 91 endogenous steroids and VolSurf + descriptors combined with a new dedicated molecular fingerprint, were developed to predict retention times of steroid structures in any gradient mode conditions. Satisfactory performance was obtained during nested cross-validation with a predictive ability (Q^2) of 0.92. The generalisation ability of the model was further confirmed by an average error of 4.4% in external

* Corresponding author. School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, Boulevard d'Yvoy 20, 1211 Geneva 4, Switzerland.
E-mail address: serge.rudaz@unige.ch (S. Rudaz).

http://dx.doi.org/10.1016/j.aca.2016.02.014

prediction. This allowed the list of candidates associated with identical monoisotopic masses to be strongly reduced, facilitating definitive steroid identification.

1. Introduction

Steroids are a family of molecules which regulate several vital functions, such as body growth, response to stress, sexual development and behaviour, and rate of metabolism [1]. The endocrine system is responsible for numerous regulatory functions, thus, a steroid imbalance involving an altered hormone production could often be related to various diseases. Several genetic disorders, such as adrenoleukodystrophy, Wolman disease and Smith-Lemli-Opitz syndrome, which are related to cholesterol synthesis and metabolism, have been identified as a cause of endocrine perturbation [2]. Other diseases (*e.g.*, hypertension, cancer, etc.) are associated with an altered regulation of hormone production [3]. Currently, environmental toxicants also constitute a potent source of endocrine disruption. In this context, a large number of drugs and synthetic chemicals have been identified as endocrine disruptors (EDCs) [4] by the Food and Drug Administration (FDA). A large-scale monitoring of steroidogenesis is therefore essential for the screening of EDCs, recommended within the Organization for Economic Co-operation and Development (OECD) guidelines.

Cholesterol can be considered as the initiating molecule of steroidogenesis [5]. From this compound, the different endocrine glands are able to synthesise a vast range of relatively homogeneous chemical structures. Hence, a complex pathway of steroidogenesis generates a wide structural diversity on the gonane skeleton (Fig. 1), dictating the need for selective analytical techniques to separate and identify these structures.

Mass spectrometry coupled with either gas chromatography (GC–MS) or liquid chromatography (LC-MS) today represents the standard detection method for steroids [6]. The main issue of the GC–MS analytical workflow is related to the derivatisation step required to enhance the volatility of the analyte, which limits throughput [7]. Today, untargeted profiling strategies based on liquid chromatography (LC) coupled to high-resolution MS (HRMS) provide an attractive analytical alternative for steroid analysis [8,9]. The resolution offered by HRMS and the peak capacity offered by modern LC instruments allows the simultaneously analysis of thousands of peaks in complex biological matrices, such as urine, blood, plasma, and cellular cultures [8,10–12]. The first step in the identification of features is usually accomplished through mass-based database searches [13]. However, this procedure has severe limitations in the presence of isotopomers [14] having the same exact mass and therefore the same molecular formula [15,16]. A recent work showed that diastereoisomers may be separated in RPLC using achiral stationary phases [17]. Therefore, the chromatographic retention time constitutes an additional molecular property that can be crucial to differentiate between isotopomers [18].

Modelling retention in LC is an active field of research both for theoretical and practical needs. Several theoretical models have already been described and implemented using various optimisation software to help the selection of optimal chromatographic conditions as reported by Héberger [19]. For reversed-phase (RP) retention, the seminal works of Dolan and Snyder [20] have

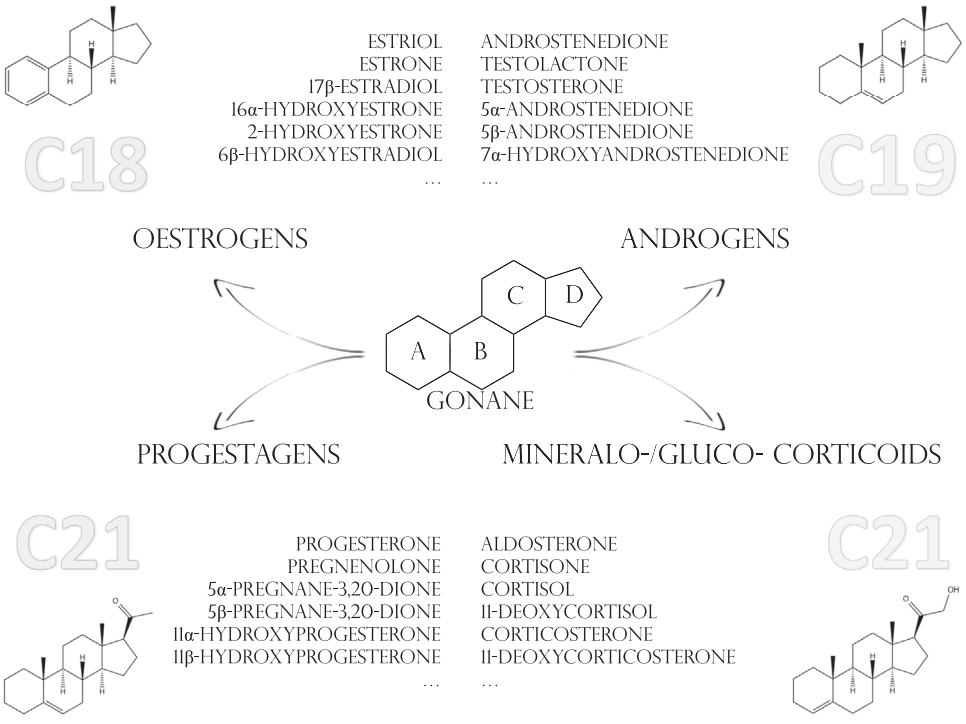


Fig. 1. Classification of steroids based on their precursors. From each scaffold, different steroids are produced by specific reactions.