Phytochemistry Letters 21 (2017) 313–319

Contents lists available at ScienceDirect

Phytochemistry Letters

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

Mini review

Strategies for dereplication of natural compounds using high-resolution tandem mass spectrometry

Tobias Kind[a,*], Oliver Fiehn[a,b,**]

[a] West Coast Metabolomics Center, UC Davis, Davis 95616, CA, USA
[b] Biochemistry Department, King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 15 July 2016
Received in revised form 26 August 2016
Accepted 10 November 2016
Available online 17 November 2016

Keywords:
Natural products
Dereplication
Tandem mass spectra
Seven Golden Rules
MS-FINDER

ABSTRACT

Complete structural elucidation of natural products is commonly performed by nuclear magnetic resonance spectroscopy (NMR), but annotating compounds to most likely structures using high-resolution tandem mass spectrometry is a faster and feasible first step. The CASMI contest 2016 (Critical Assessment of Small Molecule Identification) provided spectra of eighteen compounds for the best manual structure identification in the natural products category. High resolution precursor and tandem mass spectra (MS/MS) were available to characterize the compounds. We used the Seven Golden Rules, Sirius2 and MS-FINDER software for determination of molecular formulas, and then we queried the formulas in different natural product databases including DNP, UNPD, ChemSpider and REAXYS to obtain molecular structures. We used different in-silico fragmentation tools including CFM-ID, CSI:FingerID and MS-FINDER to rank these compounds. Additional neutral losses and product ion peaks were manually investigated. This manual and time consuming approach allowed for the correct dereplication of thirteen of the eighteen natural products.

Contents

1. Introduction

Natural compound identification commonly employs multiple analytical techniques. Especially NMR is highly useful to determine the correct connection table and stereochemistry of compounds (Wolfender et al., 2001). It is however possible to perform a dereplication of known natural products looking up such compounds in natural product or public compound databases (Corley and Durley, 1994) and using only mass spectral information. We here discuss the manual methodology for annotation of known natural products by interpreting and utilizing high resolution tandem mass spectrometry information (Kind and Fiehn, 2010). Our approach included MS/MS search for fast compound annotation and for those compounds that are not covered in MS/MS databases a manual procedure. This involved the determination of molecular formulas, the subsequent query of these molecular formulas in natural product databases and retrieval of compound isomer structures. These were ranked by in-silico fragmentation algorithms and MS/MS spectra and rankings were refined by a manual and time consuming process.

* Corresponding author.
** Corresponding author at: West Coast Metabolomics Center, UC Davis, Davis 95616, CA, USA.
E-mail addresses: tkind@ucdavis.edu (T. Kind), ofiehn@ucdavis.edu (O. Fiehn).

http://dx.doi.org/10.1016/j.phytol.2016.11.006
1874-3900/

2. Materials and methods

The challenges for the CASMI 2016 Category-1 contest are natural products from several organisms of different possible origin (plants, fungi, marine sponges, algae or micro-algae) acquired on QToF instruments from Waters and Agilent (see Fig. 1). Based on the MS and MS/MS, the goal was to determine the correct molecular core structure (without any stereo information) at the given retention time using the spectral data and the additional information provided. The contest website lists detailed results and participant lists (http://www.casmi-contest.org/2016/results-cat1.shtml). The submitted structures are then ranked according to the absolute rank of the correct solution and for tied scores the average of tied ranks is used.

Molecular formulas were determined with the Seven Golden Rules (Kind and Fiehn, 2007) and Sirius[2] (Böcker et al., 2009). In some cases the CASMI provided data was not sufficient and MS^1 and MS^2 data were extracted from the raw files using ProteoWizard

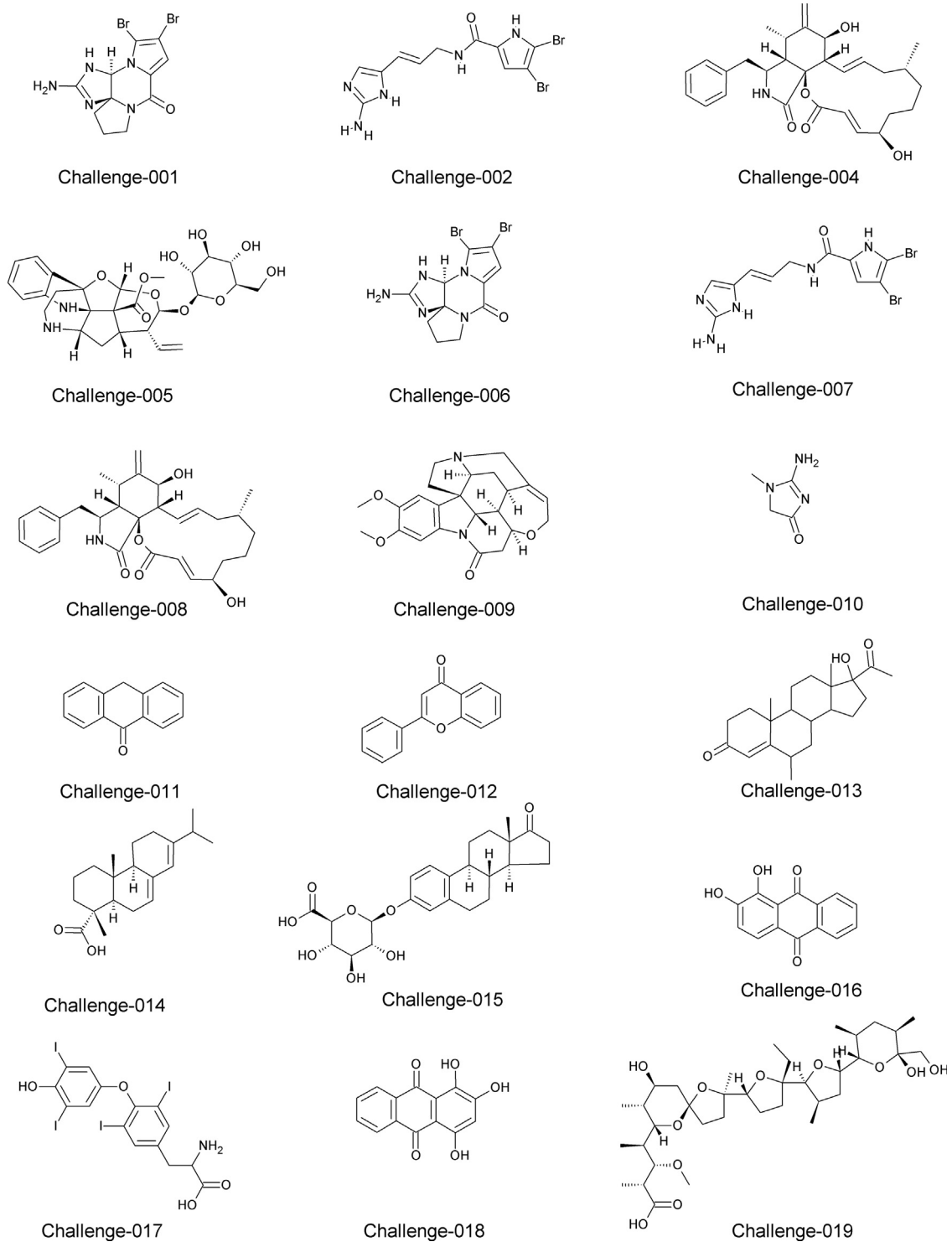

Fig. 1. The CASMI (Critical Assessment of Small Molecule Identification) 2016 contest in category one (natural products) provided the accurate MS and MS/MS spectra of eighteen compounds. Based on the mass spectral information and other metadata correct structural information needed to be assigned.