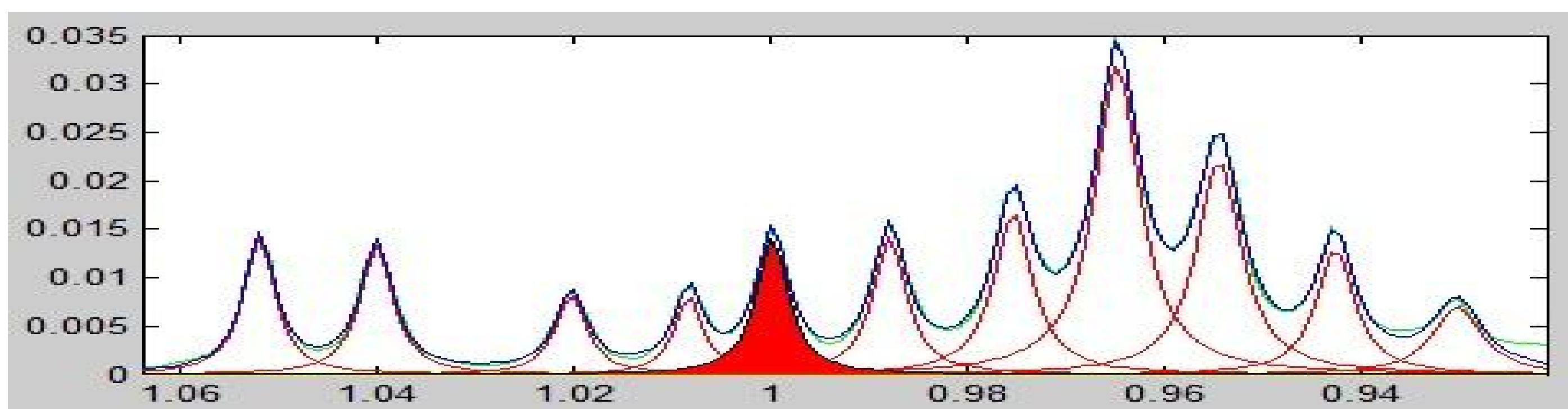


INTRODUCTION

- The field of *Metabolomics* studies small molecules, called metabolites, which are present in cells and extracellular fluids of biological organisms. Metabolomics is increasingly applied to the areas of system biology, drug discovery, early disease detection and pharmaceutical research.
- Nuclear Magnetic Resonance (NMR) spectroscopy has been established as one of the most popular tools for high-throughput characterization of metabolites present in complex biological mixtures, because it is a non-destructive, highly reproducible and versatile technique, since it allows the user to “interrogate” the same sample in different ways by selecting different pulses and acquisition parameters in order to obtain complementary information.
- Each metabolite may contribute to the NMR spectra with many individual signals (singlets and multiplets) and the peaks may seriously overlap with those from other metabolites.
- In summary, a correct identification and quantification for a set of compounds of interest (targeted profiling) can be challenging, particularly in high peak congested regions.

BACKGROUND

- *Curve fitting without a database:* NMR spectra can be viewed as a combination of signals from individual metabolites and the peaks derived from those signals can be modeled by a linear combination of Lorentzian-Gaussian lineshape functions. The problem is the lack of automatic assignment of fixed proportions of multiplets that often makes unpractical the correct identification and quantification of metabolites.



- Signals overlapped region fitted as a linear combination of Lorentzian-Gaussian line-shape functions. The lack of fixed proportions between signals usually drifts in identification and quantification errors.
- *Curve fitting with database:* Interpretation of NMR spectra as linear combinations of reference profiles from database allows the user to simultaneously identify and quantify metabolites. Automation of this strategy doesn't work successfully in regions where unknowns signals are present and difference between reference profiles and experimental data due to matrix influence in chemical shift position makes the practice of this process be viable only in a manual way.

- Valine profiling in Chenomx NMR Suite Profiler in a congested region. Manual assignments of the signals present choosing between a large list of pure compounds requires high level of user expertise.

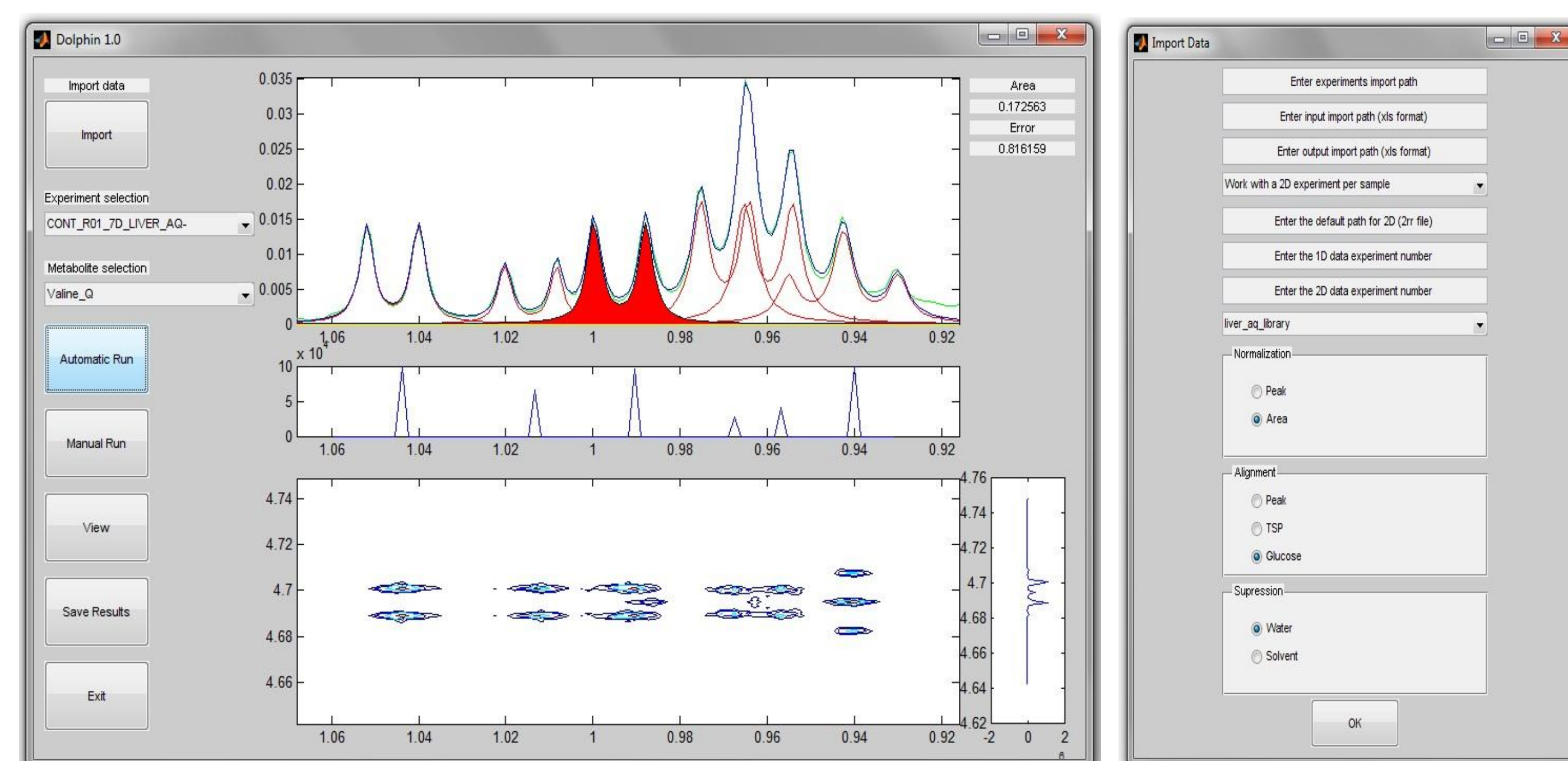
CONCLUSIONS

- Dolphin's strategy combines both 1D and 2D data in order to work with more accurate information to match and fit signals, improving the automation in both identification and quantification processes
- Results shows a really high correlation between quantifications obtained in a manual way using Chenomx as the gold standard, where the user has to check lots of combinations, choosing signals in a large library of compounds and adjusting intensities trying to fit the real spectra, in a time consuming and difficultly reproducible process, while with Dolphin time is severely reduced in a high reproducible process.

OBJECTIVES

- Design a strategy to improve automatic signals identification and quantification in peak congested regions combining 1-dimension filtered data and 2-dimension dispersed data.
- Improve current fitting methods in a realistic constrained total-line-shape fitting, based in real signals present in the samples instead of superposition of pure compounds from databases, allowing the detection of unknown signals.
- Implement this strategy in the matlab environment under a user friendly graphical interface.
- Test the method comparing results with Chenomx NMR Suite 6.1 as the gold standard for 2 samples datasets, one corresponding to a pull of standard compounds (STN) and another corresponding to liver-aqueous extracts from rats (EXP).

RESULTS



- Dolphin's running panel where the user can select the experiments and metabolites to check and watch how the fitting has been done (left). Importing data panel where the user can choose the paths and parameters for data pre-processing (right).

- Radial graphs of the R2 values for all targeted metabolites present in STN (left) and EXP (right). Both figures shows a really good correlation values between Dolphin automated method and Chenomx manual method.