Replacing Pen and Paper in the Analytical Laboratory

A Database for analytical chemistry

Jürgen Bienert
Max-Planck-Institute for Biophysical Chemistry; Facility for Synthetic Organic Chemisty
Göttingen, Germany

Introduction

At the Max-Planck Institute for Biophysical Chemistry, the main task for the Facility for Synthetic Organic Chemistry is the synthesis and subsequent analysis of various organic compounds that are not commercially available. This allows other departments to spend more time focused on their research, rather than spending time on these non-trivial tasks.

By using state-of-the-art technology, the facility is able to handle all kinds of requests from start to finish: preparation and literature research, classical or microwave synthesis of the desired compounds, and the application of different analytical techniques to ensure quality. The facility uses many different analytical techniques, including: Chromatography (DC, Flash), NMR, MS (ESI, GC-MS, EI or CI), FT-IR, UV/Vis, Fluorescence, GC and HPLC. This wide array of techniques allows syntheses in an efficient and timesaving manner.

During the whole process – from the initial synthesis request to the finished compound – all kinds of different data is being used, generated and stored: literature, safety data sheets, chromatograms, spectra, images, reports, and our own notes are among them. But a key question is, what happens with this data? The classical method is of course to write everything down in your paper lab journal and to print your chromatograms and spectra and paste them in:

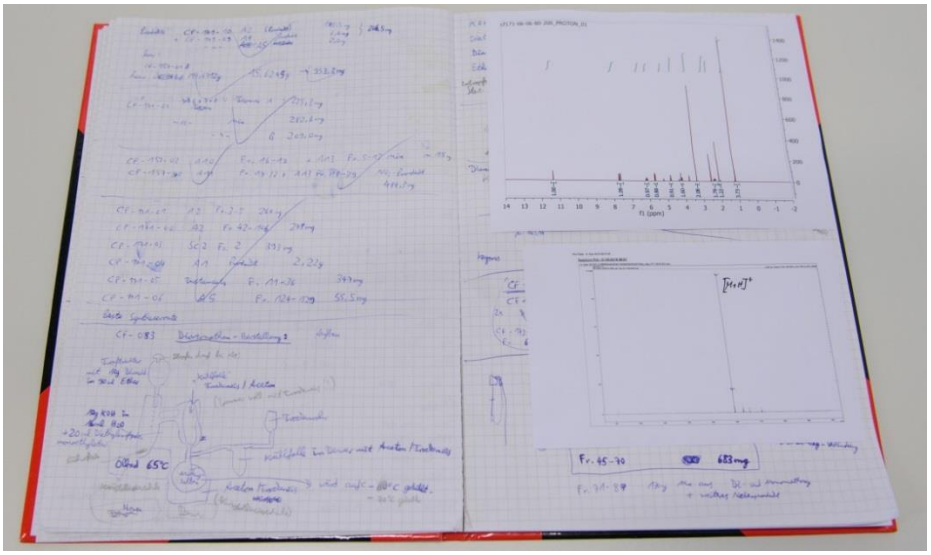


Figure 1
Over 70% of all industrial laboratories and up to 96% of all academic or governmental laboratories are still utilizing paper-based documentation.[1]

Paper notebooks as data graveyards

Many scientists have experienced the frustration of not knowing if the problem they are currently struggling with in the laboratory was already solved by someone else. Additionally, if experiments need to be completely repeated because the handwriting of a colleague, who has since left, could not be deciphered, one quickly realizes: **the paper notebook has served its time.**

In any industry, the use of paper notebooks creates a substantial problem. Data that has been reported on paper only, cannot be easily searched or shared with peers. Because only a small fraction of that data will be reprocessed for publication, the vast majority will disappear in the laboratory's mountain of files, where it is not even of use for the scientists who originally created that data. This indirect loss of data is one of the biggest infrastructural problems in science, and sometimes known as the "File Drawer Effect".

What can be done to overcome the File Drawer Effect?

Digital archiving systems like Electronic Lab Notebooks (ELNs) offer one solution: scientists are now able to use a computer keyboard and mouse instead of a pen to write into their lab journal. An ELN is very often used by more than just one scientist – thus ELNs can control restrictions for data creation and modification, along with passwords or digital signatures to keep track of all changes made by different individuals. Elaborate search functions then enable the ELN users to quickly search and find all kinds of data that has been saved digitally: experimental parameters, instrument settings, actual data such as chromatograms or spectra, metadata like text or tables, and even pictures.

However, commercially available ELNs are often relatively inflexible and offer only minimal user customization.

Every laboratory works in its own way, and many use instruments from different vendors, thereby creating data in different data formats.[2] Therefore the ideal software can handle a lot of data formats, and offers detailed user customization in order to create group- or department-specific interfaces, focusing on that particular group's workflow.

Figure 2
With so many different analytical techniques utilized by a modern synthetic chemistry laboratory, how does an ELN-type software bring all the data together in a unified manner, while allowing customization by the user?

To address these issues in the Facility for Synthetic Organic Chemistry, for many years we have utilized the ACD/Spectrus Platform (ACD/Spectrus Processor and Workbooks). With this software we are able to access all of our data via an interface that we created and optimized by ourselves.

Figure 3

The ACD/Spectrus Platform is designed to be a vendor-neutral analytical software environment, importing and processing data from most major vendors in a uniform manner.

All data, reaction schemes, spectral and/or chromatographic data, and metadata, is collected and merged in databases that can be accessed directly via the ACD/Labs software.

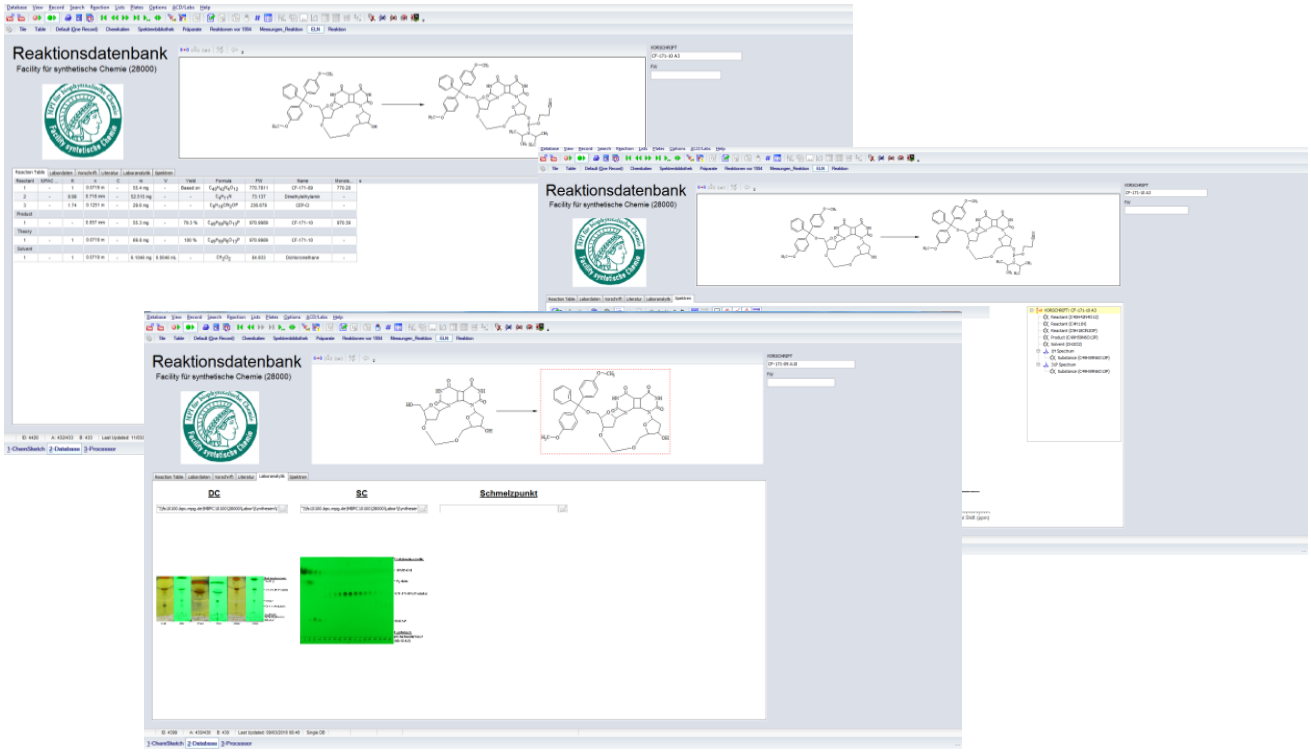


Figure 3: The Reactions database of the Facility for Synthetic Organic Chemistry.

Reprocessing of saved datasets is also possible directly from the respective database: spectral or chromatographic data can be processed again or modified with ACD/Spectrus Processor, while other information like an embedded text document or a spreadsheet table can be opened with their respective editors.

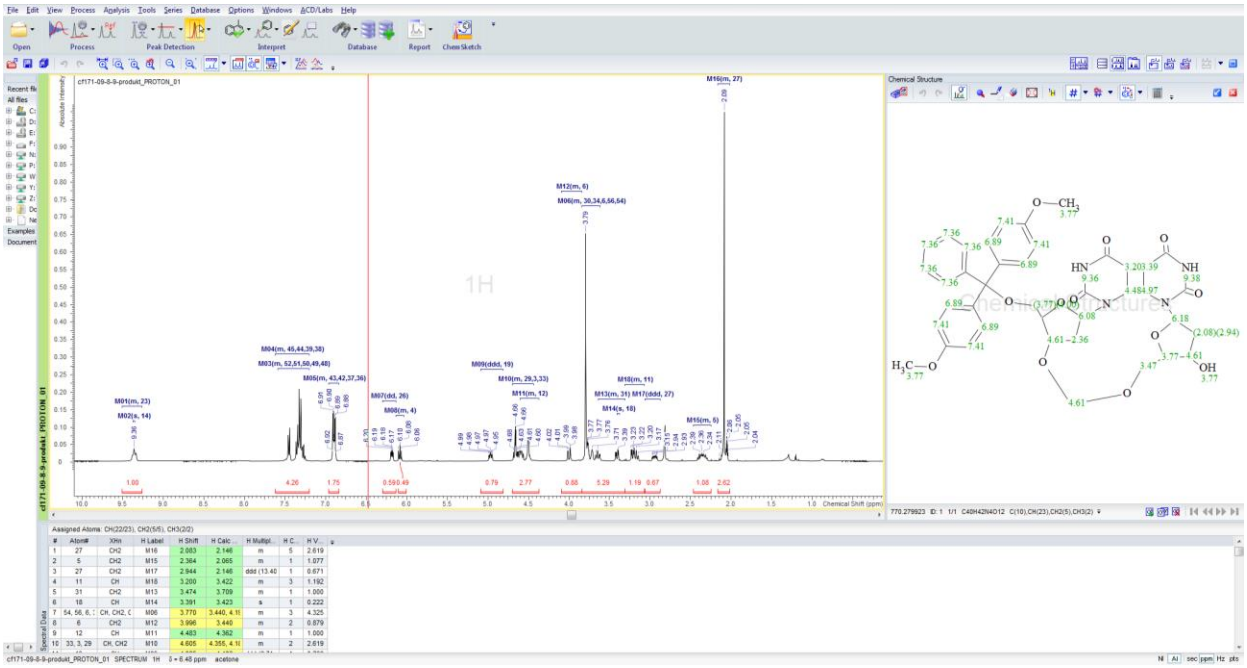


Figure 4: ACD/Spectrus Processor: Interpretation of an NMR Spectrum.

With the ACD/Spectrus Platform we have a powerful tool for processing and interpreting analytical data (NMR, GC/MS, LC/UV/MS, IR, Raman, X-Y Data and more) from different vendor instruments in a homogeneous interface. Database searches, even across multiple databases at the same time, are usually performed by searching for a structure, a spectrum, or a CAS number; a multi-parameter search or setting of specific search filters is also possible.

Exchange of data across departments is possible by using Spectrus Processor to browse and search the database(s). In our case, our colleagues from the different scientific groups who submit syntheses and analytics jobs are able to immediately access all the data that is generated for their specific job. At the same time, they would also have the option to directly process that data using the same software. In this way, the use of different software tools for different analytical techniques is an issue of the past.

References:

1. Dr. Florian Hauer. “Laboratory Book 2.0: Digital and Networked”, GIT-Labor – Portal für Anwender in Wissenschaft und Industrie, http://www.git-labor.de/forschung/informationstechnologie-it/laborbuch-20-digital-und-vernetzt **2013**.
2. The Pains of Analytical Data Management in R&D, Advanced Chemistry Development, Inc., Toronto, ON, Canada, http://www.acdlabs.com/surveyresults/ **2015**.

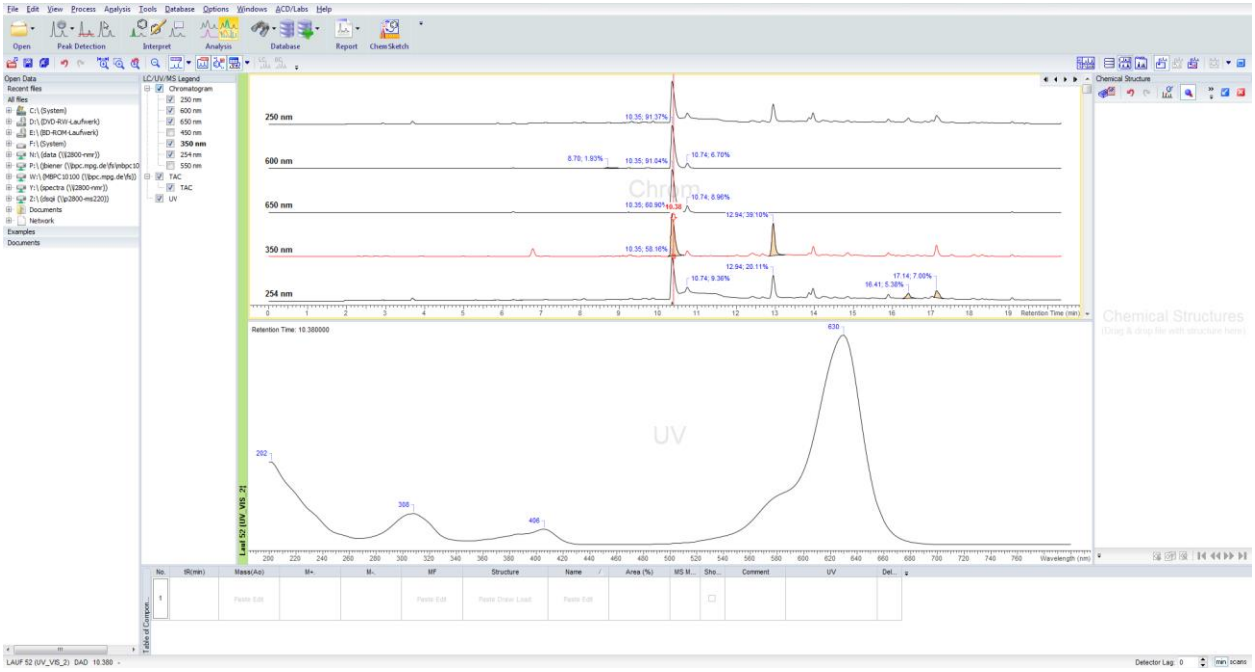
Appendix

Below are examples of different techniques handled in ACD/Spectrus Processor:

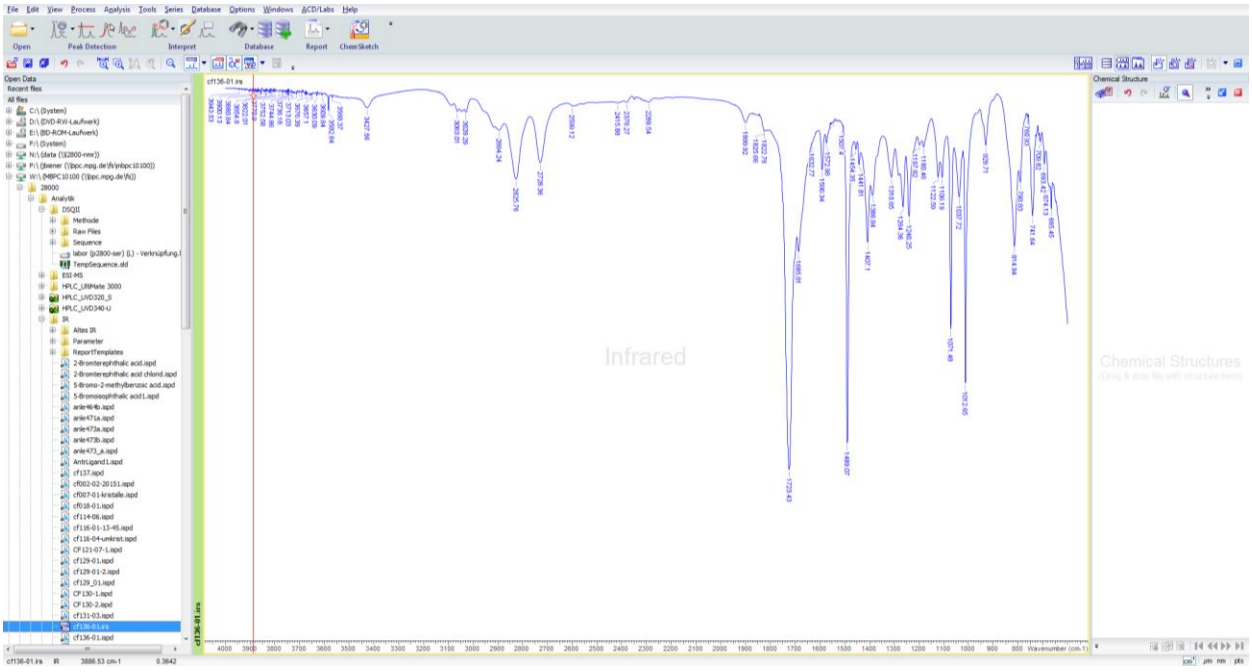
ESI-MS:

GC/MS:

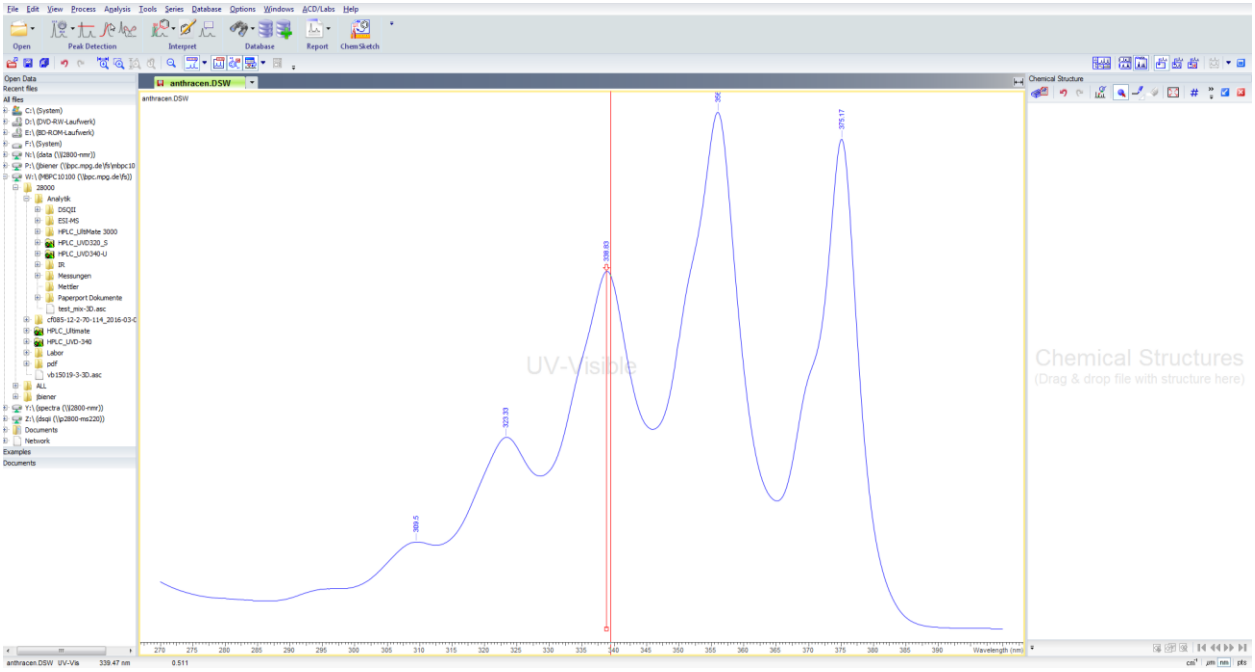
HPLC:


250 nm
600 nm
650 nm
350 nm
254 nm
UV
Chemical Structures

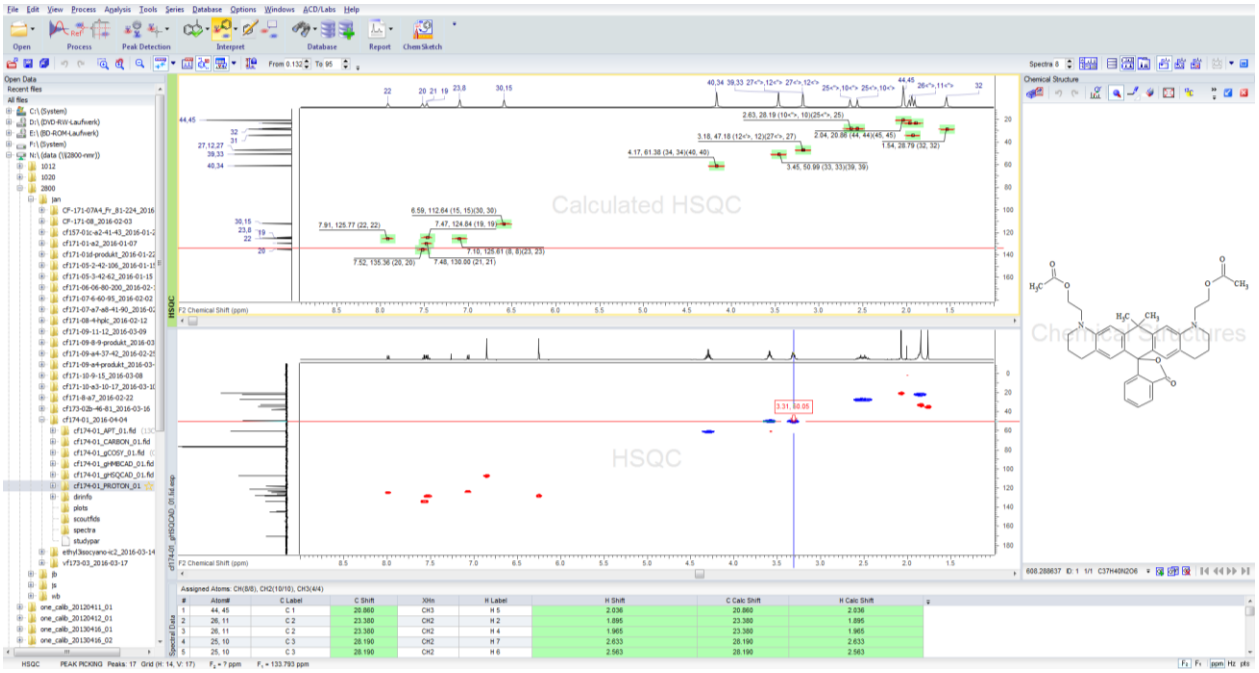
FT-IR:


Infrared
Chemical Structures

UV:


anthracen.DSW
UV-Visible
Chemical Structures
(Drag & drop file with structure here)

NMR:


Calculated HSQC
HSQC
Chemical Structures