Selecting a chemical registration system

A look at out-of-the-box and custom systems, in theory and in practice at Allergan and Aventis

Chemical registration is at the core of the discovery process. All chemical and pharmaceutical research organizations require that compounds be documented, laying the foundation for integrating diverse information on chemical structures with test results to support analysis and eventual lead selection.

But despite the centrality of the registration process, there are multiple ways to approach it and endless variation in detail. The most common categories of registration systems in use today are out-of-the-box products and custom-built solutions, which can be developed in-house or provided by external consultants.

Factors such as organizational complexity, the availability of IT resources, and strategic priorities often dictate an approach. This article examines key factors for selecting between the types of systems and presents case studies of recent implementations by Allergan and Aventis.

Functionality: what constitutes a registration system?

The basic registration process consists of a novelty check on a structure and the assignment of a unique identifier, as well as the recording of associated data such as chemist or supplier name, lab notebook number, and date.

Most companies now also require the registration of each new batch of a compound so that biological data can be associated with specific batches instead of the compound itself. On a structural level, registration entails checking and normalizing structures, checking for novelty, assigning corporate identifiers, and data validation.

> *"We don't have a big-pharma research IT staff to support us... so if an out-of-the-box solution fits our needs we will use it."*
>
> *Robert Cain, senior scientist, Allergan*

In the past, dedicated registrars managed centralized systems, approving data before it was entered. With decentralized systems now allowing scientists to register data as they generate it, tasks such as enforcing corporate drawing conventions are accomplished by applying structure representation business rules in the registration system.

Within the registration process, organizations must set policy on a number of specific issues: Will the database be organized by parent structures or with salts and addends? How will mixtures, partially defined structures, and stereochemistry be handled? (For a white paper on MDL's Enhanced Stereochemical Representation, visit **www.mdl.com**.) Will the system assign identifiers using a generic or "intelligent" system? Will classes of data (traditionally synthesized compounds vs. combinatorial libraries) be segregated? How will system access, security, and data review be managed?

A company's workflow and business rules will influence the answers to these questions, and each of the three basic types of systems has varying degrees of flexibility to accommodate company-specific needs. But often broader considerations—organizational complexity, corporate culture and history, IT resources—dictate whether a company adopts an out-of-the-box or custom solution, either built in-house or by external consultants.

Decision points: finding the right system

A global organization trying to integrate disparate legacy systems and add extensive custom functionality is unlikely to be satisfied with a standard solution. Conversely, a start-up biotech with well-defined research parameters might have neither the specialized needs to warrant a custom-built system nor the resources to build and maintain one.

What specific criteria can be applied to determine the right system? The three approaches generally taken in industry today (buy a standard application, build in-house, or build a custom system using external expertise) are

(continued on page 10)

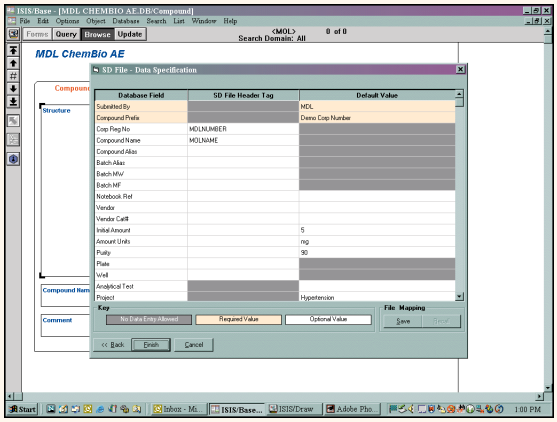

Figure 1: MDL ChemBio AE is an out-of-the-box storage system for registering single compounds or entire libraries. The system automatically checks for duplicates, strips salts and solvents, assigns ID numbers, and calculates chemical properties.

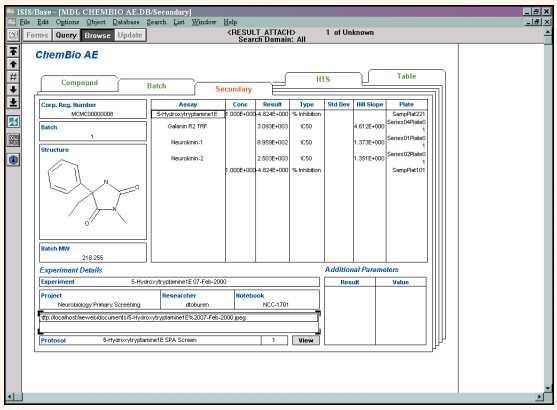

Figure 2: MDL ChemBio AE allows scientists to retrieve and view chemical structures and information alongside biological data from high-throughput and secondary screening.

(continued from page 9)

typically differentiated by five key factors: acceptance, time, flexibility, cost, and support.

- **User acceptance:** Are the scientists convinced by the functionality? Do they see a benefit for themselves in the workflow? Does the system bring workflow improvements?
- **Time to deployment:** How long will it take to become productive with the system?
- **Flexibility:** Can the system accommodate changing needs? Does it integrate with existing systems?
- **Costs:** What is the return on investment? Does it lower the total cost of ownership?
- **Maintenance/support:** Is support available for training, user advice, and enhancement?

Table 1 evaluates the three basic approaches using the above factors, based on tendencies observed by MDL consultants and account representatives during projects and in discussions with customers. Each system entails inherent trade-offs; an informed decision on a registration system involves both R&D and IT management in a frank analysis of strategic priorities and corporate resources.

Factor	Out-of-the-box	Custom Internal IT	Custom MDL Consulting
User acceptance	++	+++	++++
Time	++++	+	+++
Flexibility	+	++++	++++
Costs	++++	++	++
Maintenance/support	++++	++	+++

Table 1 assigns rankings of one through four (with four representing the most favorable ranking) to key factors that differentiate between the basic approaches to chemical registration systems.

Notes:

User acceptance—In-house builds can suffer from poor communication between science and informatics. Since consultants speak the language of the scientist, MDL-supported projects generally meet with high acceptance. Out-of-the-box solutions are occasionally accepted only with the caveat that modifications will be made—though in cases where an organization has a well-defined research focus or the system represents a significant improvement over an existing registration process, acceptance can be high.

Time—Provided no customization is undertaken, an out-of-the-box solution can rapidly be made operational. In-house builds do not necessarily take longer than 3rd-party consultant projects to become productive, but factors such as lack of specialized experience, lack of training support, and changing informatics priorities can impede rapid productivity.

Flexibility—Custom systems are inherently more flexible. However, inexperience in building systems can hamper the flexibility through errors in architecture and build. Experience and broad understanding of the technology are critical to ensuring a cost effective and flexible system.

Cost—Out-of-the-box systems with no customization are by far the most inexpensive solutions. However, cost benefit analysis for a custom solution will frequently reveal a higher return even though the initial outlay is higher. Improvements in process through custom process design can drive significant productivity improvements.

Maintenance/support—Maintenance and support are critical to moving scientists rapidly up the learning curve and maintaining their proficiency with a system. In-house builds often rely on a cadre of power users to spread the word, a process that can require scientists to dedicate time to training tasks. Professional external organizations offer training and management support to facilitate process and organizational changes.

Out-of-the-box case study: Allergan

Allergan Inc. (NYSE: AGN), with headquarters in Irvine, California, is a specialty pharmaceutical company that develops innovative products for the eye care, neuromodulator, skin care, and other markets.

Though known for Botox (botulinum toxin is being tested with encouraging results as a treatment for ailments ranging from stroke paralysis to migraine headaches and carpal tunnel syndrome, according to the *New York Times)*, Allergan also produces small molecule pharmaceuticals. The company recently decided it was time to update the registration system used by its small molecule research division.

"We weren't catching much batch level information," said Robert Cain, senior scientist, Spectroscopy, for Allergan. "And our data model wasn't extensible."

According to Cain, the company considered pursuing a better data model and incorporating additional features, but realized that it was going to be difficult to make the existing in-house system into a professional level system.

"We didn't have the resources in house," Cain recalled. "We aren't big-pharma. We are a mid-sized, specialty pharmaceutical company and thus we don't have a big-pharma research IT staff to support us. We try not to develop things that are already developed, so if an out-of-the box solution meets ours needs we will use it."

Allergan had several requirements for an off-the-shelf registration system in addition to a better data model. The system had to be extensible. It had to integrate well with other applications, including a 3rd party system for biological information. And to replace the quality control function of registrars in the existing centralized system, it had to offer built-in validation tools for incorporating business rules.

In the end, Allergan selected MDL® ChemBio AE as the best fit between its needs and resources (see Figures 1 and 2).

"We went with ChemBio because it was an MDL application and we knew that they were a reliable software partner," said Cain. "MDL helped us design small, easy-to-maintain customizations that we needed. These will allow us to maintain support and lower the amount of in-house development we need to do."

MDL ChemBio AE (*http://www.mdl.com/products/chembio.html*) is an out-of-the-box registration and storage system for chemical data that lets scientists retrieve and view chemical and biological data side by side. Scientists can register single compounds or entire libraries while ChemBio AE automatically checks for duplicates, strips salts and solvents, assigns ID numbers, and calculates chemical properties.

MDL ChemBio AE combines easily with MDL® Assay Explorer and MDL® Report Manager to offer a complete solution for managing and reporting chemical and biological information.

"On the West Coast I see a lot of small or mid-sized companies, particularly biotechs, that need a flexible, extensible registration system that can be implemented with a minimum of internal resources," said Bob Olszewski, the western regional sales manager for MDL in North America. "Because ChemBio AE integrates with other applications and has the potential to grow with companies, it's the right solution for many discovery research organizations."

In Allergan's case, the flexibility of ChemBio AE was an important factor since it had to integrate with their existing, 3rd-party system for biological information—something it has done well.

ChemBio AE was rolled out to approximately 40 scientists at Allergan's Irvine location.

Because it offers clear advantages over the previous system, scientists adopted the new technology and found it to be a considerable improvement.

"It's saving time for the synthetic chemists to register compounds and get a corporate number—and avoiding mistakes," Cain says. "And the support has been very good since we've moved to ChemBio AE."

According to Cain, Allergan is already looking to improve their workflow by tying spectroscopic data into the system. "When people register compounds they'll register identifiers for their spectral data," Cain said. "It's just in the planning stage at this point, but we think we know how it will be done."

Custom solution case study: Aventis

In a global business climate where mergers and acquisitions are commonplace, research organizations often face the challenge of consolidating disparate information systems.

On the bright side, mergers and other large-scale transformations present opportunities to replace outdated or inefficient systems, and implement best practice business rules and standard operational procedures.

When Hoechst Marion Roussel and Rhône-Poulenc Rorer merged to form Aventis (NYSE, Paris, FRA: AVE), the resulting company had scientists at research labs in three countries, speaking different languages and running multiple systems.

"We realized very early that we had to deal with many sites, many chemical information systems, and many different processes for registering chemical data," said Dr. Ludwig Franzisket, global project manager for Aventis Pharma. "We saw the need to consolidate this—to define a common process for chemical registration at all sites."

With mergers, adopting one of the existing systems is not always practical or politically desirable. In Aventis' case, the range of chemistry techniques employed in various labs and the desire for additional functionality beyond chemical registration required management to start from scratch.

"How we started this project was not to think about an IT system, but to define the process we have in the labs," said Dr. Franzisket. "We looked at what are the similar steps and what is different."

Beyond registering chemical batches and linking 5,000 scientists and support staff at multiple locations on two continents, the system had to provide a repository for structural data that was linked to test results, in order to support structure-activity relationship analysis. The long-term goal was to enable an electronic notebook capability to capture all chemically relevant experimental information.

To design and build the new Aventis Registration Process (ARP) system, Aventis turned to MDL Consulting.

Charles Buse was the project leader of the MDL consulting team. "The primary technical challenge we faced was finding a database structure flexible enough to handle classical synthetic experiments, parallel synthesis, and combichem," said Buse.

To meet these challenges and connect with all of the Aventis labs on the wide area network, MDL designed a hybrid system with a distributed architecture that is between a client-server and a true multi-tier system. It consists of a thick client with some services running on NT servers, and Oracle® databases on UNIX platforms. The ARP system is built on a framework of MDL® ISIS (Integrated Scientific Information System) and incorporates MDL® Chemistry Rules Interface to automate chemical structural analysis and manipulations.

Now chemists at every Aventis lab can register synthesized, isolated, or acquired physical batches into a central repository, through manual entry of single batches or automated processing of multiple batches. Each batch is normalized in accordance with Aventis' global business rules and assigned a unique ID.

"We use the same business rules implemented at each site, so the data is of high quality and is comparable between the sites," said Dr. Franzisket.

Scientists have adopted the new system because it is specifically designed to do everything they need. "ARP really covers the whole range of different technologies of chemical synthesis," said Dr. Franzisket. "That means traditional experiments and multi-step syntheses, as well as the parallel approach of chemical synthesis and the combichem approach that generates hundreds of thousands of compounds at the same time."

The ARP system helps chemists plan and document their experiments, search and review others' work, analyze results, and generate reports to share with colleagues. To further simplify scientists' day-to-day tasks, the system interface integrates with Aventis' global logistics and analytical systems.

"MDL delivered on our primary requirement," said Pete Loupos, vice president Aventis Drug Innovation & Approval Information Solutions. "They built a system that conformed to the specifics of our scientific workflow, not the confines of software or technology."

Making the right choice

Finding the right registration system requires an awareness of an organization's operational needs and available resources—and an understanding of the attributes of the various systematic approaches.

The unheralded workhorse of an informatics infrastructure, a well planned and executed registration system can improve efficiency and ensure consistency in even the most Byzantine organizations. But don't expect a celebration if you get it right.

"With a registration system, the baseline expectation is one of seamless infallibility," said Aleksander Ruzicic, chemistry practice lead for MDL Consulting. "People tend not to think or talk about it unless there's a problem. But its function is critical to successful discovery research."

For more information on chemical registration systems, or guidance on your unique situation, contact an MDL consultant or your MDL sales representative. ■

The next generation in custom registration

The recent introduction of the n-tier MDL® Discovery Framework opens the door to a new generation of registration systems.

MDL® Core Interface—the middleware integration engine of MDL Discovery Framework that provides a standard interface for custom development—combined with MDL chemistry registration services enable consultants to build extremely flexible systems designed with a custom user interface and a maintained and supported middle layer registration service.

This next-generation solution offers the advantages of an earlier deliverable combined with customization. A European pharmaceutical company recently selected to implement this new Discovery Framework-based solution. The project, managed by Cap Gemini Ernst & Young in collaboration with MDL Consulting, features a custom interface and custom functionality, including a simple plate management and analytical requests interface.