Electronic Laboratory Notebooks

What are they, and what do they need to do?

Presented by: K. T. Taylor, D. Hughes, P. McHale
At: ACS National Meeting, Washington DC
Date: August 31, 2005
Evolution of the products

- Often grew from another need
  - IntelliChem – process documentation
  - Synthematix – synthesis planning
  - Tripos – sample management

- Often complex
- Try to manage more than experiment data
  - LIMS
  - E-Records management
  - Structure registration
  - Synthesis planning
  - Electronic signatures

- But often do not understand reactions
Definitions

CENSA (www.censa.org)
An electronic laboratory notebook is a system to create, retrieve and share fully electronic records in ways that meet all, legal, regulatory, technical and scientific requirements

PRISM forum (www.prismforum.org)
A hybrid electronic laboratory notebook is a system to create, retrieve and share electronic records, which uses printed copies with handwritten signatures for archiving and IP protection

Records are the collection of information or data associated with an experiment to enable a suitably skilled person to repeat it. Records also contain a time and date-stamp.
Inadequacies of the pLN

- Computerization of R&D produces incompatible outputs
- Paper lab notebooks (pLN) only capture part of the research process
  - NMR spectra filed separately
  - Assay results too numerous to enter manually
  - Transcription of a parallel synthesis experiment is time consuming
- Information is entered in many places
  - Re-entering of data is tedious and can introduce errors
    - FDA abhors discrepancies
    - Adversarial attorneys live for discrepancies
Loss of corporate knowledge

- Information in a pLN is soon lost to the company
- Archived but not easily searchable
  - Unless you know who did it or when it was done
- If you do find the information it is often
  - Illegible
  - Incomplete
- Negative information
  - What experiments failed
  - What can I learn from the failures
  - What do I do to make my experiment a success?
- The information is very valuable
  - The context of the experiment is known
  - It is unique to your organization
  - It was expensive to collect
Retrieving all the data

How do you know that you have all the data needed to make or defend a patent claim or prove regulatory compliance when you cannot be sure that you have obtained all the paper notebooks?

Collecting the data is very time-consuming and expensive

Collection of information for patents often takes 6 months of work

Once a product exceeds $100 million a year someone will sue you and will have to collect the historic data.
What do we know about compound X?

What compounds are related to X?

And what does the analysis of Y look like?

What batches of X exist, and where are they?
What is driving ELN adoption? (1)

- Computerization of R&D
  - Experiment design uses electronic databases
    - Bioinformatics, SciFinder, DiscoveryGate
  - And software systems
    - Tripos, Accelrys, BLAST, SRS
  - Experiment execution uses robots
  - Experiment confirmation uses computer-driven analytical equipment
    - LIMS
  - Experiment evaluation uses software
    - Microsoft Excel, Spotfire
  - Experiments are reported in electronic documents
    - Microsoft Word, Microsoft PowerPoint
  - Key information is archived in an electronic database
    - MDL ISIS, Oracle
- But the legal record is still a paper document
What is driving ELN adoption? (2)

- Knowledge Management
  - Knowledge is available to all of the organization
- Productivity gains
  - Enter data only once
  - Repeated experiments can be copied electronically, not transcribed manually
  - Related experiments can be used as templates, rather than transcribed manually
  - Reduced time collecting the information and assembling a legal record
- Completeness of the information
  - I have searched all our databases and I am confident that I have retrieved everything
- Computer literacy of the scientists
  - I cannot read my own handwriting, so you do not have a hope!
What do the scientists want?

- It must make my job easier, not more tedious
  - Touch everything once only
- A blank page, because that is what I am used to
  - But I need pre-defined features to make my job easier
    - Synthesis chemists want stoichiometry tables that automate the task
- It must be accessible at my desk or in the laboratory
MDL Élan
Based on Microsoft Word
User can drop in relevant modules

A blank page ELN

**Synthesis of 2,6-Dichloro-9-thiabicyclo[3.3.1]nonane**

Electrophilic addition

Step: 1

108.18
181.15
C₆H₁₆
C₆H₃Cl₂S

**Apparatus:** 200 ml 3-neck flask equipped with a Teflon-coated magnetic stir bar, a 25-ml pressure-compensating addition funnel with nitrogen inlet, thermometer, and standard taper stopper. Heated entire apparatus under dry nitrogen flow to remove all visible water.

<table>
<thead>
<tr>
<th>unit</th>
<th>mmol</th>
<th>equ.</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>100.0 g</td>
<td>905.9</td>
<td>1.00</td>
<td>L₄-cyclooctadiene (Fluka - 25598; 98 %)</td>
</tr>
<tr>
<td>94.70</td>
<td>905.9</td>
<td>1.00</td>
<td>Sulfur dichloride (974665_D1_3_134; 98.5 %)</td>
</tr>
<tr>
<td>660 ml</td>
<td></td>
<td></td>
<td>Hexane (Fluka 52765)</td>
</tr>
</tbody>
</table>

A 200 ml 3-neck flask was equipped with a Teflon-coated magnetic stir bar, a 25-ml pressure-compensating addition funnel with nitrogen inlet, thermometer, and standard taper stopper. The stopper was removed and dry nitrogen was flowed through the apparatus while heating all glass surfaces with a heat-gun until all visible water condensation was driven from the surface. The nitrogen flow was then switched from flow to vent and the stopper was replaced. After the apparatus was cool, the flask was charged with the sulfur dichloride and 80 ml of dry hexane, and the flask was immersed in a dry-ice-acetone bath. The cyclooctadiene was then charged to the addition funnel. The magnetic stirrer was started, and the cyclooctadiene was added dropwise to keep the reaction temperature below -40 deg C. After the addition of cyclooctadiene was complete, the addition funnel was rinsed with ca. 10ml of additional dry hexane; this was added to the reaction mixture. The dry-ice-acetone bath was removed, and the mixture allowed to come to room temperature.

**TLC of the warmed reaction mixture showed a nearly quantitative conversion of the 1,5-cyclooctadiene.**

The product crystallized directly from the reaction mixture. The reaction mixture was filtered, and the resulting white, fluffy crystals were recrystallized from fresh, dry hexane.

**Final analytes:**

<table>
<thead>
<tr>
<th>Name</th>
<th>NMR</th>
<th>IR</th>
<th>MS</th>
<th>FAB</th>
<th>GC/MS/HRMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (2H₄)</td>
<td>1H-248.04</td>
<td>S</td>
<td>M+248.04</td>
<td>M+248.04</td>
<td>EA-248.04</td>
</tr>
</tbody>
</table>

**Physical properties:**

<table>
<thead>
<tr>
<th>product</th>
<th>amount</th>
<th>purity</th>
<th>physical properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (2H₄)</td>
<td>177.8 g</td>
<td>99%</td>
<td>m.p. = 35-35°C, 194.3 g</td>
</tr>
</tbody>
</table>

**Comments:** Very exothermic reaction; use caution at larger scales

**Reaction Specification:** Optimized
Driving user adoption

- It is not enough for the ELN to do what the pLN does but slicker
- Users will not change because the pLN
  - Does the job
  - Is understood
  - Is portable
  - Is accessible
- The ELN must be fully integrated with my workflow
- There must be a major benefit to my productivity
The ELN at the center of the workflow

The gateway to systems

Integrates systems and Aggregates information

The repository of the information
Why do people use paper laboratory notebooks?

- Because they have to!
  - Defend a US patent
    - First to invent
  - Prove compliance with predicate rules
    - Example: US FDA
- Two environments
  - US Patent & Trademark Office (PTO)
  - US Food & Drug Administration (FDA)
The traditional requirements

- Satisfy the legal record:
  - Why
    - Conception / invention
  - Who
    - Authenticate, authorize
  - How
    - Repeatable by a competent person skilled in the art
  - When
    - Priority of invention / compliance with regulated procedure
  - What
    - Support reduction to practice
Some requirements are similar

- Produce a legal record
  - Defined by the Federal Rules of Evidence (FRE)
    - Authenticate the author or authors
    - Audit the information
    - Prove chain-of-custody
Some requirements differ

- FDA is rules based
  - Whatever you do must be documented
  - If it is not documented, it did not happen, and you are in trouble!
  - If it happened, and you did not document it, you are in trouble

- Patents
  - Only document what is needed to support the patent claim
  - Patent infringements are adversarial
  - Patent records are subject to the legal discovery process
  - Need to restrict areas that are open to discovery
What does this mean?

- **Regulated environment**
  - Everything that was done must be recorded and available
    - If it is not documented then it did not happen
      - And you are in trouble
    - If it was done and it is not documented
      - You are in trouble

- **Patent environment**
  - Only record what is needed to defend a patent claim
    - Be careful about linking to other systems from within the notebook record
      - Linked systems become exposed to the *discovery* process
ELN: the central application

- The application that is used by all scientists, everyday, for most of the day
  - It will become the hub of the R&D desktop
- It must deliver much more than the pLN
- More than an electronic version of paper notebook. It is the **gateway** to:
  - Information required to design experiments - both internal and external
  - Reagents and equipment required to perform experiment
  - Analytical equipment required to analyze results
  - Information systems required to register materials
From paper to electronic

**Standardization of KOH Solution**

**Purpose:** The purpose of this experiment is to standardize a KOH solution by titration using potassium hydrogen phthalate as a primary standard acid.

Potassium hydrogen phthalate (KHP) = \(\text{C}_8\text{H}_4\text{O}_4\text{K} \cdot \text{H}_2\text{O}\) from 1069, 204.22 g/mol.

\[\text{KOH} + \text{C}_8\text{H}_4\text{O}_4\text{K} \rightarrow \text{C}_8\text{H}_4(\text{COO})_2 + \text{H}_2\text{O}\]

<table>
<thead>
<tr>
<th>Trial</th>
<th>Mass of KHP, g</th>
<th>Initial Vol, mL</th>
<th>Final Vol, mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.7543</td>
<td>21.01</td>
<td>39.88</td>
</tr>
<tr>
<td>2</td>
<td>0.7742</td>
<td>0.80</td>
<td>20.16</td>
</tr>
<tr>
<td>3</td>
<td>0.7435</td>
<td>2.45</td>
<td>21.01</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Samples of KHP were weighed into 250-mL Erlenmeyer flasks, about 100 mL of distilled water were added to each. The flasks were covered to dissolve the samples. Phenolphthalein (3 drops) was added to each flask as an indicator. At this point, the solutions were colorless. A burette was rinsed with KOH solution, then filled. The first sample was titrated until a light-pink color. The color was judged to be slightly darker than desired. The second sample was titrated, with the color of the endpoint being lighter. The third sample was titrated to a satisfactory endpoint.

Conrad J. Shibata
4/27/00
An ELN built on a solid foundation technology
The ELN can stand alone

The ELN is central to scientists’ workflow
Integration delivers the most value

Process & summary data captured in ELN over time

Login | Research & Design | Source | Make | Analyze | Store | Sign & Witness | File

ELN

- Authentication
- Search External & Internal Content
- Reagent Logistics
- Sample Logistics
- LIMS
- Assay Results
- Analytical Repository
- Summary Repository
- Corporate Registry
- Registration
- Signature Capture
- E-Records System

Integration delivers the value needed to drive adoption
Integration delivers the most value

Key:
Orange = Available from MDL
Blue = external service

Systems available from Elsevier MDL
Designing for integration

- Integration, scalability, and extensibility must be part of the original design
- Elsevier MDL is experienced in the delivery of mission critical, enterprise systems to multi-site, multi-national organizations
- Elsevier MDL provides:
  - Workflow solutions
    - MDL® Notebook, MDL® Logistics, MDL® Assay Explorer
  - Chemistry technology
    - MDL® Direct, MDL® ISIS/Host
  - The integration platform for scientific applications
    - MDL® Isentris®
- Many vendors are building on MDL’s integration platform though the Isentris® Alliance
ELN Benefits (1)

- Experiment records are:
  - Lost no more
    - All the information is captured in a secure archive that is fully searchable
  - Legible
    - Not handwritten
  - Complete
    - Now that notebook records are being re-used, the authors are tending to make them more complete
      - Result is closer to what a patent requires
  - Secure
ELN Benefits (2)

Productivity gains

- *Failed* experiments are not repeated
- Successful procedures are re-used
- Integrated workflow delivers most benefit

“The MedChem Tools make the parallel synthesis process faster and easier from conception through registration. It used to take 17-23 hours to search for the reagents, draw the structures, set up the purification file, select the fractions for QC, run the QC analysis, and finally register the compounds. With the tools, this process takes 4-6 hours.” – Medicinal Chemist, San Diego

Source Electronic Laboratory Notebooks, IQPC conference, Atlanta, 2005
ELN Benefits (3)

- Retrieving information is faster
  - Collecting all the information for a patent
    - Typically takes 6 months with paper notebooks
  - Project reporting
    - Click on a button to activate a report template
  - Retrieved information is complete
    - When the electronic searches are done, you are confident that you are done
ELN Benefits (4)

- **Enforcing GxP**
  - **Validation**
    - Is the person trained in the technique?
    - Is the equipment validated for the technique?
    - Is the action appropriate at this time?
    - Have all preceding steps been completed satisfactorily?
  - **Support *Due Diligence in Reduction to Practice* for an invention**
    - Timestamp all activities
    - Analytical confirmation requested and completed in a timely manner
    - Requested reason for delays and inactivity
**ELN Benefits (4)**

- **Substantial Return on Investment (ROI)**
  - less time planning
  - less time recording
  - Less time reporting
  - Less time collecting information for IP protection and regulatory compliance
  - eliminate duplicate experiments

- **Typical ROI:**
  - 100 scientists using ELN
    - Savings ~$850k per year
Conclusions

- ELNs deliver substantial productivity gain & ROI
- ELNs enhance knowledge management
- ELNs encourage better quality records
- ELNs give you confidence that you have all the data
- ELNs support *Due Diligence* and SOPs
- ELNs deliver the most value when they are an integral component of a collaborative e-R&D environment