

Designing Data Repositories to Support Preservation and Publication for the Chemistry Community

Simon Coles
EPSRC National Crystallography Service
School of Chemistry
University of Southampton



Social Networks for Chemists

The screenshot displays the nanoHUB website, a platform for nanoscience and technology. The header features the nanoHUB logo and the text "an NCN project". The navigation bar includes links for Home, My nanoHUB, Resources, Contributors, Events, About, and Support. A search bar and a help icon are also present.

The main content area features a "Nanotechnology 101" section with a thumbnail image of a scanning electron micrograph (SEM) showing a surface at 1 nanometer scale. Below the image is a brief description of the series: "Nanotechnology 101 is a series of lectures designed to provide an undergraduate level introduction to nanotechnology. Our Nanotechnology 501 series offers lectures directed at the graduate student/ professional level." A "Learn more" button is available for both series.

On the right side, there is a sidebar with the text: "A resource for nanoscience and technology, the nanoHUB was created by the NSF-funded NCN and its development is driven by research themes in:" followed by three categories: "Nanoelectronics" (marked as "new"), "NEMS/Nanofluidics", and "Nano-Bio Devices".

The main content area is divided into several sections:

- Simulate**: Includes links for "Nanoelectronics Tools for nanoelectronics", "NEMS/Nanofluidics Tools for NEMS and Nanofluidics", and "Nano-Bio Devices Tools for nano-bio devices". A "More" link leads to "Browse all available tools".
- Research**: Includes links for "Seminars Browse research seminars", "Collaborate Work with your colleagues", "Web Meetings Right in your browser", and "User Groups Share with your colleagues".
- Teach & Learn**: Includes links for "Nano 101 / Nano 501 Introductory tutorials", "Nanocurriculum Curriculum on Nanoelectronics", "Learning Modules Self-paced web instruction", and "Teaching Materials Graduate, Undergrad, K-12".
- Contribute**: Includes links for "Contribute Content Upload your own materials", "Give us Feedback Success story? Suggestions?", "Take a Poll What aggregator do you use for RSS feeds?", and "Donations Contribute your financial support".

The NCN logo is visible in the bottom right corner of the sidebar.

Google generation: new behaviour and approach

Sharing Rich Media

• Video + Paper = Pubcast

New Approaches to 'Sharing Experiments'



- Specialised domain-oriented innovations

Formation of Open Communities

The following institutions are listed on the button:

- Europe**
 - [Universiti Teknologi Malaysia \[show\]](#)
 - [Bristol University \[show\]](#)
 - [CeMM \(Vienna\) \[show\]](#)
 - [CNRS \[show\]](#)
 - [Cambridge Research Institute \[show\]](#)
 - [ETH Zurich, Switzerland \[show\]](#)
 - [FH Campus Wien \[show\]](#)
 - [Imperial College \[show\]](#)
 - [King's College London \[show\]](#)
 - [Norwich Research Park, UK \[show\]](#)
 - [PRBB \[show\]](#)
 - [Technische Universiteit Delft \[show\]](#)
 - [Univ of Cambridge \[show\]](#)
 - [Univ of Edinburgh \[show\]](#)
 - [Univ of Ljubljana \[show\]](#)
 - [Univ of Manchester, UK \[show\]](#)
 - [MH Hannover, Germany \[show\]](#)
 - [Univ Paris 5-INSERM \[show\]](#)
- Oceania**
 - [CAMBIA \[show\]](#)
 - [Univ of Otago \[show\]](#)
- Labs in development**
 - [Labs in development \[show\]](#)

[\[edit\]](#) [\[edit\]](#) [\[edit\]](#)

Los Alamos National Laboratory [show]

MIT [show]

Oregon Health & Science University [show]

Penn State University [show]

Purdue University [show]

Rice University [show]

Rochester Institute of Technology [show]

The Rockefeller University [show]

Texas A&M [show]

Tufts University [show]

UC Berkeley [show]

UC Davis [show]

UC San Diego [show]

UC San Francisco [show]

UI Urbana-Champaign [show]

Univ Alabama Birmingham [show]

Univ of Chicago [show]

Univ of Connecticut Health Center [show]

Univ of New Mexico [show]

Univ of Notre Dame [show]

Univ of Oregon [show]

Univ of Rhode Island [show]

Univ of Tennessee [show]

Univ of Toronto [show]

Univ of Victoria [show]

Univ of Virginia [show]

UT Austin [show]

Vanderbilt University [show]

Add your lab to OWW [\[edit\]](#)

Interested in starting a lab wiki on OpenWetWare?

See the [starting a lab wiki tutorial](#)

- New approaches surfacing and growing FAST

Open Notebook Science

Useful Chemistry [XML](#) [subscribe with bloglines](#) [UsefulChem molecules](#) [UsefulChem wiki](#)

Google™ Custom Search [Search](#)

This is an open source science project in chemistry. Post specific problems in chemistry that Post specific partial solutions to these problems. Or execute a suggested step. NOTE: ANYTHING SUBJECT TO A SHARE-ALIKE WITH ATTRIBUTION CREATIVE COMMONS LICENSE (see bottom of page).

UsefulChem [Exp025](#) [Protected](#) [page](#) [discussion](#) [history](#) [notify me](#)

Oc1ccc(Oc2ccc(O)cc2)cc1C[C@H](N)C $\xrightarrow[\text{116-118°C; 1h}]{\text{H}_3\text{PO}_4}$ Oc1ccc(Oc2ccc(O)cc2)cc1

Objective
To convert [adrenaline](#) to [DOPAL](#) by acid catalysis

Procedure
A solution of [adrenaline](#) (1.0g 5.5mmoles) in 85% phosphoric acid was heated (116-118°C; 1h) in a heating mantle then removed from heat and allowed to cool. The solution was stirred for 1h, then saturated with NaCl. It was taken up in ethyl ether, and dried over anhydrous MgSO₄. The ether was then removed by rotovap to obtain DOPAL (80 mg 0.53mmol, 9.5% yield)

Characterization:
1. [TLC of 25A](#) in 3:1 MeCl₂/MeOH and in 6:1 MeCl₂/MeOH, and [stained with CAM](#) (Varian inova). The integration is good enough to not require further purification.

Results
1. [TLC of 25A](#) in 3:1 MeCl₂/MeOH and in 6:1 MeCl₂/MeOH, and [stained with CAM](#) (Varian inova). The integration is good enough to not require further purification.

Discussion
This is the first time that DOPAL was obtained pure (by NMR integration) immediately after heating. There are several factors that may have contributed to this: preheating the phosphoric acid then adding the reaction under nitrogen and a careful temperature control. It is likely that the main impurity in previous attempts ([EXP016](#), [EXP023](#)) is the carboxylic acid ([Exp016HNMR](#), [Exp023HNMR](#)).

Conclusion
DOPAL can be obtained pure in 9.5% yield by heating adrenaline in 85% phosphoric acid hydrolysis and extraction into ether.

Help me improve my notebook
I'm also trying to come up with some [Open Notebook Science](#) requirements and suggestions to help move towards an open notebook that is intelligible to people besides myself. I'd be interested for your comments/suggestions on what else should be included.

How I construct my notebook
A few people have written asking how I make the document itself. The document is made in [LaTeX](#) (specifically [pdflatex](#)) and [sample.pdf](#) is the compiled document. If you don't have [pdflatex](#) installed, you can download [Introduction to LaTeX](#).

RR RESEARCH
THINKING ABOUT OUR RESEARCH INTO THE MECHANISM, FUNCTION AND EVOLUTION OF DNA UPTAKE BY HAEMOPHILUS INFLUENZAE AND OTHER BACTERIA.

SATURDAY, JANUARY 19, 2008

New microarray data

The post-docs have finished the first-pass analysis of how *E. coli* gene expression is affected by both the *E. coli* Sxy and the *H. influenzae* Sxy proteins. I suppose I shouldn't be surprised that it's more complicated than I had hoped. For example, unlike the situation in *H. influenzae*, in *E. coli* there are also groups of genes whose expression goes down when Sxy is present.

One complication is that these cells are probably seriously OVER-producing Sxy. Unlike *H. influenzae*, where we've only done arrays of cells expressing a single-copy *sxy* gene under its natural promoter, these *E. coli* studies used a *sxy* gene on a high-copy plasmid and under a highly inducible promoter. We know that prolonged expression of Sxy from this plasmid produces large quantities of denatured Sxy (in inclusion bodies) and we don't know the extent to which even the 30-minute expression used for the array studies might create a situation unlike that of natural *sxy* expression.

POSTED BY ROSIE REDFIELD AT 6:25 PM

2 COMMENTS:

Bayork said...
Remember what you've said about sxy in the past. Can you make a null version of

ABOUT ME
 **ROSIE REDFIELD**
CANADA
I run a microbiology research lab in the Life Sciences Centre at the University of British Columbia.
[VIEW MY COMPLETE PROFILE](#)

PREVIOUS POSTS
[Thermodynamics of home heating](#)
[Gene transfer agent](#)
[Sorry, wrong link](#)
[Sorry for the paucity of posts](#)
[Genespring progress and problems](#)
[Data on *E. coli* protein](#)

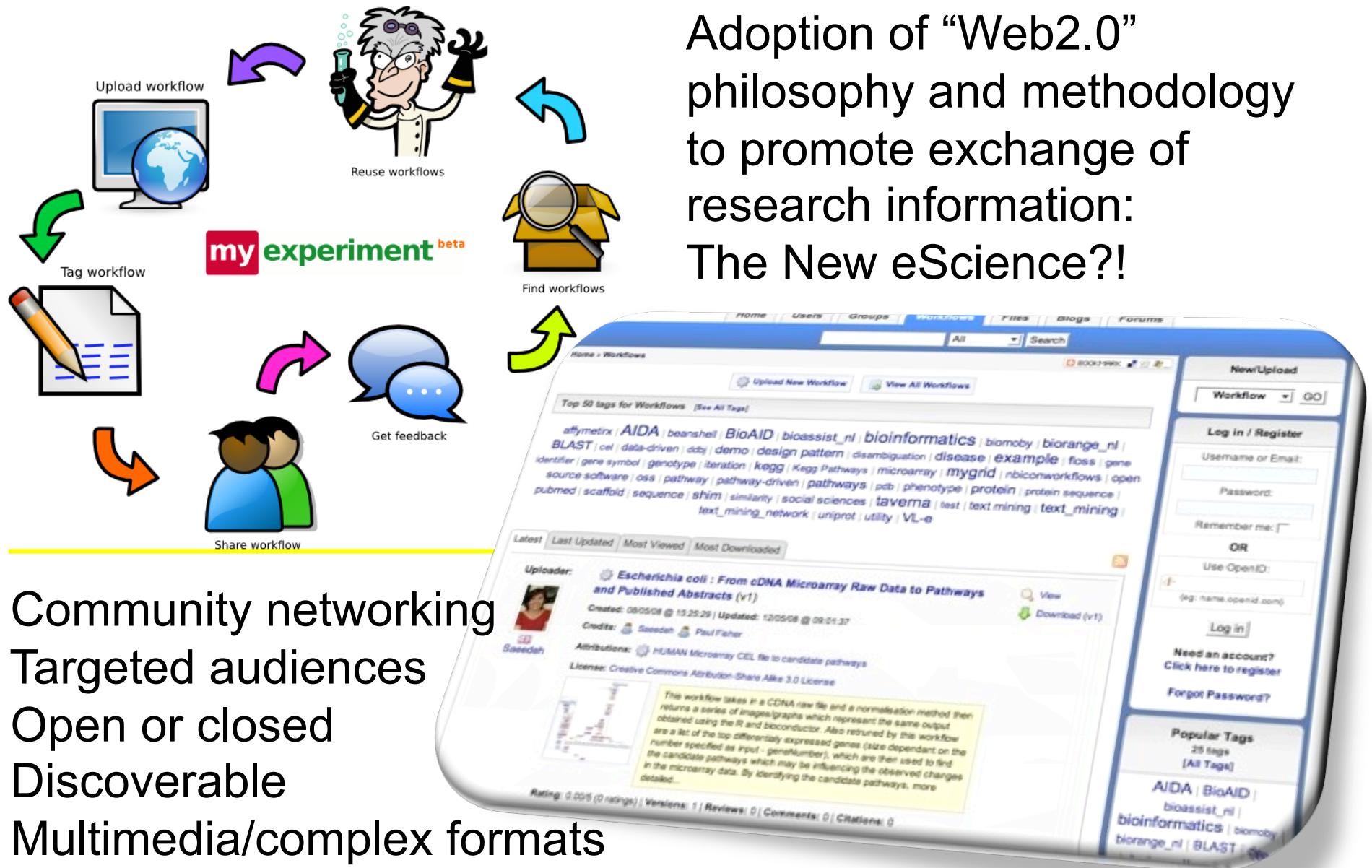
- Immediate sharing of experimental information & data

New Information Exchange Environments



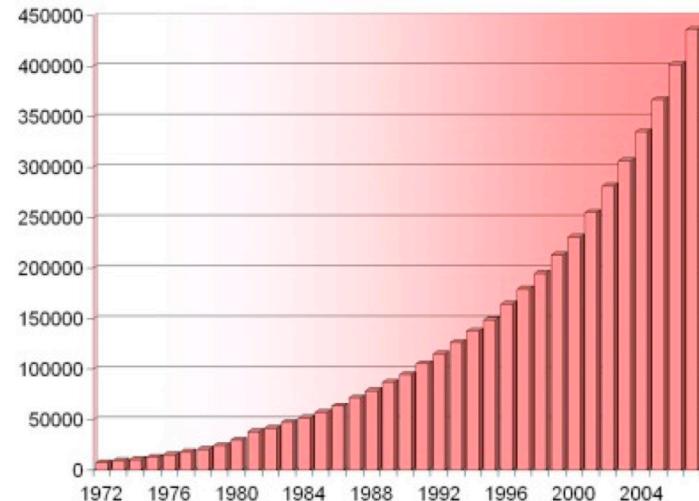
- Immersive alternative to conventional browsing & interaction

A Virtual Research Environment



Data Deluge

- 40 years ago a PhD student would determine about 3 crystal structures during the course of their study – this can now be easily achieved in a day
- There are approx. 30 million known chemical compounds
- Approx. 2 million crystal structures have been determined
- There are less than 0.5 million published crystal structures residing in (licensed) curated databases
- There are just a few thousand ‘open’ crystal structures
- The primary cause of this is the current data publication process, which is tied to journal articles and peer review



Data Publishing & Open Access

- Short communications
- Electronic only
- Rapid publication
- Open access (01/2008)
- Highly cited
- Written in CIF
- Freely available tools
- **Still cannot keep up!**
- Journal ‘article’ format required
- Not all crystal structures are of primary importance to the underlying chemistry: by-products / unexpected results / tracking reactions



The Solution

Intellect & Interpretation (Journal article, report, etc)



Structural investigations of phosphorus-nitrogen compounds. 4. Steric and electronic effects in dibenzylamino derivatives of hexachlorocyclotriphosphazatriene and 4,4,6,6-tetrachloro-2,2-diphenylcyclotriphosphazatriene¹

Sarp Bedi,^a Simon J. Coles,^b David R. Davies,^a Michael A. Huthnance,^a Adem Kole,^a Thomas A. Mayer,^a Robert A. Shaw^a and Aysin Ustun^a

^a Department of Chemistry, Celal Bayar University, Faculty of Sciences, Faculty of Education, Faculty of Engineering, Faculty of Medicine, 44130, Manisa, Turkey

^b Department of Chemistry, University of London, 22 Gordon Square, London, WC1E 6BT, United Kingdom

Correspondence e-mail: s.j.coles@kcl.ac.uk

Received 11 December 2001
Accepted 27 February 2002

Part 5: Akyol et al.
(1988)

A systematic study is presented on the products of aminolysis of $\text{N}_3\text{P}(\text{Cl})_3$ and $\text{N}_3\text{P}(\text{Ph})_3$ with dibenzylamine (1) and (4), namely, $\text{N}_3\text{P}(\text{Cl})_3[\text{N}(\text{C}_6\text{H}_5)_2]_2$ (2) and $\text{N}_3\text{P}(\text{Ph})_3[\text{N}(\text{C}_6\text{H}_5)_2]_2$ (3) and $\text{N}_3\text{P}(\text{Ph})_3[\text{N}(\text{C}_6\text{H}_5)_2]_3$ (4) and $\text{N}_3\text{P}(\text{Ph})_3[\text{N}(\text{C}_6\text{H}_5)_2]_2$ (6) [where (2), (3), (5) and (6) are the products of the reaction of (1) with (4), (3) and (6) are the products of the reaction of (1) with (3)]. Steric and electronic effects prevail in the formation of dibenzylamino-substituted cyclophosphazenes. The influence of an electron-releasing group (i.e. phenyl) on the stereochemistry and degree of substitution of the products is analysed by comparison with those two series. The electronic effect in unimERICALLY substituted endocyclic P–N bond length, Δ , is used as a measure of the degree of the electronic contribution in combination with basicity constants, to quantify the degree of the electron-releasing capacity of the R group. In order to compare geminal versus non-geminal substitution, a difunctional secondary amine was used to form the compound $\text{C}_6\text{H}_5\text{N}(\text{Me})\text{C}_6\text{H}_5$ (7) (a reinvestiga-

tion of a previously reported compound).

© 2002 RSC Publishing

DOI: 10.1039/b107540k

ISSN 0959-991X

10.1039/b107540k

© 2002 RSC Publishing

The eCrystals Data Repository

• Quick & simple to deposit

• Software tools

• Laboratory archive

• Community involvement

• ‘Embargo’ facility

• Structured foundations

• Discoverable & harvestable

Sample Originator: Esther Rousay* and Jeremy G. Frey*.

Data Collection: Simon J. Coles*

Structure Determination: Simon J. Coles* and Michael B. Hursthouse*.

University of Southampton*

C₁₄H₂₀O₅

InChI=1/C14H20O5/c1-2-4-14-13(3-1)18-11-9-16-7-5-15-6-8-17-10-12-19-14/h1-4H,5-12H2

Identification Number: 10.3737/ecrystals.chem.soton.ac.uk/145

Controlled keywords: crown ethers, crown

Date Created: 07 October 2004

Deposited On: 21 Jan 2008 15:29

By: Dr Simon J Coles

Available Files

Final Result
04sjc0831.cif
04sjc0831.cml
04sjc0831.fcf.txt

Collection parameters

Chemical formula	C14 H20 O5
Crystallisation Solvent	Plate
Crystal morphology	Orthorhombic
Crystal system	Orthorhombic
Space group symbol	Pbca
Cell length a	16.4963(18)
Cell length b	8.325(3)
Cell length c	20.061(6)
Cell angle alpha	90.00
Cell angle beta	90.00
Cell angle gamma	90.00
Data collection temperature	120(2)

Processing

04sjc0831.hkl
04sjc0831.htm
04sjc0831_0kl.jpg
04sjc0831_l0l.jpg
04sjc0831_h00.jpg

Data collection

04sjc0831_crystal.jpg

Other files

04sjc0831.doc
04sjc0831.ins
04sjc0831.mol
04sjc0831.p4p
04sjc0831.pcif.txt
04sjc0831_ellipsoid.gif

Citation: Rousay, Esther and Frey, Jeremy G. and Coles, Simon J. and Hursthouse, Michael B. (2004) University of Southampton, Crystal Structure Report Archive. [doi:10.3737 ecrystals.chem.soton.ac.uk/145]

UKOLN

DCC

eprints

A Thorough Approach to Dissemination

- Using simple Dublin Core protocol (OAI-PMH)
 - Crystal structure
 - Title (Systematic IUPAC Name)
 - Authors
 - Affiliation
 - Creation Date
- Additional **chemical** information through Qualified Dublin Core
 - Empirical formula
 - International Chemical Identifier (InChI)
 - Compound Class & Keywords
- Specifies which 'datasets' are present in an entry
- Application Profile [*http://www.ukoln.ac.uk/projects/ebank-uk/schemas/*](http://www.ukoln.ac.uk/projects/ebank-uk/schemas/)
- DOI links [*http://dx.doi.org/10.1594/ecrystals.chem.soton.ac.uk/145*](http://dx.doi.org/10.1594/ecrystals.chem.soton.ac.uk/145)
- Rights & Citation [*http://ecrystals.chem.soton.ac.uk/rights.html*](http://ecrystals.chem.soton.ac.uk/rights.html)

A Thorough Approach to Preservation

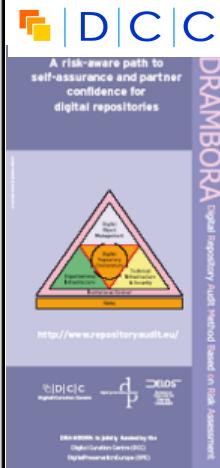
A study of Curation and Preservation Issues in the eCrystals Data Repository and Proposed Federation

eBank-UK Phase 3: WP4
September 2006 - June 2007

Final Version (Revised): 7th September 2007

Manjula Patel
UKOLN, DCC
University of Bath, UK

Simon Coles
National Crystallography Centre
University of Southampton, UK



KEEPING RESEARCH DATA SAFE

A COST MODEL AND GUIDANCE FOR UK UNIVERSITIES

Neil Beagrie, Julia Chruszcz, and Brian Lavoie

with case studies contributed by the Universities of Cambridge, Southampton, King's College London, and the Archaeology Data Service University of York.

Final Report - April 2008

Prepared by:

Charles Beagrie Limited

www.beagrie.com

A study funded by



Large repositories need failure resilience — the ability to suffer multiple types of failures without risk of losing data integrity. Through parallelism, Honeycomb can provide a level of reliability that is arguably better than what customers can get in any SAN environment.

The last major unsolved challenge for these customers is the management of metadata, the rich set of attributes that describe the data and allow it to be recalled instantly.

Scaling Up: A Community Solution

Interviews & analysis

Synthesis: IR Policy & Practice, Laboratory Practice & Workflows, Technical Interoperability & Standards, Metadata Schema & Application Profiles, Semantic Interoperability, Data Citation, Identifiers & Linking, Federation Architectures & Third Party Services, Rights & Licensing, Data Quality & Validation, Preservation, Curation & Sustainability



Recommendations, commentary

Matters Arising: Diverse lab practice, LIMS and proprietary formats, Data policy should reflect lab practice & institutional model, Data quality criteria/validation, “Prior publication” problem, We need scalable assignment of “terms” for data discovery, No discipline preservation model

Scaling Up: Towards a Federation of Crystallography Data Repositories

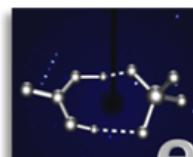
Document details

Author:	Liz Lyon, Simon Coles, Monica Duke, Traugott Koch
Date:	12th May 2008
Version:	1.0 Final
Document Name:	ebank-phase3-report-final.doc
Notes:	

<http://wiki.ecrystals.chem.soton.ac.uk>

The eCrystals Federation

Chem Seer



RSC



Centre For Molecular Science Informatics



The University of Sydney

Reciprocal Net



UNIVERSITY OF
Southampton



Useful Chemistry



Science & Technology
Facilities Council

JISC



EPSRC

Engineering and Physical Sciences
Research Council



$$n\lambda = 2 d \sin\theta$$

blogs@ChemTools

Chemistry of Crystals



D|C|C

CrystalEye (beta)



eBank UK



Chemical Database Service



General Chemistry Issues: Data Generation

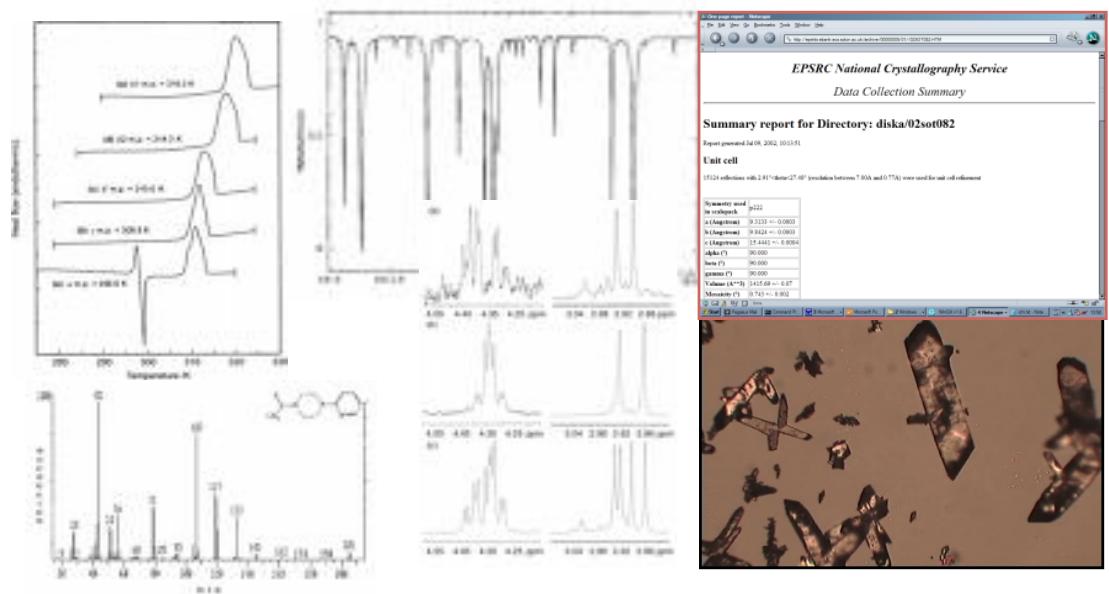
Compound Details		For X-ray	22.07.2004
Researcher	Gozal YENILMEZ (GIFTCI)		
Code	303-489-a		
Compound	C ₁₂ H ₁₂ Cl ₂ N ₂ O ₂ P		
Empirical			
Formula			
Prepared			
Structural			
Formula			
Melting Point	132-135 °C		
Reagents (and	J,3-propenoid, tetrahydrofuran, NaH, dichloromethane, Hexane, D ₂ O		
reactions) (and			
General	This product is soluble in CH ₂ Cl ₂ , CHCl ₃		
Properties for			
light, thermal			
and chemical			
stability			
Crystallisation	CH ₂ Cl ₂ /Hexane, D ₂ O (3:1)		
Source			
Appearance of	Red colorless		
Compound			
Mass Calc.	Cal. 303.489-a (760)		
Mass Measured			
Calcd (g/mol)			
Mass Measured (g/mol)			
Calcd (g/mol)			
Reaction Scheme			
NMR Data			
UV-vis			



Synthesis



Characterisation



Shortfalls in Data Management

“Data from experiments conducted as recently as six months ago might be suddenly deemed important, but those researchers may never find those numbers – or if they did might not know what those numbers meant”

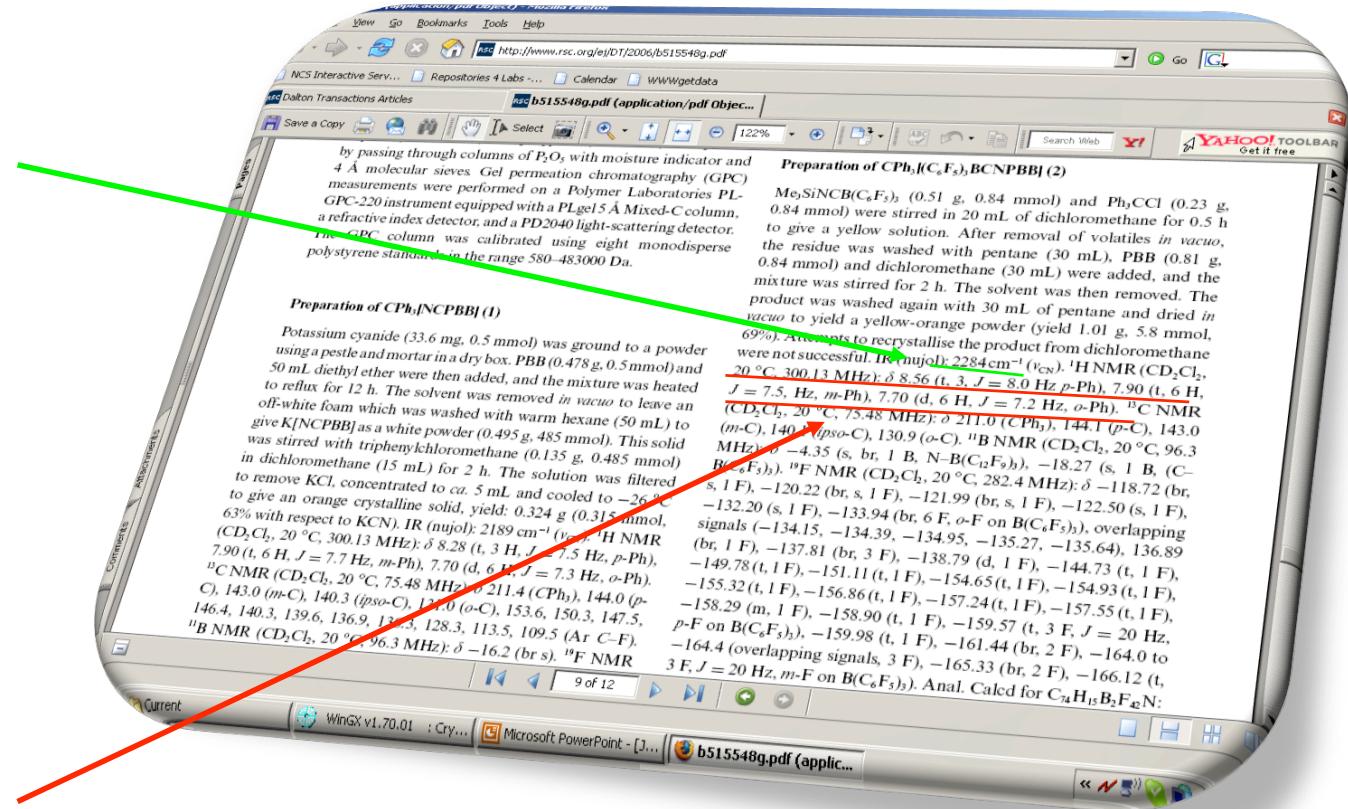
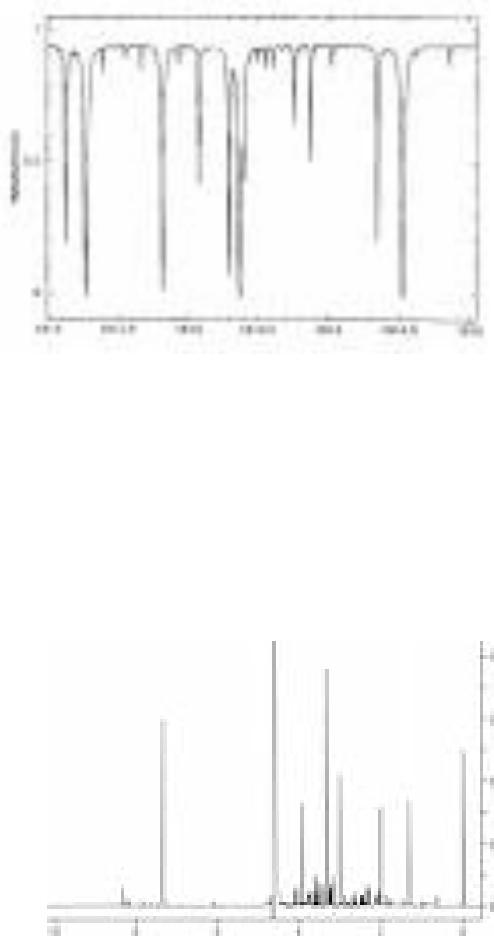
“Lost in some research assistant’s computer, the data are often irretrievable or an undecipherable string of digits”

“To vet experiments, correct errors, or find new breakthroughs, scientists desperately need better ways to store and retrieve research data”

“Data from Big Science is ... easier to handle, understand and archive. Small Science is horribly heterogeneous and far more vast. In time Small Science will generate 2-3 times more data than Big Science.”

‘Lost in a Sea of Science Data’ S.Carlson, The Chronicle of Higher Education (23/06/2006)

Analysis Data Publication & Information Loss



Spectroscopic analysis is often performed to ensure a reaction is proceeding according to plan – as a result <5% are published (via a process with heavy information loss)

A New Kind of Electronic Lab Notebook

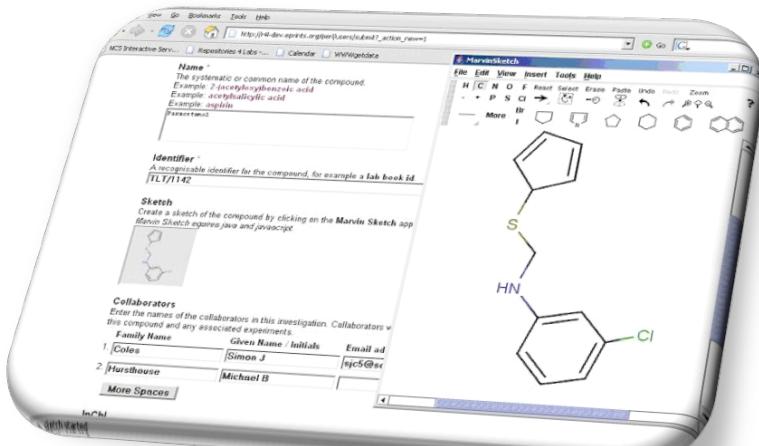
the Smart Tea Project



Weigh-Station #1		Sep 19, 2003 2:51:28 PM																															
mrg		mrg/3828/4																															
Experiment Details																																	
<table border="1"> <thead> <tr> <th>Name</th> <th>Planned</th> <th>Actual</th> </tr> </thead> <tbody> <tr> <td>Europium acetate</td> <td>0.2500 g</td> <td>0.2516 g</td> </tr> <tr> <td>4chloro dipicoline acid monoh...</td> <td>0.5000 g</td> <td>0.4950 g</td> </tr> <tr> <td>Triethylamine</td> <td>0.6 ml</td> <td>0.6 ml</td> </tr> <tr> <td>methanol</td> <td>100.0 ml</td> <td>110.0 ml</td> </tr> <tr> <td>Acetone</td> <td>30.0 ml</td> <td>0.0 ml</td> </tr> </tbody> </table>		Name	Planned	Actual	Europium acetate	0.2500 g	0.2516 g	4chloro dipicoline acid monoh...	0.5000 g	0.4950 g	Triethylamine	0.6 ml	0.6 ml	methanol	100.0 ml	110.0 ml	Acetone	30.0 ml	0.0 ml														
Name	Planned	Actual																															
Europium acetate	0.2500 g	0.2516 g																															
4chloro dipicoline acid monoh...	0.5000 g	0.4950 g																															
Triethylamine	0.6 ml	0.6 ml																															
methanol	100.0 ml	110.0 ml																															
Acetone	30.0 ml	0.0 ml																															
<table border="1"> <tr><td>7</td><td>8</td><td>9</td></tr> <tr><td>4</td><td>5</td><td>6</td></tr> <tr><td>1</td><td>2</td><td>3</td></tr> <tr><td>0</td><td>.</td><td></td></tr> </table>		7	8	9	4	5	6	1	2	3	0	.		Enter	Del																		
7	8	9																															
4	5	6																															
1	2	3																															
0	.																																
All measurements completed.																																	
<input type="button" value="123 - 1111 - Water"/> <input type="button" value="222 - 2222 - Tens"/> <input type="button" value="333 - 3333 - Sugar"/> <input type="button" value="444 - 4444 - Coffee"/> <input type="button" value="555 - 1233 - Mrs Bar"/>		<input type="button" value="Escape"/>																															
<input type="button" value="Quit"/>	<input type="button" value="Weigh"/>	<input type="button" value="Liquid-Measure"/>	<input type="button" value="Bench"/>																														
Bench-Station #1																																	
22-Sep-2003 16:01:39																																	
dj	djbj3403																																
Experiment Details																																	
<table border="1"> <thead> <tr> <th>Stage</th> <th>Instructions</th> <th>Done</th> </tr> </thead> <tbody> <tr><td>1</td><td>Dissolve 4-fluorinated biphenyl in butanone</td><td><input type="checkbox"/></td></tr> <tr><td>2</td><td>Add K₂CO₃ powder</td><td><input type="checkbox"/></td></tr> <tr><td>3</td><td>Heat at reflux for 1.5 hours</td><td><input type="checkbox"/></td></tr> <tr><td>4</td><td>Cool and add Br11OCB</td><td><input type="checkbox"/></td></tr> <tr><td>5</td><td>Heat at reflux until completion</td><td><input type="checkbox"/></td></tr> <tr><td>6</td><td>Cool and add water (30ml)</td><td><input type="checkbox"/></td></tr> <tr><td>7</td><td>Extract with DCM (3x40ml)</td><td><input type="checkbox"/></td></tr> <tr><td>8</td><td>Combine organics, dry over MgSO₄ & filter</td><td><input type="checkbox"/></td></tr> <tr><td>9</td><td>Remove solvent in vacuo</td><td><input type="checkbox"/></td></tr> </tbody> </table>				Stage	Instructions	Done	1	Dissolve 4-fluorinated biphenyl in butanone	<input type="checkbox"/>	2	Add K ₂ CO ₃ powder	<input type="checkbox"/>	3	Heat at reflux for 1.5 hours	<input type="checkbox"/>	4	Cool and add Br11OCB	<input type="checkbox"/>	5	Heat at reflux until completion	<input type="checkbox"/>	6	Cool and add water (30ml)	<input type="checkbox"/>	7	Extract with DCM (3x40ml)	<input type="checkbox"/>	8	Combine organics, dry over MgSO ₄ & filter	<input type="checkbox"/>	9	Remove solvent in vacuo	<input type="checkbox"/>
Stage	Instructions	Done																															
1	Dissolve 4-fluorinated biphenyl in butanone	<input type="checkbox"/>																															
2	Add K ₂ CO ₃ powder	<input type="checkbox"/>																															
3	Heat at reflux for 1.5 hours	<input type="checkbox"/>																															
4	Cool and add Br11OCB	<input type="checkbox"/>																															
5	Heat at reflux until completion	<input type="checkbox"/>																															
6	Cool and add water (30ml)	<input type="checkbox"/>																															
7	Extract with DCM (3x40ml)	<input type="checkbox"/>																															
8	Combine organics, dry over MgSO ₄ & filter	<input type="checkbox"/>																															
9	Remove solvent in vacuo	<input type="checkbox"/>																															
2 of 10 tasks completed.																																	
		<input type="button" value="Escape"/>																															
<input type="button" value="Quit"/>	<input type="button" value="Weigh"/>	<input type="button" value="Liquid-Measure"/>	<input type="button" value="Bench"/>																														

A General Chemistry Laboratory Repository

Create new compound
(parent record)



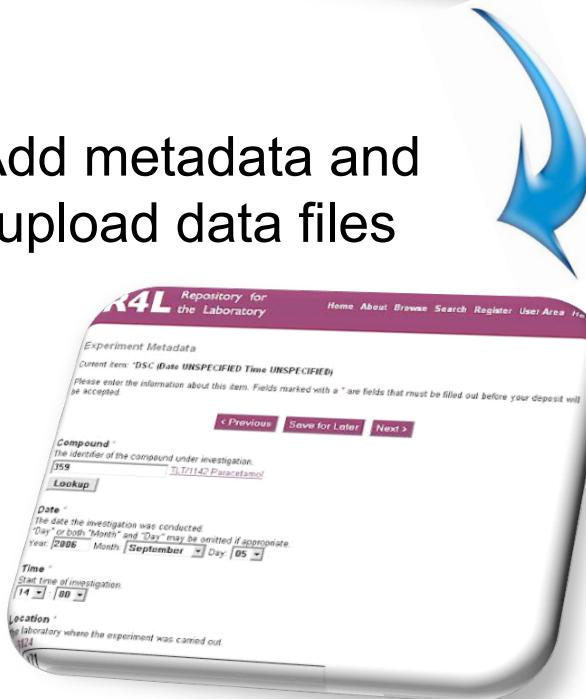
The screenshot shows the R4L Repository interface for creating a new compound. The MarvinSketch app is open, displaying the chemical structure of 2-(4-chlorophenyl)benzenesulfonamide. The structure is shown with a benzene ring attached to a sulfonamide group (-NH-SO₂-CH₃) which is further attached to a 4-chlorophenyl group. The 'Identifier' field contains 'JSL771142'.

Add new experiment type



The screenshot shows the R4L Repository interface for adding a new experiment type. A modal dialog box is open, listing various experimental methods: Single Crystal Diffraction, Powder X-Ray Diffraction, IR, UV-Vis, Mass Spec, Raman, Optical Microscopy, DSC, TGA, NMR, Solid State NMR, and Elemental Analysis. The 'Add Experiment' button is visible at the bottom right.

Add metadata and
upload data files



The screenshot shows the R4L Repository interface for adding experiment metadata. The 'Experiment Metadata' form is filled out with fields for 'Compound' (JSL771142-Paracetamol), 'Date' (2006-09-05), 'Time' (14:00), and 'Location' (JSL). The 'Save for Later' button is visible at the bottom right.

A General Chemistry Laboratory Repository

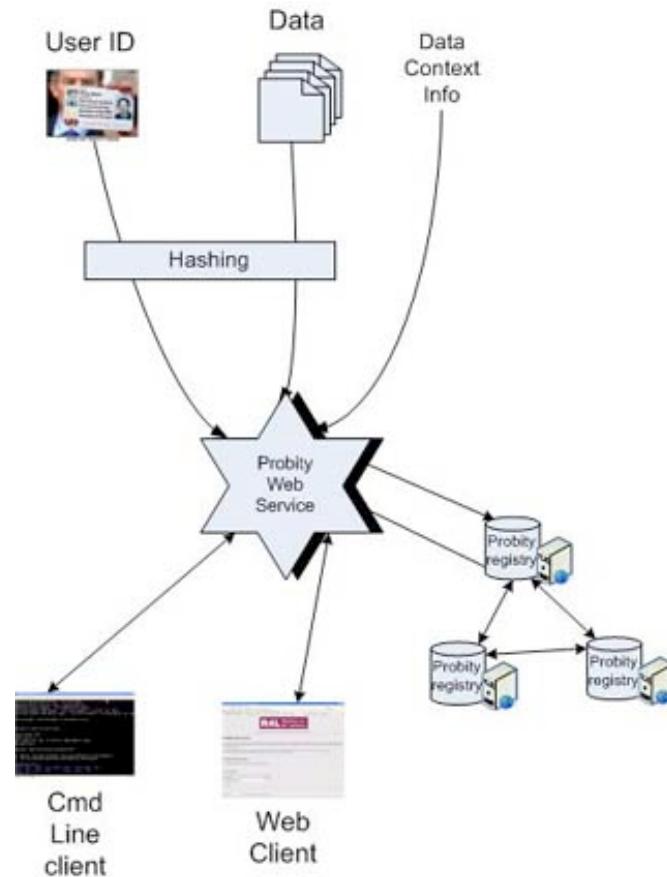
- Probit: A process to assert originality of a data record

Register a new claim

File for user ID Hash

File data Hash

Data Context Information



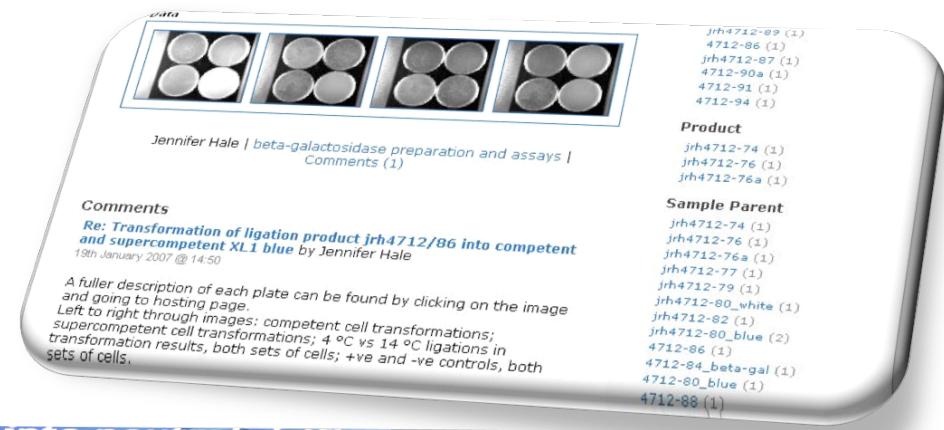
Search / Browse



Analysis & Discussion: Blogging Experiments

A repository can...

- Allow one to put, store and get
- Provide search and browse functionality
- **NOT** provide the presentation and discussion functions essential to working up a scientific study



Investigations into neutral drift

Test PCR (different primer, dNTP and enzyme stocks)
16th May 2008 @ 14:26

Post Type: PCR
Risk assessment: Mutagenesis/PCR risk assessment

Reaction	DNA template	µL Water	µL Buffer	µL Primer 1	µL Primer 2	µL dNTPs	µL Miscellaneous	µL Enzyme	µL	Product
1	Purified plasmid from the X-glu positive colonies in 5025/46 (experiment round 4)	Sterile filtered 13.05.08	8.8 GoTaq Buffer (04/08)	4 Beta-glu fwd	1 Beta-glu rev	Normal made 04/08	2 MgCl ₂	1.2 GoTaq (04/08)	1 (diluted 1in first)	Test PCR (different primer, dNTP and enzyme stocks) product 1
2	Purified plasmid from the X-glu positive colonies in 5025/46 (experiment round 4)	Sterile filtered 13.05.08	8.8 GoTaq Buffer (04/08)	4 Beta-glu fwd 04/08	1 Beta-glu rev 04/08	Normal made 04/08	2 MgCl ₂	1.2 GoTaq (04/08)	1 (diluted 1in first)	Test PCR (different primer, dNTP and enzyme stocks) product 2
3	Purified plasmid from the X-glu positive colonies in 5025/46 (experiment round 4)	Sterile filtered 13.05.08	8.8 GoTaq Buffer (04/08)	4 β-glu fwd (13.05.08)	1 β-glu rev (13.05.08)	Normal made 04/08	2 MgCl ₂	1.2 GoTaq (04/08)	1 (diluted 1in first)	Test PCR (different primer, dNTP and enzyme stocks) product 3
4	None	Sterile filtered 13.05.08	9.8 GoTaq Buffer (04/08)	4 β-glu fwd (13.05.08)	1 β-glu rev (13.05.08)	Normal made 04/08	2 MgCl ₂	1.2 GoTaq (04/08)	1 (diluted 1in first)	Test PCR (different primer, dNTP and enzyme stocks) product 4
	Purified plasmid from the X-glu positive colonies in 5025/46 (experiment)	Sterile filtered 13.05.08	8.8 GoTaq Buffer (04/08)	4 Beta-glu fwd	1 Beta-glu fwd	Normal made 13.05.08	2 MgCl ₂	1.2 GoTaq (04/08)	1 (diluted 1in first)	Test PCR (different primer, dNTP and enzyme stocks)

Search Find

Archives

- May 2008 (58)
- April 2008 (77)
- March 2008 (72)
- February 2008 (113)
- January 2008 (67)
- December 2007 (61)
- November 2007 (181)
- October 2007 (323)
- September 2007 (181)
- August 2007 (67)
- July 2007 (89)
- June 2007 (123)
- May 2007 (128)
- April 2007 (98)
- March 2007 (116)

Sections

- Materials (71)
- Notes (31)
- Procedure (338)
- Product (1274)
- Safety (11)
- Templates (29)

Post Type

- Digestion (31)
- Enzyme (15)
- Cell_strain (28)
- Buffer (18)
- Primer (11)
- Thermocycler_program (1)
- Template (27)

Facilitating Research

- Enables 'geographically distributed collaborative research'
- Can be open or private
- A useful approach for sharing 'failed' experiments?

PCR of beta-galactosidase third attempt by Jennifer Hale
14th December 2006 @ 11:10

Unfortunately the purification appears not to have gone well. Though I also can't get any consistency from the figures given by the nano-drop. These are the results I got:

-	reading 1	reading 2	reading 3	reading 4	reading 5	reading 6	average
PCR product before*	282.3 ng/µL	283.4 ng/µL	281.1 ng/µL	N/A	N/A	N/A	282.3 ng/µL
PCR product after*	7.8 ng/µL	12.9 ng/µL	17.6 ng/µL	85.4 ng/µL	22.4 ng/µL	12.8 ng/µL	?

*Both reactions combined together after PCR

I'm going to do another PCR again. That step is working really well. I'm just not sure what to do about purifying it. The only other thing I can try is eluting in TE buffer rather than water (which it says you can also elute into)

In this purification I used preheated water and followed the instructions closely. Perhaps the DNA will elute into TE more effectively.

Re: PCR of beta-galactosidase third attempt by David Neylon
14th December 2006 @ 18:32

I would definitely compare these on a gel so as to see whether it is just the nanodrop that is the problem. It might help also if you are explicit about how much solution you are trying to purify and what the final volume is.

4712-86 (1)
4712-84_beta-ga
4712-80_blue (1)
4712-88 (1)

Sample Parent2
jrh4712-80_blue (1)
jrh4712-80_white (2)
4712-84_pBad (1)
4712-80_white (1)

Sample Parent 3
4712-74 (1)

Search

Find

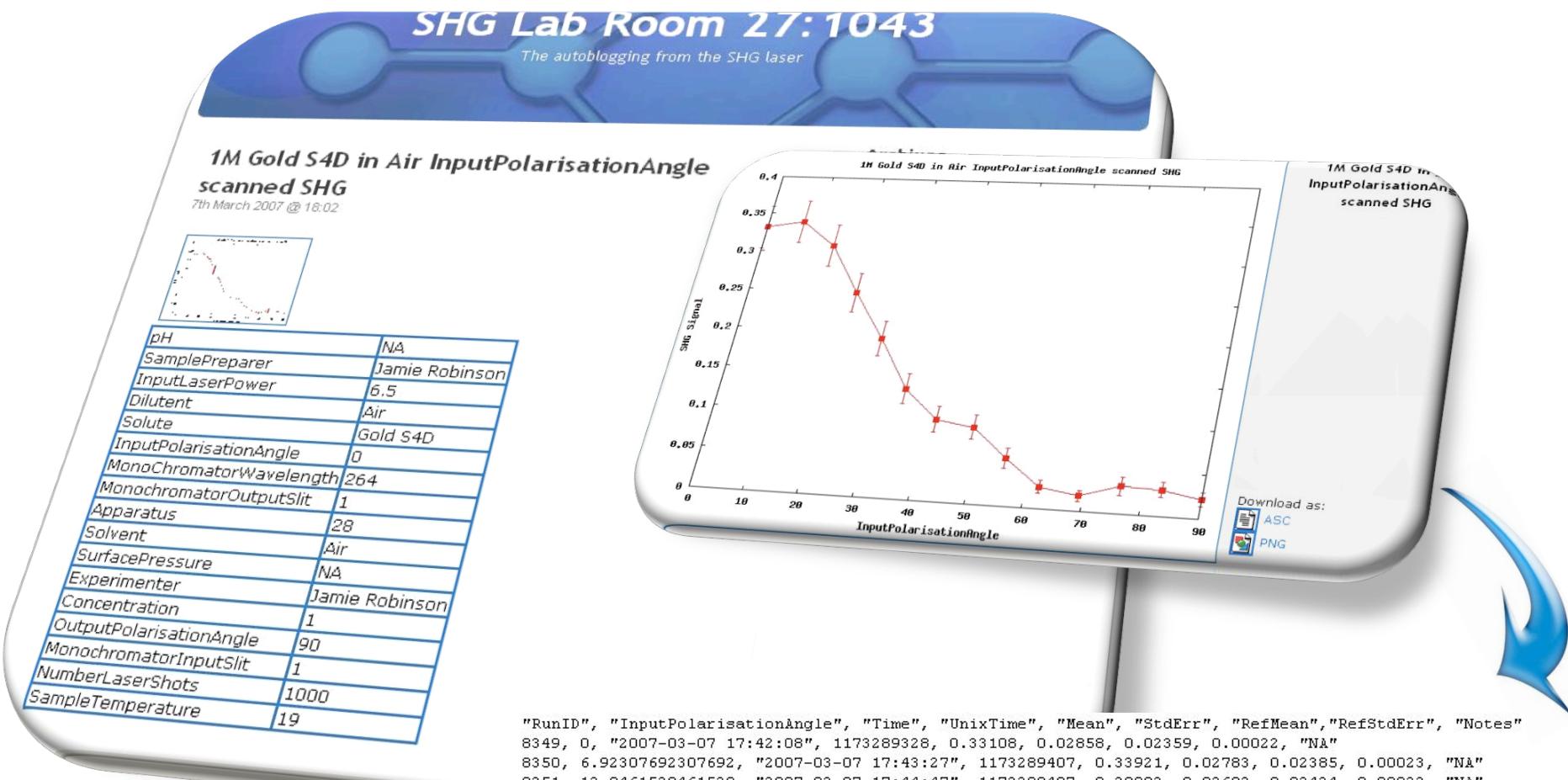
Links

Admin

New Post

Live Copy

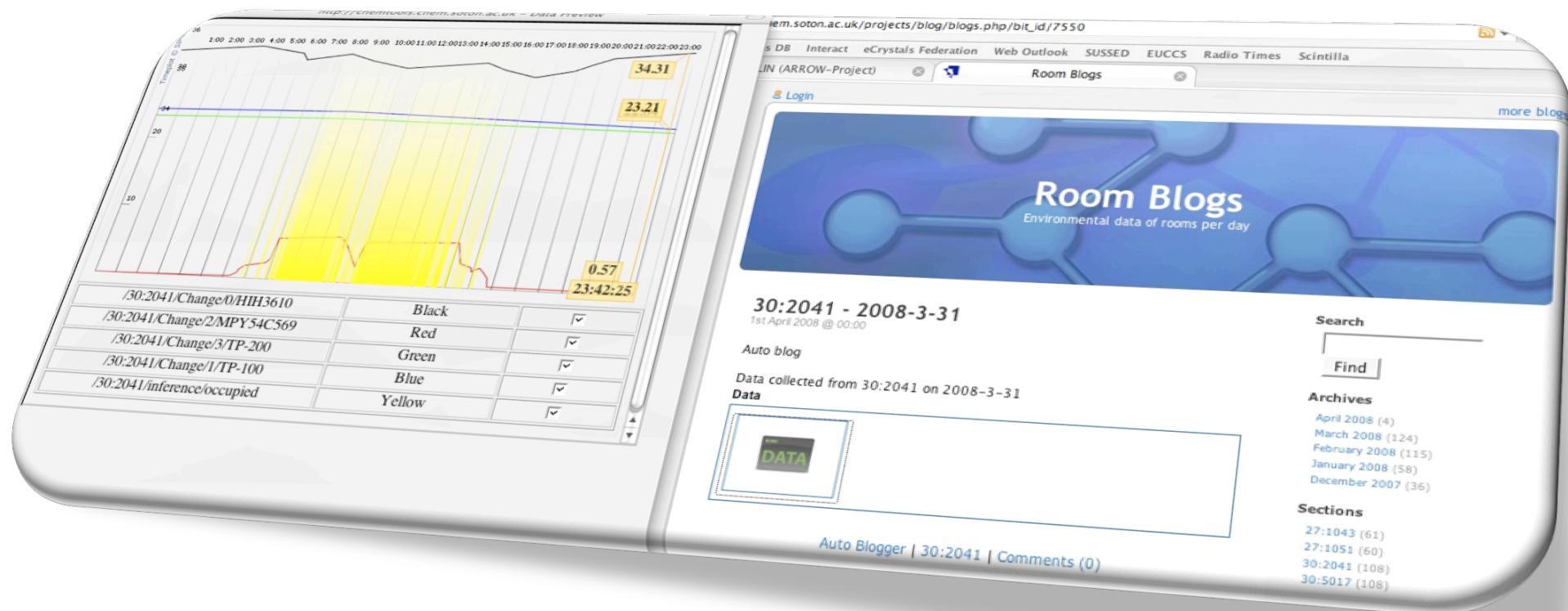
Automatic Blogging by Machines



```
"RunID", "InputPolarisationAngle", "Time", "UnixTime", "Mean", "StdErr", "RefMean", "RefStdErr", "Notes"  
8349, 0, "2007-03-07 17:42:08", 1173289328, 0.33108, 0.02858, 0.02359, 0.00022, "NA"  
8350, 6.92307692307692, "2007-03-07 17:43:27", 1173289407, 0.33921, 0.02783, 0.02385, 0.00023, "NA"  
8351, 13.8461538461538, "2007-03-07 17:44:47", 1173289487, 0.30893, 0.02692, 0.02424, 0.00023, "NA"  
8352, 20.7692307692307, "2007-03-07 17:46:06", 1173289566, 0.24921, 0.02485, 0.02396, 0.00023, "NA"  
8353, 27.6923076923076, "2007-03-07 17:47:26", 1173289646, 0.19186, 0.02197, 0.02372, 0.00022, "NA"  
8354, 34.6153846153845, "2007-03-07 17:48:45", 1173289725, 0.13147, 0.01842, 0.02377, 0.00022, "NA"  
8355, 41.5384615384614, "2007-03-07 17:50:05", 1173289805, 0.09684, 0.01536, 0.02407, 0.00023, "NA"  
8356, 48.4615384615383, "2007-03-07 17:51:24", 1173289884, 0.08984, 0.01470, 0.02406, 0.00023, "NA"  
8357, 55.3846153846152, "2007-03-07 17:52:43", 1173289963, 0.05578, 0.01183, 0.02456, 0.00023, "NA"  
8358, 62.3076923076921, "2007-03-07 17:54:03", 1173290043, 0.02520, 0.00770, 0.02401, 0.00023, "NA"  
8359, 69.230769230769, "2007-03-07 17:55:22", 1173290122, 0.01750, 0.00661, 0.02438, 0.00023, "NA"  
8360, 76.1538461538459, "2007-03-07 17:56:41", 1173290201, 0.03136, 0.00998, 0.02397, 0.00023, "NA"  
8361, 83.0769230769228, "2007-03-07 17:58:01", 1173290281, 0.03012, 0.00846, 0.02393, 0.00023, "NA"  
8362, 89.9999999999997, "2007-03-07 17:59:20", 1173290360, 0.02174, 0.00789, 0.02383, 0.00023, "NA"
```

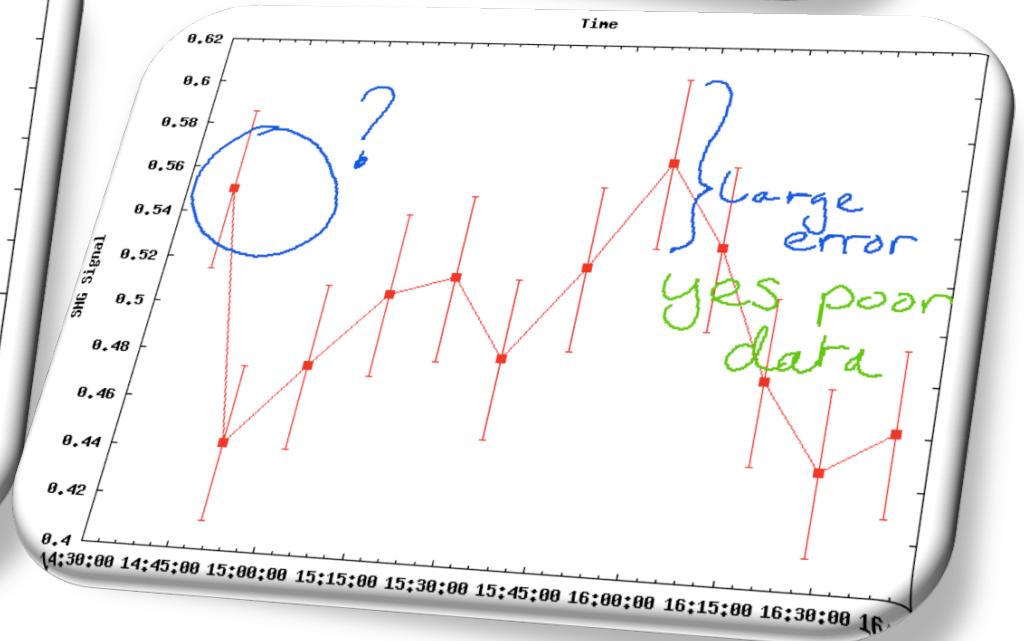
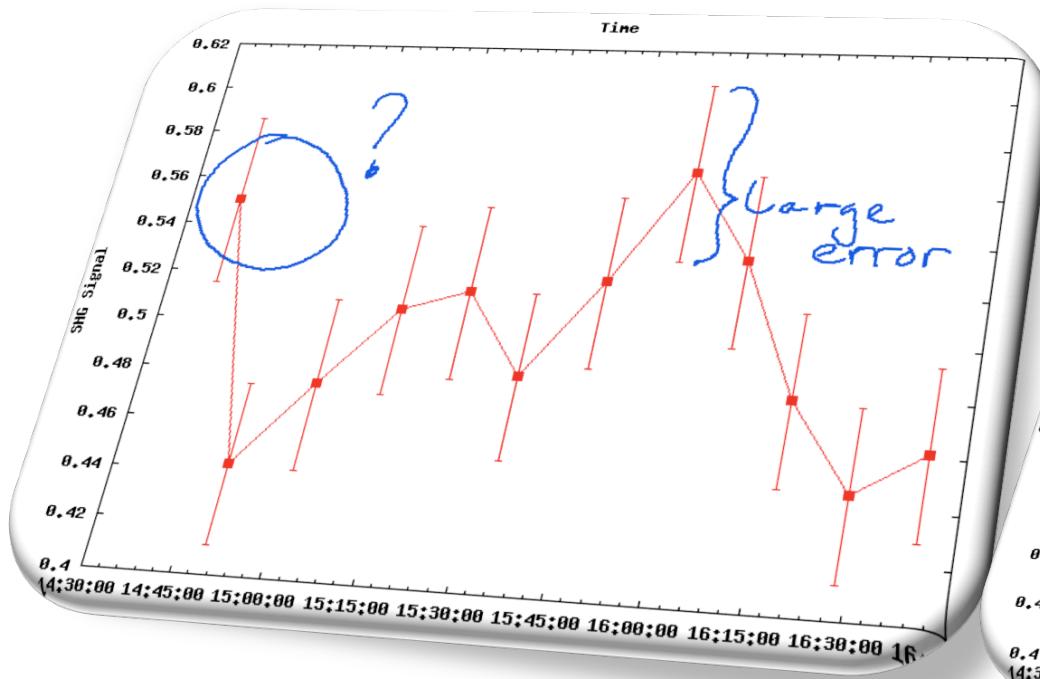
Automatic Blogging by Sensors

- Continuous log of 'environmental' conditions in a laboratory
- Instant detection of erroneous events
- Correlate with inconsistencies in datasets



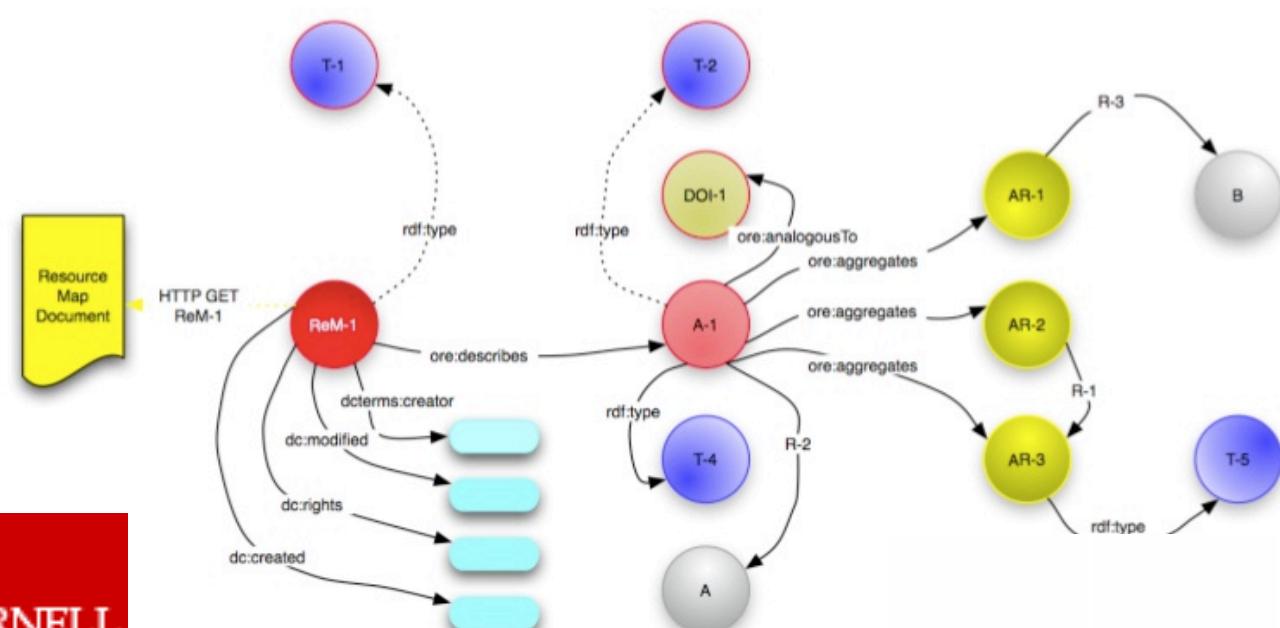
Comments and Collaborative Tools

- Annotation tools allow comments and foster collaboration and / or communication
- Need for more advanced Blog tools / technology around data



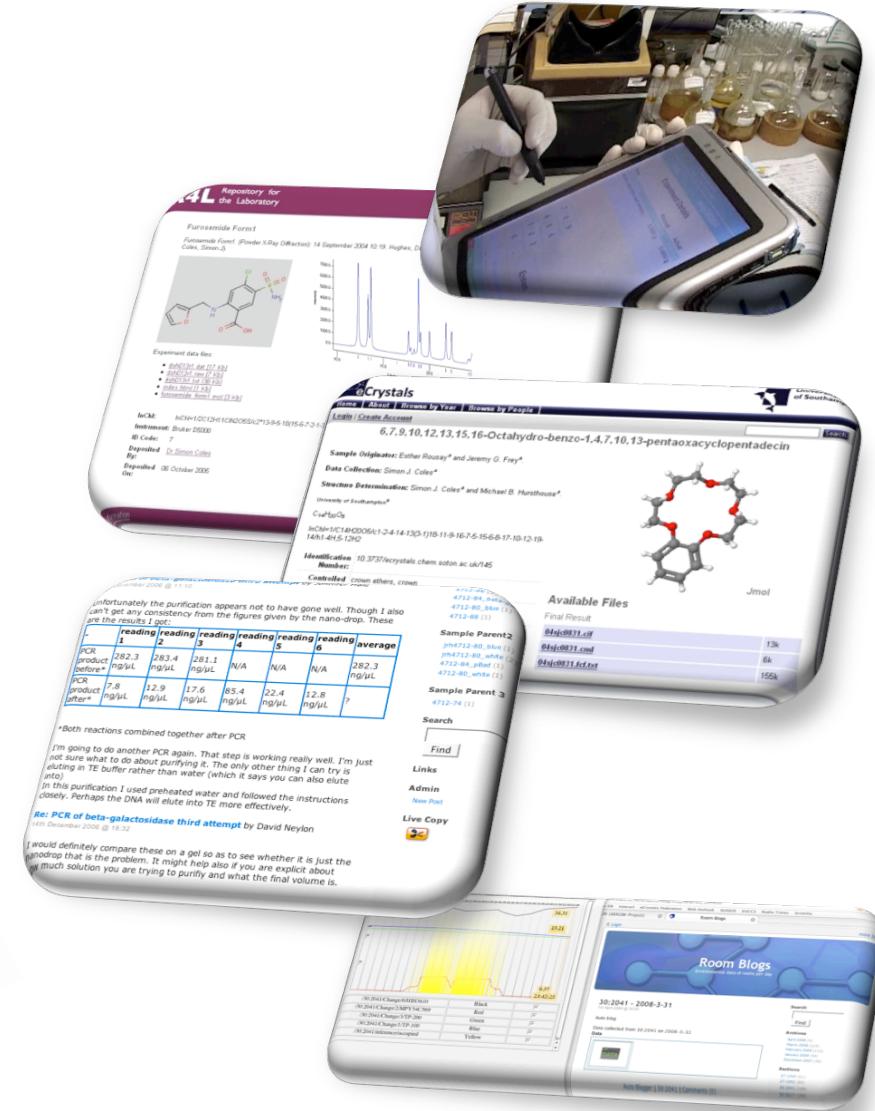
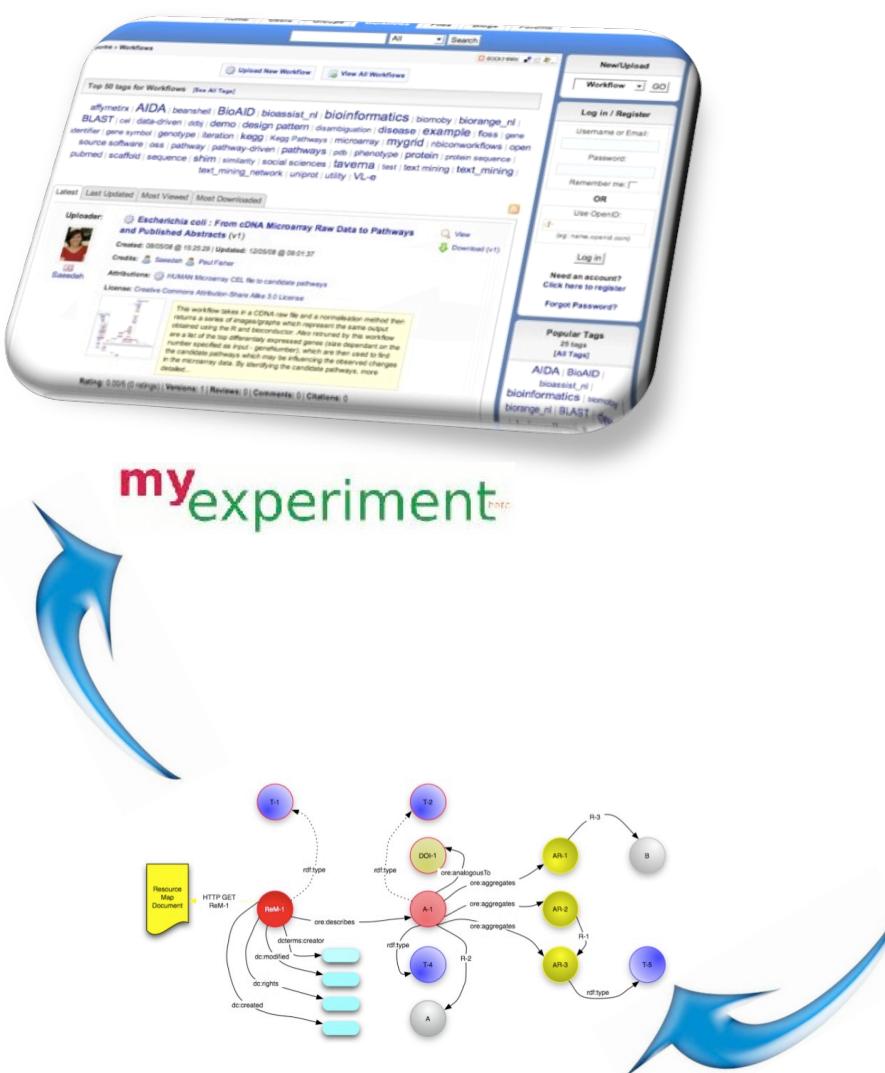
Packaging and Interoperability

- New moves in Digital Libraries community to enable distributed repositories to exchange content
- OAI-ORE (Open Archives Initiative – Object Reuse and Exchange)
- <http://www.openarchives.org/ore/>
- Describes an aggregation of objects in an exchangeable format
- Microsoft funded eChemistry testbed project



Microsoft

Towards a New Model for Chemical Information Exchange



A solid foundation for Open/Self-Publishing of Chemistry Data???

Thanks to:

- **Jeremy Frey**, Andrew Milsted, Richard Stephenson, Cameron Neylon, Jamie Robinson, Steven Wilson 
- **Dave DeRoure, Les Carr**, Chris Gutteridge, Tim Myles-Board, Arouna Woukei 
- **Liz Lyon**, Rachel Heery, Monica Duke, Michael Day, Traugott Koch, Manjula Patel 

