



Hypothesis-driven drug design using wiki-based collaborative tools

Graeme Robb
AstraZeneca, Alderley Park



- Philosophy Discussion
 - Hypothesis-Driven Design
 - A Solution from the Zeitgeist
 - Chemistry (finally)
 - Real Data; Real Value



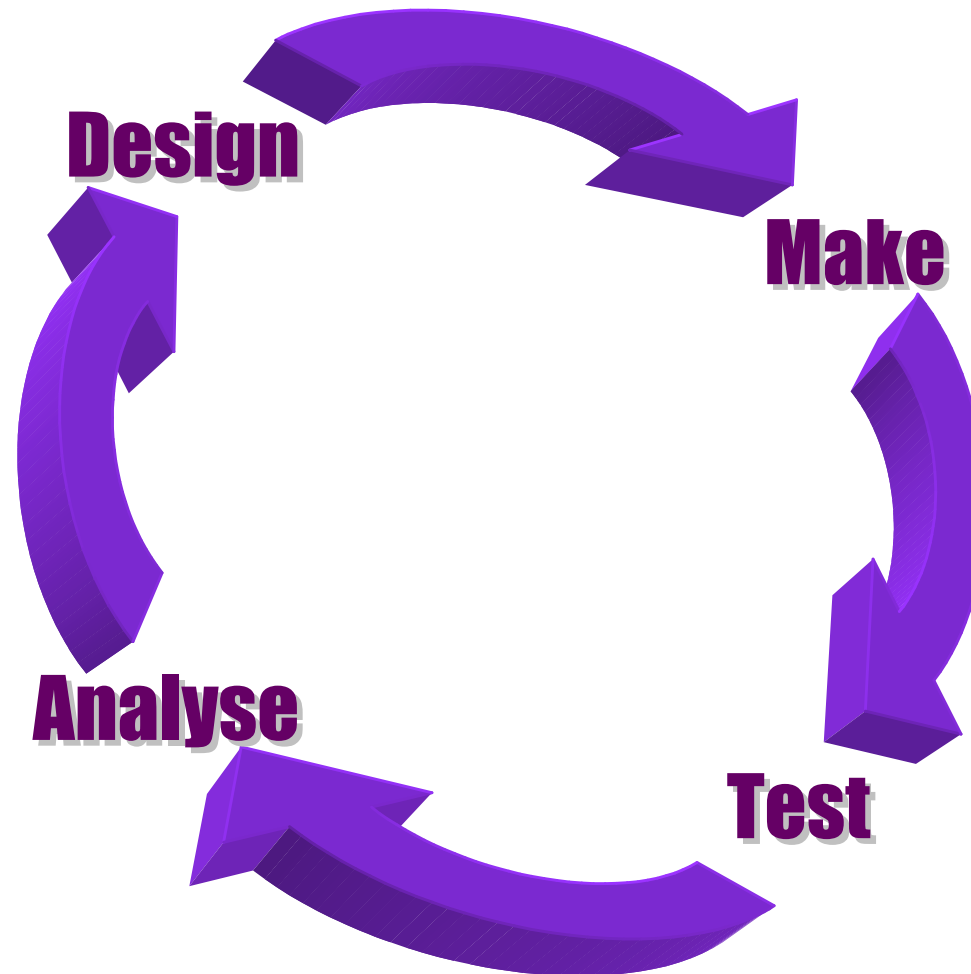
The Quantity Argument...

- Focus on: making larger numbers of compounds
- In a nutshell: 'there are more needles in a larger haystack'
- Constrained by: only using easy chemistry / available reagents
- Opportunity: rapid SAR exploration
- Risk: never explore 'hard' chemistry; never find best compounds

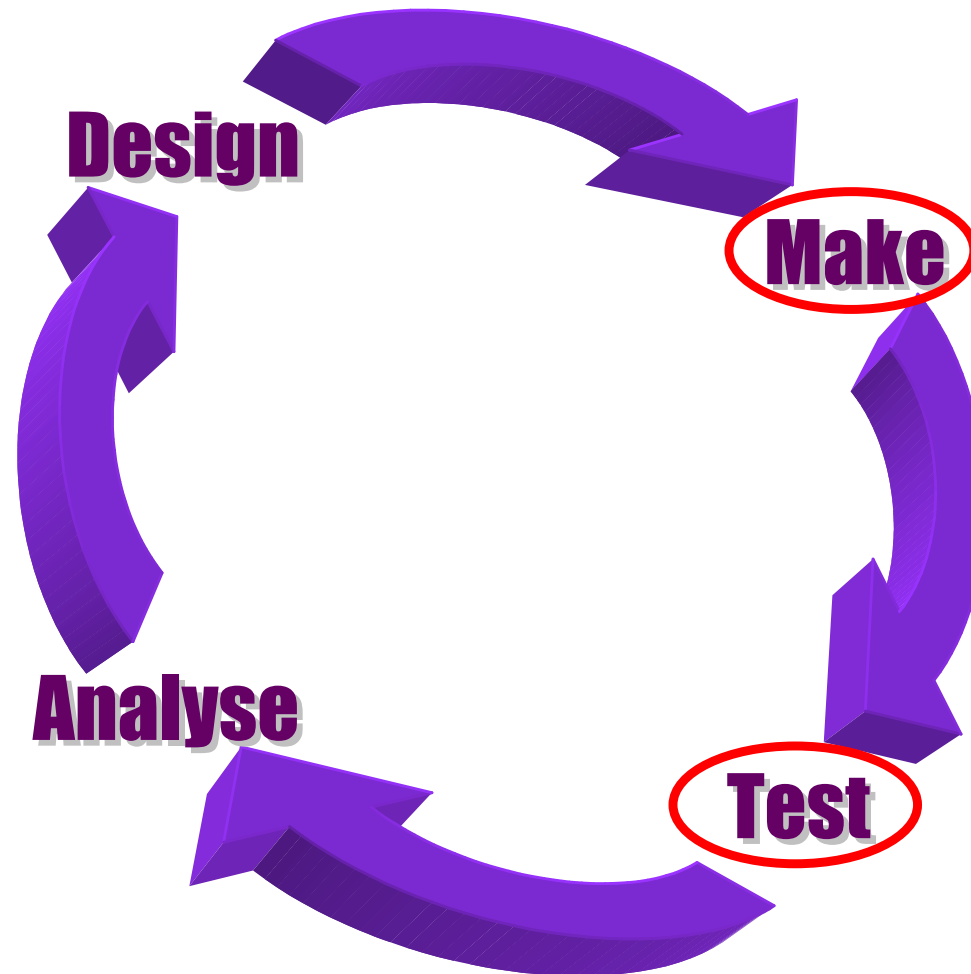
• The Quality Argument...

- Focus on: making the right compounds
- In a nutshell: 'using all our expertise we can do much better than chance'
- Constrained by: our imaginations!
- Opportunity: waste less on non-productive avenues
- Risk: bespoke synthesis means fewer compounds

The Circle of Progress



The Circle of Progress



Quantity

Puts bottleneck at 'Make', and 'Test' must keep up. Very resource intensive

Creative thinking and innovation are underused

The Circle of Progress



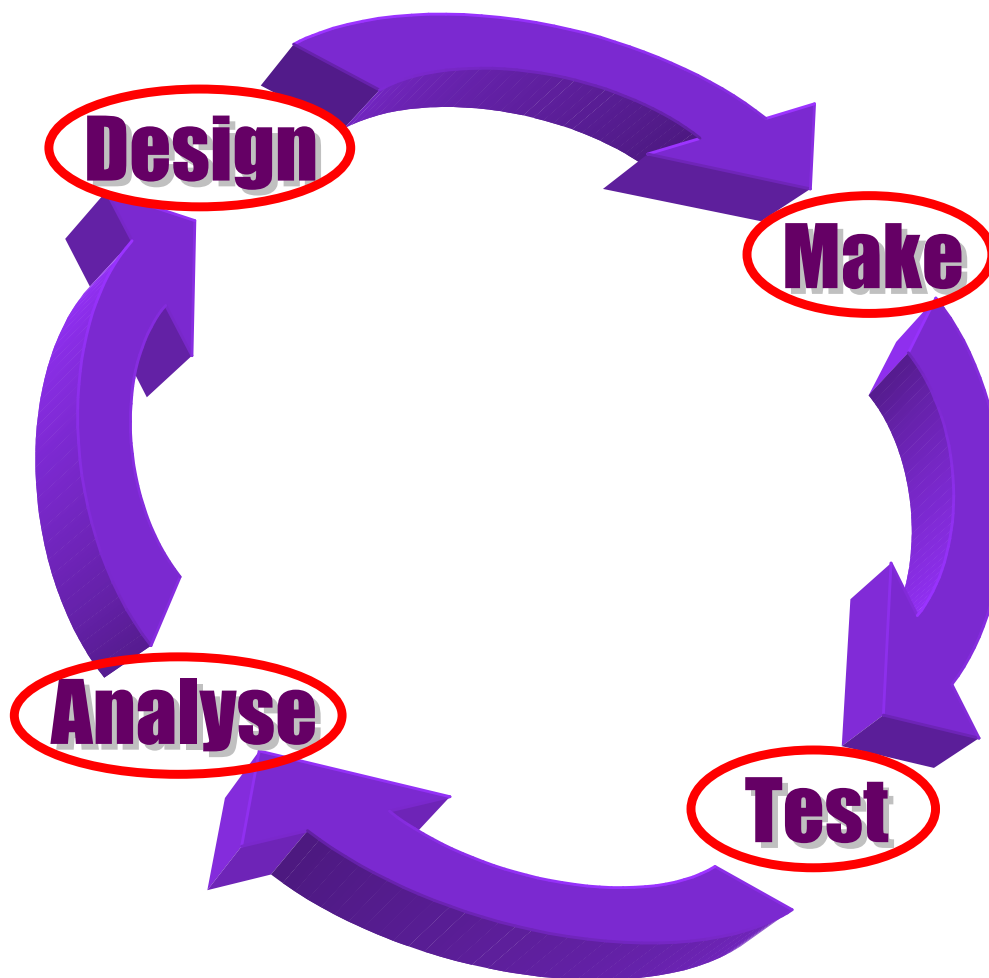
Quality

Puts emphasis on design, to pick the right compounds.

Analysis is key component of good design.

Make and Test still at capacity, but less resource intensive as fewer compounds.

No bottleneck!





- The best method (opinion!)
 - Certainly best for Lead Optimisation (LO)
 - May not always be appropriate for Lead Generation (LG)
 - Makes best use of limited make and test resources
- Have to ensure you make the right compounds
 - Take the shortest route from lead to CD (candidate drug)
 - Be systematic, and transparent about rationale
 - Only make key compounds to address issues and add maximum value

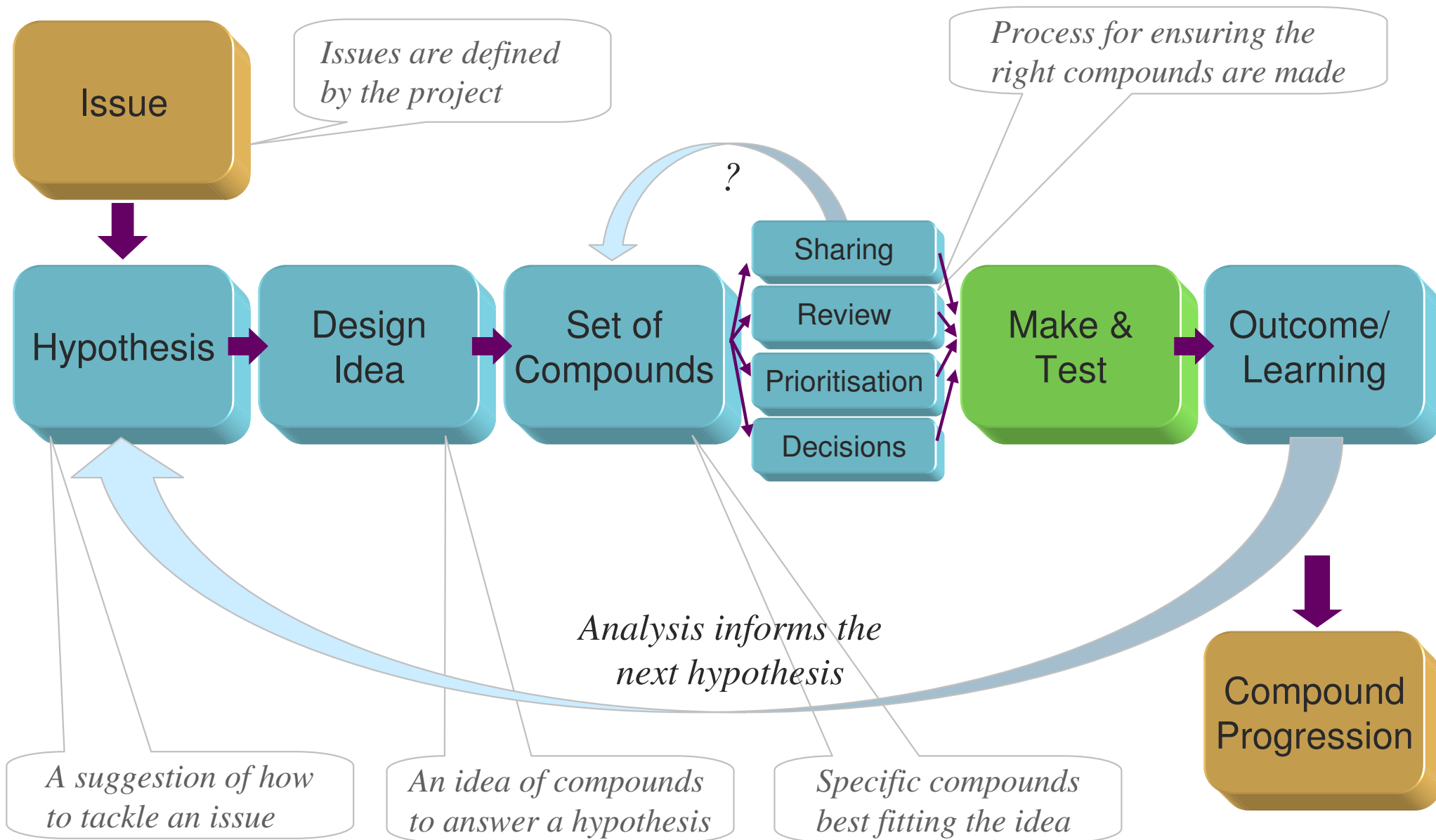
➔ Hypothesis Driven Drug Design

Be Hypothesis-Driven



- Compound design is focused on addressing specific issues
 - Ensures chemistry is aligned with project issues
 - Formulate a hypothesis of a potential solution
 - Design compounds to test this hypothesis
 - Analyse results to assess hypothesis → add to collective knowledge
- Lean Design
 - Make minimum number of compounds necessary to answer the question
 - Do not make compounds with predicted poor outcomes
 - Do not make a compound without a purpose
- Exponential success:
 - The more we *do*, the more we *understand*, the better we *design*.

Be Hypothesis-Driven





- Focus on designer role as key
 - “Quality of a compound is fixed at the moment of conception”
 - Need experts with time dedicated to design
 - Need multi-disciplines (esp. in LO) to see whole picture
 - Need to collaborate effectively
- Traditional methods are often entrenched
 - Move from individual to collective...
 - Move from ‘what *can* we make’ to ‘what *should* we make’...
 - We need a paradigm shift.....
 - We need...



In 1972, a crack chemistry team began making drugs. These men promptly decided on the best methods and stuck to them. Today, still using these outdated methods, they miss opportunities for innovative drug design. If you have this problem, if no one else can help, and if you can make them, maybe you should form...

A DESIGN -TEAM

Design Teams



- Now adopted across much of AstraZeneca
 - Especially in LO
- Multidisciplinary team of experts
 - Meet for regular, short meetings
 - Time allocated to design and analysis
- New way of working – need new tools
 - Need to communicate within design teams
 - Need to efficiently share / collaborate
 - Need to make decisions
 - Need to have transparency

An Old Solution



- There is nothing new under the sun
 - Some have tried this before...
 - But often fails, or effectively ends up with a team of one
- Problem is the tools
 - Shared spreadsheets and documents on a server or in email
 - People lose the links, don't read the email or can't be bothered to update
 - Multiple copies of data can get out of synchronisation
 - Files are not always accessible by everyone (esp. in meetings)
 - Data often hard to find, being 'hidden' inside files.
 - Not integrated into workflows
 - Requires extra effort → People just don't have time

A New Solution



- Take inspiration from the zeitgeist!
- Current growth in IT tools is on the internet
 - Social networking, online encyclopaedias, blogs, wikis, *etc.*
- “Web 2.0”
 - Facilitate communication and collaboration
 - Constantly updating
 - Interoperative – bringing data in from multiple source
 - Remix content (‘mashup’) for different purposes and tasks
 - Increased benefits from ever greater numbers of users
 - All online – no client software; works from any PC



- Compound Design Database
 - Based on Wiki / MySQL technologies
- **Why use Wiki?**
 - Web-based, so no client software to develop
 - Easy to prototype
 - Fast and easy to use
 - Easy to customise for our requirements
 - Can handle multiple simultaneous edits
 - Always up-to-date
 - Has all the history/version features needed
 - Can plug in oechem for added functionality
 - It's free (mediawiki)



A CoDD

- Each project uses a page in the wiki to record, discuss and comment upon design ideas
 - Common platform for all users
 - Single one-stop-shop gives access to all relevant data



my preferences my watchlist my contributions log out

article discussion edit history move watch

Pretend Project Compound Design Database

[Compound Design Database Feedback]

- [CONTENT SUMMARY](#)
- [VIRTUAL COMPOUND LIST](#)
- [PROGRESS CHART](#)
- [INTERACTIVE VIEW](#)
- [Tracking Summary](#)
- [User-Defined Cmpd Matches](#)
- [CoDD Training Manual](#)
- [Pretend matchable cmpds](#)
- [Pretend Design Team Output](#)
- [Install Canvas](#)
- [DMT Tool](#)
- [Testing Schedule](#)

On this page are detailed synthesis target ideas for the Pretend project, with links to appropriate data and comments. To add an idea, simply go to the appropriate section and click the *[edit]* tab. Or, if you're less confident about editing the page simply send an email to [Graeme Robb](#) with all the relevant information. Some useful links are:

- Linked files are all stored [here](#) in the eRoom.
- Excluding remakes, the highest target ID used so far is: **0 1 7 9**

Contents [hide]

1 Ideas Addressing Potency

1.1 Explore P1 pocket

main portal



- Each project uses a page in the wiki to record, discuss and comment upon design ideas
 - Common platform for all users
 - Single one-stop-shop gives access to all relevant data



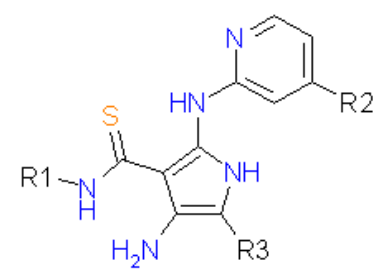
Ideas Addressing Potency [\[edit\]](#)

Explore P1 pocket [\[edit\]](#)

DesignSet Number One [\[edit\]](#)

Project Aim: Increase in-vitro potency

[DBview Link for ID=1](#)

Date <small>(xx/xx/xx)</small>	Generic Structure	Rationale	Status/Outcome
<p>CONCEPT 10/11/2008</p> <p>STARTED 15/11/2008</p> <p>COMPLETION</p>	 <p>Where R1 and R2 are small or polar groups.</p>	<p>Docking shows that there is a largely unexplored P1 pocket close to the ligand. Small increases in the size of the groups at R1 or R2 may better fill this pocket to give greater potency. {J.Bloggs}</p> <p>Polar groups may be preferred as these may pick up an additional interaction in the P1 pocket while keeping physprops in the right area. {C.Smith}</p> <p>This is all pretend and just an example of what you might write in order to share your ideas and form a basis for discussion. You can also include links to supporting data.</p>	<p>Priority = 2</p> <hr/> <p>Status = In Progress</p> <hr/> <p>Chemist = Geoff</p>

[Upload SMILES](#) [Chemist comments](#)

Designset
template



- Each project uses a page in the wiki to record, discuss and comment upon design ideas
 - Common platform for all users
 - Single one-stop-shop gives access to all relevant data



Classification hierarchy

Ideas Addressing Potency [edit]			
Explore P1 pocket [edit]			
DesignSet Number One [edit]			
<i>Project Aim: Increase in-vitro potency</i>			
DBview Link for ID=1			
Date <small>(xx/xx/xx)</small>	Generic Structure	Rationale	Status/Outcome
CONCEPT 10/11/2008	 <small>Where R1 and R2 are small or polar groups.</small>	Docking shows that there is a largely unexplored P1 pocket close to the ligand. Small increases in the size of the groups at R1 or R2 may better fill this pocket to give greater potency. {J.Bloggs}	Priority = 2
STARTED 15/11/2008		Polar groups may be preferred as these may pick up an additional interaction in the P1 pocket while keeping physprops in the right area. {C.Smith}	Status = In Progress
COMPLETION		This is all pretend and just an example of what you might write in order to share your ideas and form a basis for discussion. You can also include links to supporting data.	Chemist = Geoff
Upload SMILES		Chemist comments	

Designset template



- Each project uses a page in the wiki to record, discuss and comment upon design ideas
 - Common platform for all users
 - Single one-stop-shop gives access to all relevant data



Classification hierarchy

Representative structure

Ideas Addressing Potency [\[edit\]](#)

Explore P1 pocket [\[edit\]](#)

DesignSet Number One [\[edit\]](#)

Project Aim: Increase in-vitro potency

[DBview Link for ID=1](#)

Date <small>(xx/xx/xx)</small>	Generic Structure	Rationale	Status/Outcome
<small>CONCEPT</small> 10/11/2008	 <small>Where R1 and R2 are small or polar groups.</small>	Docking shows that there is a largely unexplored P1 pocket close to the ligand. Small increases in the size of the groups at R1 or R2 may better fill this pocket to give greater potency. {J.Bloggs}	<i>Priority = 2</i>
<small>STARTED</small> 15/11/2008		Polar groups may be preferred as these may pick up an additional interaction in the P1 pocket while keeping physprops in the right area. {C.Smith}	<i>Status = In Progress</i>
<small>COMPLETION</small>		This is all pretend and just an example of what you might write in order to share your ideas and form a basis for discussion. You can also include links to supporting data.	<i>Chemist = Geoff</i>

[Upload SMILES](#) [Chemist comments](#)

Designset template



- Each project uses a page in the wiki to record, discuss and comment upon design ideas
 - Common platform for all users
 - Single one-stop-shop gives access to all relevant data



Classification hierarchy

Representative structure

Compound upload viewing links

Ideas Addressing Potency [\[edit\]](#)

Explore P1 pocket [\[edit\]](#)

DesignSet Number One [\[edit\]](#)

Project Aim: Increase in-vitro potency

[DBview Link for ID=1](#)

Date <small>(xx/xx/xx)</small>	Generic Structure	Rationale	Status/Outcome
<small>CONCEPT</small> 10/11/2008	<p>Where R1 and R2 are small or polar groups.</p>	Docking shows that there is a largely unexplored P1 pocket close to the ligand. Small increases in the size of the groups at R1 or R2 may better fill this pocket to give greater potency. {J.Bloggs}	Priority = 2
<small>STARTED</small> 15/11/2008		Polar groups may be preferred as these may pick up an additional interaction in the P1 pocket while keeping physprops in the right area. {C.Smith}	Status = <i>In Progress</i>
<small>COMPLETION</small>		This is all pretend and just an example of what you might write in order to share your ideas and form a basis for discussion. You can also include links to supporting data.	Chemist = <i>Geoff</i>

[Upload SMILES](#) [Chemist comments](#)

Designset template

A CoDD

- Each project uses a page in the wiki to record, discuss and comment upon design ideas
 - Common platform for all users
 - Single one-stop-shop gives access to all relevant data



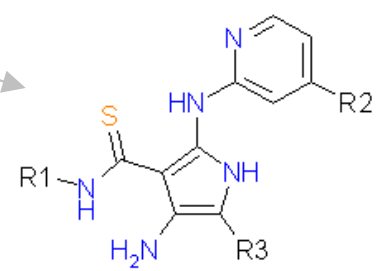
Classification hierarchy

Representative structure

Compound upload viewing links

Data to help track progress, decisions, responsibility

Designset template

Date (xx/xx/xx)	Generic Structure	Rationale	Status/Outcome
CONCEPT 10/11/2008		Docking shows that there is a largely unexplored P1 pocket close to the ligand. Small increases in the size of the groups at R1 or R2 may better fill this pocket to give greater potency. {J.Bloggs} Polar groups may be preferred as these may pick up an additional interaction in the P1 pocket while keeping physprops in the right area. {C.Smith}	Priority = 2 Status = In Progress Chemist = Geoff
STARTED 15/11/2008		This is all pretend and just an example of what you might write in order to share your ideas and form a basis for discussion. You can also include links to supporting data.	
COMPLETION			

Where R1 and R2 are small or polar groups.

Upload SMILES

Chemist comments

A CoDD

- Each project uses a page in the wiki to record, discuss and comment upon design ideas
 - Common platform for all users
 - Single one-stop-shop gives access to all relevant data



Classification hierarchy

Representative structure

Compound upload viewing links

Data to help track progress, decisions, responsibility

Explanation, discussion, comments

Designset template

Date (xx/xx/xx)	Generic Structure	Rationale	Status/Outcome
<i>CONCEPT</i> 10/11/2008	 <i>Where R1 and R2 are small or polar groups.</i>	Docking shows that there is a largely unexplored P1 pocket close to the ligand. Small increases in the size of the groups at R1 or R2 may better fill this pocket to give greater potency. {J.Bloggs} Polar groups may be preferred as these may pick up an additional interaction in the P1 pocket while keeping physprops in the right area. {C.Smith} This is all pretend and just an example of what you might write in order to share your ideas and form a basis for discussion. You can also include links to supporting data.	Priority = 2 Status = <i>In Progress</i> Chemist = <i>Geoff</i>

Upload SMILES

Chemist comments

A CoDD

- Each project uses a page in the wiki to record, discuss and comment upon design ideas
 - Common platform for all users
 - Single one-stop-shop gives access to all relevant data



Classification hierarchy

Representative structure

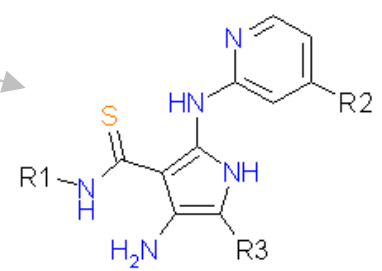
Compound upload viewing links

Designset template

Data to help track progress, decisions, responsibility

Explanation, discussion, comments

Link out to other data

Date (xx/xx/xx)	Generic Structure	Rationale	Status/Outcome
<i>CONCEPT</i> 10/11/2008		Docking shows that there is a largely unexplored P1 pocket close to the ligand. Small increases in the size of the groups at R1 or R2 may better fill this pocket to give greater potency. {J.Bloggs} Polar groups may be preferred as these may pick up an additional interaction in the P1 pocket while keeping physprops in the right area. {C.Smith}	Priority = 2 Status = <i>In Progress</i> Chemist = <i>Geoff</i>
<i>STARTED</i> 15/11/2008		This is all pretend and just an example of what you might write in order to share your ideas and form a basis for discussion. You can also include links to supporting data.	
<i>COMPLETION</i>			

Where R1 and R2 are small or polar groups.

Upload SMILES

Chemist comments



- But more than just a wiki....
 - Wiki is good for managing the editing
 - Less good for visualisation, data- or structure-searching
 - So extract and reformat data from wiki mysql tables
 - Merge with in-house data on real compounds

- Display anyway you like





- But more than just a wiki....
 - Wiki is good for managing the editing
 - Less good for visualisation, data- or structure-searching
 - So extract and reformat data from wiki mysql tables
 - Merge with in-house data on real compounds



- Display anyway you like

The screenshot shows a web browser window with a URL bar containing 'http://apoc11.rd.astrazeneca.net/az_dbview/az_dbview.py/display/mid=...'. The page content includes a chemical structure on the left and a table of results on the right. The chemical structure is a 5-membered ring with a sulfur atom, two nitrogen atoms, and three substituents labeled R1, R2, and R3. The table has columns for ID, IDEA, STATUS, CPDS, MADE, and CHEMIST. The table shows 21 records out of 40 of 174 total records.

ID	IDEA	STATUS	CPDS	MADE	CHEMIST
136		Parked	None	None	None
132		Parked	None	None	None
129		Analysis Complete	3	2	None
128		Parked	None	None	None
120		Analysis Complete	2	1	None
116		Parked	8	0	None
110		Analysis Complete	12	11	None
111		Parked	6	0	None
108		Analysis Complete	3	3	None
101		Analysis Complete	2	1	Team
98		Parked	6	0	None
94		Analysis Complete	4	4	None
86		Parked	2	0	Team
71		Parked	None	None	Unassigned
68		Parked	None	None	Unassigned
64		Analysis Complete	2	2	None
66		Parked	None	None	Unassigned
67		Parked	None	None	Unassigned
1		Analysis Complete	23	23	Chris Sheldon
36		Parked	7	1	Jane Moore



- But more than just a wiki....
 - Wiki is good for managing the editing
 - Less good for visualisation, data- or structure-searching
 - So extract and reformat data from wiki mysql tables
 - Merge with in-house data on real compounds



- Display anyway y

The screenshot shows a web browser displaying a table of chemical compounds. The table has columns for <uID>, <_smiles>, <Compound_Name>, <Reg_Date>, <CoDD_ID>, <Target_Details>, <Dates>, <Status>, <Priority>, and <Risk>. The table contains three rows of data, each with a chemical structure and a status. The status bar at the top of the table includes buttons for All, New, In Design, Design Locked, Work in Progress, Synthesis Complete, Analysis Complete, and (Parked).

<uID>	<_smiles>	<Compound_Name>	<Reg_Date>	<CoDD_ID>	<Target_Details>	<Dates>	<Status>	<Priority>	<Risk>
0195-0001				195	from: Combination Matrix		Design Locked	1	
0195-0002				195	from: Combination Matrix		Design Locked	1	
0182-0001		AZ-ID_1234 Direct	2009-04-08	182	from: Combination Matrix		Synthesis Complete	4	M

A CoDD

- But more than just a wiki....

- Wiki is good for managing the editing
- Less good for visualisation, data or structure searching
- So extract and reformat data from
- Merge with in-house data on real

- Display anyway

The screenshot displays the AstraZeneca CoDD system interface. It features a grid of chemical structures, each with associated data fields such as AZ_No, IBIS, bind_IC50, Avail_mg, uID, and yyyy-mm-dd. The interface includes navigation controls like 'First', 'Prev', 'Next', and 'Last'. A table at the bottom shows analysis results for various compounds.

Analysis	Complete	Count	Assigned To
Analysis Complete 2	2	None	None
Parked	None	None	Unassigned
Parked	None	None	Unassigned
Analysis Complete 23	23	Chris Sheldon	
Parked	7	1	Jane Moore

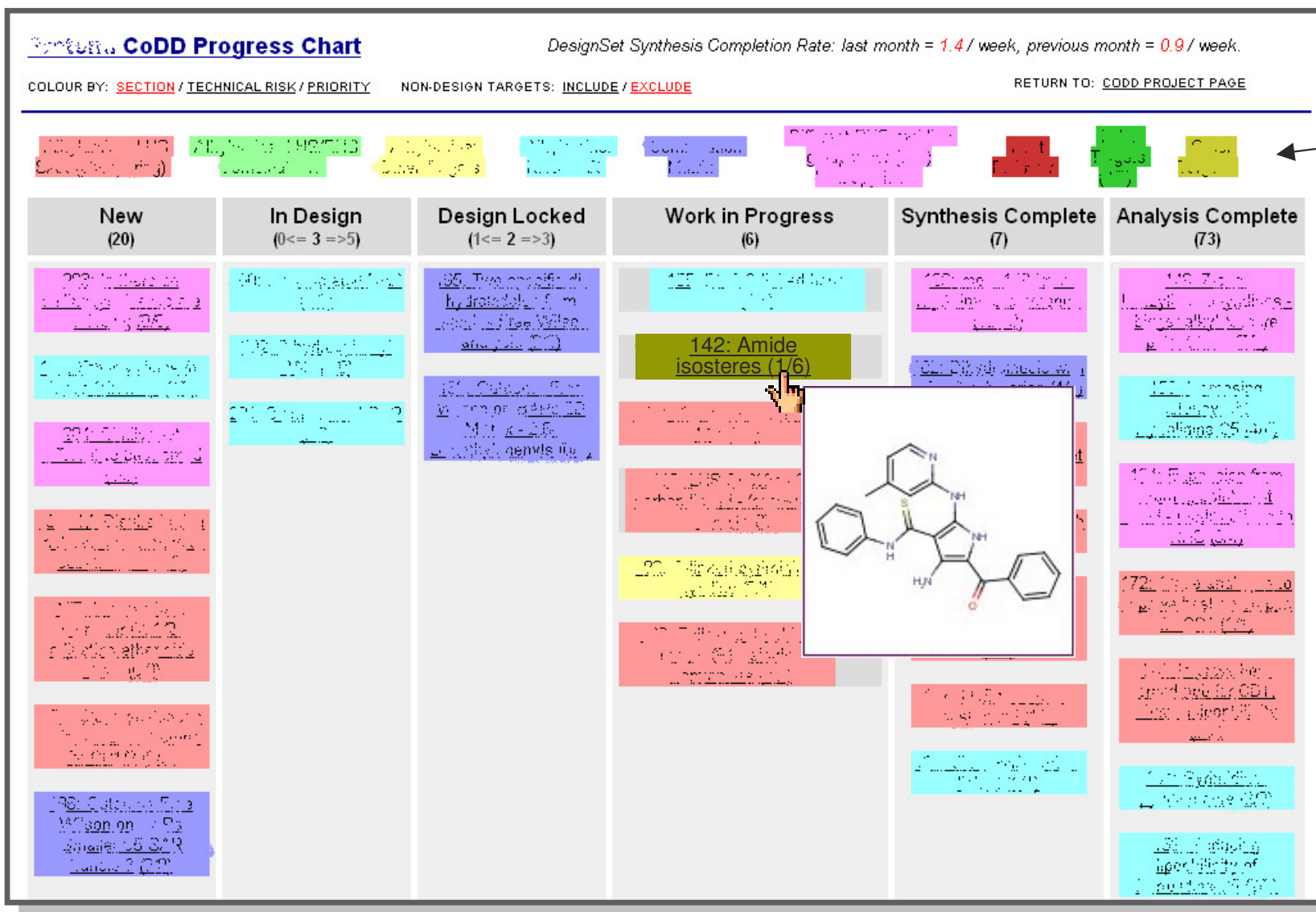


- Crucially, the system does not 'do' design for you
 - Only keeps track and organises
 - Make things easier to find / share
 - Saves time
- Frees up more time for being creative
 - Do better design
 - Do more thorough analysis
- Linked in to all parts of the D-M-T-A cycle
 - Track all designs for easy decision making
 - Assign synthesis and testing for compounds from the tool
 - Alerting when data is ready for analysis



Visual Planning

- Progress chart view for visual planning



Colour scheme indicates different areas of chemistry here.

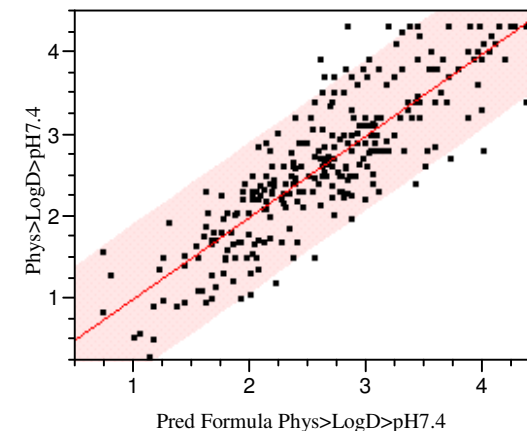


- A powerful computational chemistry platform
 - Not just an organisational tool
- Contains a database of all past, present and future compounds
 - Huge scope for predictive modelling
 - Can highlight 'good' compounds for rapid advancement
 - Can caution against 'bad' compounds for possible problems
 - Perform well in advance of decision to sanction synthesis
- Using 'Web2.0' principles
 - Capture more data → inform models → better prediction

Property Prediction

- QSAR modelling using descriptor based PLS

- Nothing new about that
- Some drawbacks
 - Imperfect descriptors give only locally correct models
 - Errors increase for compounds dissimilar to the training set



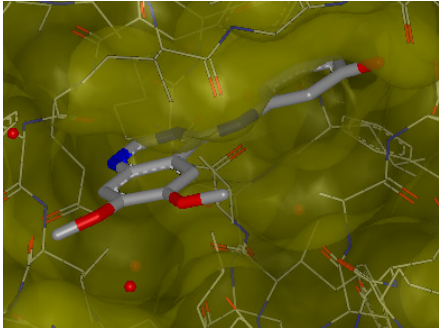
- But here designs are generally following up on previous compounds
 - All previous compounds in the training set
 - Local model is appropriate (especially in LO)

- Accurate predicted properties for compounds returned shortly after being entered in the tool.
 - Lipophilicity, Solubility, Plasma Protein Binding, *etc*
 - Use these to inform progression decisions

Auto-Docking and Visualisation



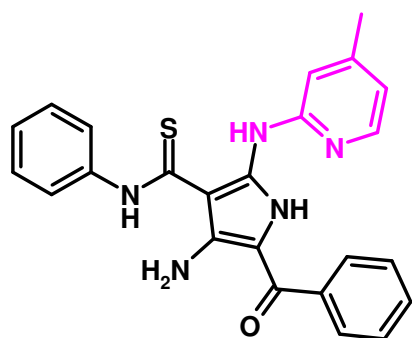
- Same principle applies to docking
 - Expert-created restricted docking script
 - Ensures you get the 'right' answer or none
 - Many docking programs output multiple scores per pose
 - Fit the pIC50 using docking scores as descriptors
 - Can give reasonable potency prediction for new compounds
 - Same caveats about being close to training set apply



Results	SimpleQuery	LingoQuery	ExpertQuery	Description
records 1 - 20 of 46				
_id	predicted_IC50	Clip_Score	Overlap_Tanimoto	Excluded_Vol_Overlap
<input checked="" type="radio"/>	0.393298	-0.378542	0.779031	18.2969
<input type="radio"/>	0.437164	-0.438254	0.708249	25.2656
<input type="radio"/>	0.402665	-0.3682	0.715103	22.6562
<input type="radio"/>	0.448417	-0.308775	0.425597	112.891
<input type="radio"/>	0.392631	-0.378542	0.747905	37.8281
<input type="radio"/>	0.389121	-0.438881	0.736518	45.0469
<input type="radio"/>	0.385311	-0.372662	0.743234	45.0781
<input type="radio"/>	0.471264	-0.241588	0.452884	130.969
<input type="radio"/>	0.0907513	-0.435477	0.55816	92.8906
<input type="radio"/>	0.0907513	-0.435477	0.55816	92.8906
<input type="radio"/>	0.439906	-0.45123	0.728426	30.0
<input type="radio"/>	0.436308	-0.481209	0.664578	32.1875
<input type="radio"/>	0.442979	-0.436916	0.73138	33.1562
<input type="radio"/>	0.0907513	-0.435477	0.55816	92.8906
<input type="radio"/>	0.399246	-0.428782	0.725042	41.5312
<input type="radio"/>	0.395849	-0.246969	0.722253	46.4062
<input type="radio"/>	0.392665	-0.373326	0.727806	46.0938

Toxophore Alerting

- We have a wealth of experience identifying and solving problems caused by toxic substructures ('toxophores')
 - But transfer of this knowledge is often poor and limited to post-rationalisation after (expensive) in-vivo findings.
 - Here we have begun encoding knowledge as precise toxophores with accompanying explanation and potential solutions
 - Every compound you could potentially make is screened and alerts emailed to design teams with explanation of the problem.

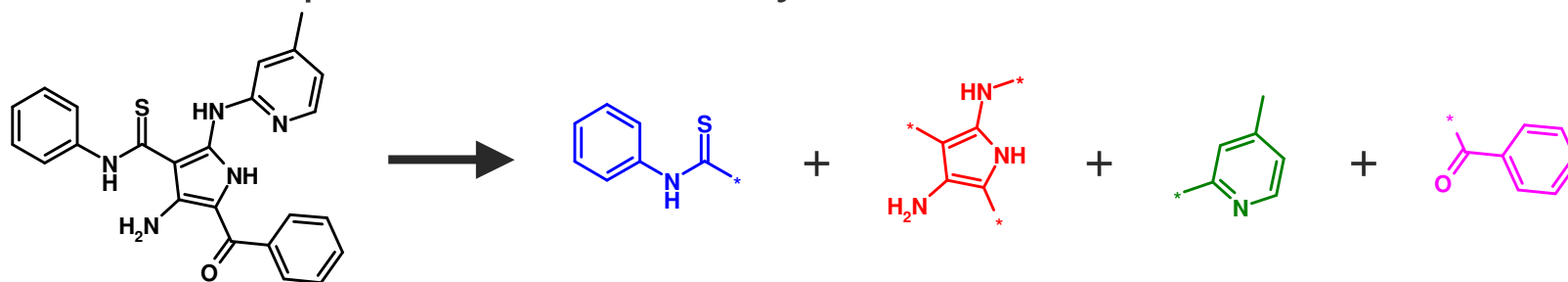


This is a compound from designset 201 and has been identified as containing a toxophore relating to **genotoxicity***. Possibly not every compound containing this substructure is toxic, but you may want to investigate further.

- CoDD link: [Pyridine Variations](#)
- Click for a [summary of why this is a toxophore](#).

New Compound Suggestions

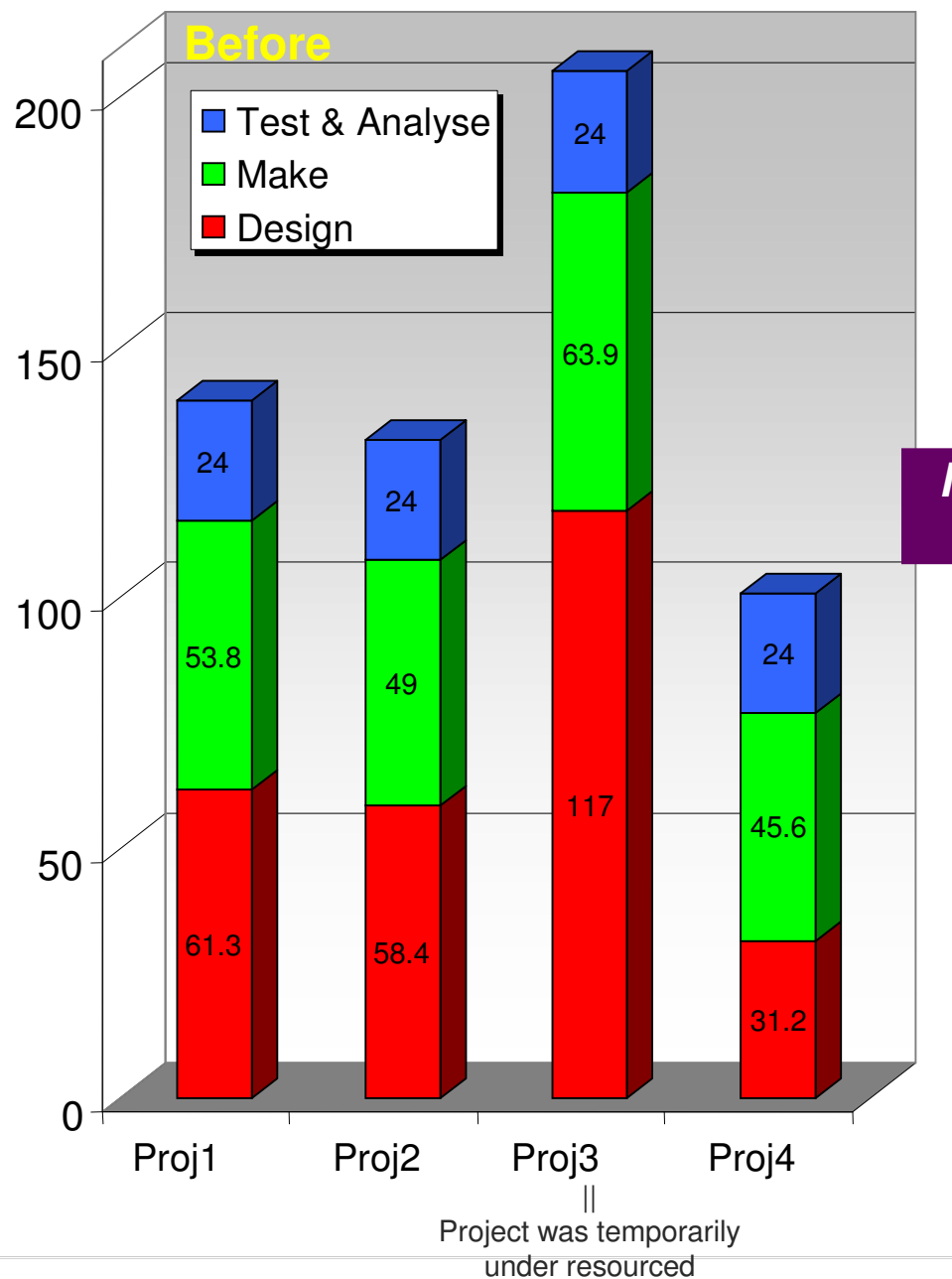
- Can accurately model compound properties using a Free-Wilson approach
 - Each group/fragment in a molecule contributes to a property
 - Whole molecule property is a sum of the group contributions
 - Makes assumption about additivity – *need to test*



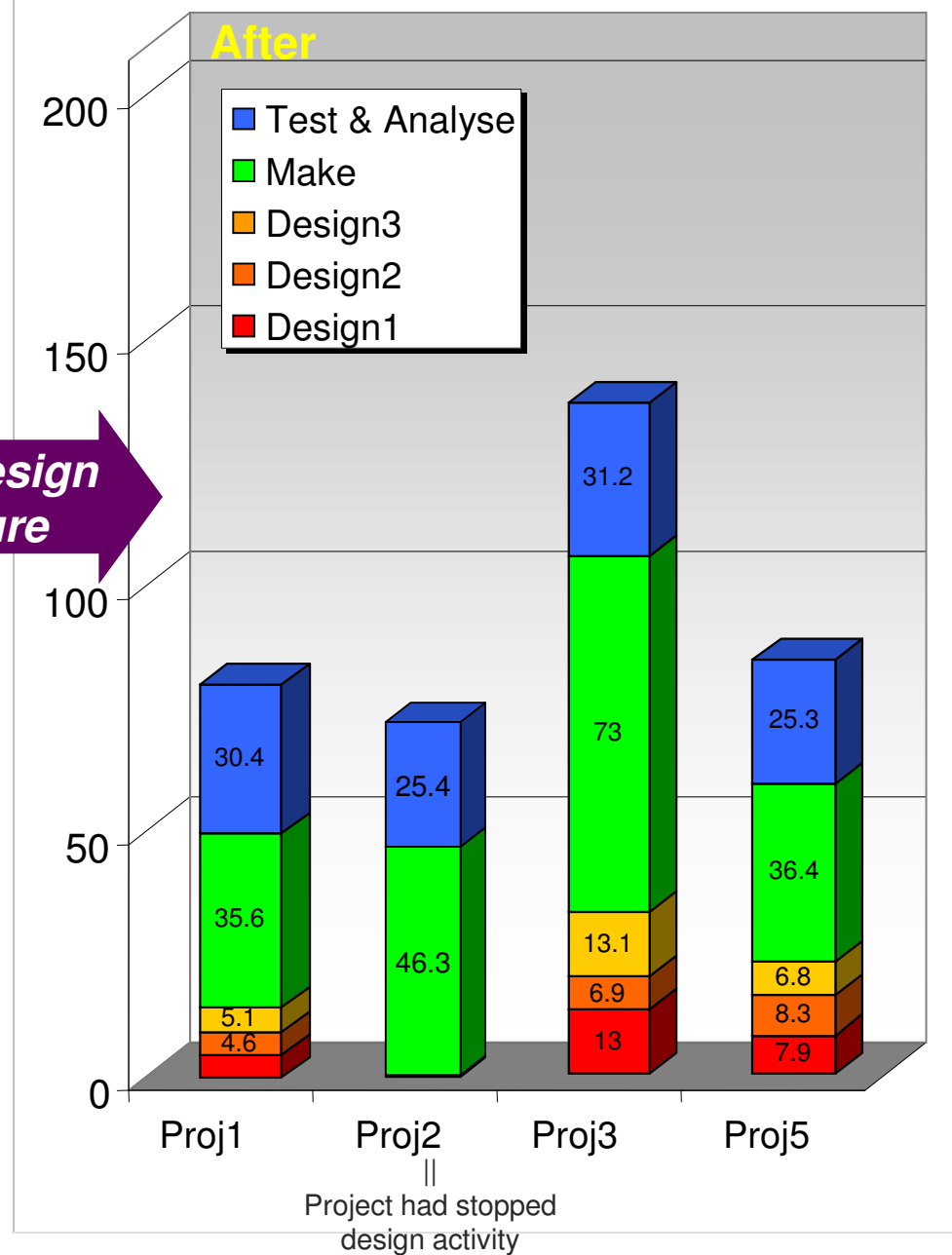
- Generate all possible fragment combinations
 - Predict properties from Free-Wilson models
- Automatically generate the top, *unmade* compounds
 - Useful compliment to the creative design process

Real Value

- Mean Time spent on each DesignSet (set of compounds) for 5 projects in LO



New Design Culture





- Greater need superior design quality
 - Form a design team of experts
- Wikis and Facebook inspire the idea of 'Social Design' tools
 - Integrate with all systems
 - Use very visual display to make understanding easier
 - Frees up time for more creativity
- Powerful computational platform
 - Add value to ideas generated by any designer
 - Make more-informed decisions