CCP CompMedChem - Meeting Minutes (10/03/17)

Executive Summary:

A collective effort that would improve the ease-of-use and productivity of computational chemistry in compound discovery was discussed. In the morning, three core topics were discussed. **1)** Is this a useful ambition and is there will in the room to proceed? **2)** What are the short term (InnovateUK - deadline April 12th) and long term ambitions and what are people able to commit to? **3)** What are the actions and the timelines? In the afternoon we discussed the proposed technical setup in more detail and how CCP CompMedChem would fit into the framework of CCPs and STFC's strategic goals.

From the meeting it was agreed that **1)** CCP CompMedChem would be beneficial for all parties. **2)** Industrial parties were willing to contribute in-kind to grant applications. Academic parties were willing to contribute time and resources in adopting the platform and generating tools for it. **3)** The InnovateUK work packages should be made more focussed and core concerns around the application would be answered over the next week. Longer term - other research councils should be approached for funding for both core development activities and outreach/training.

At Diamond (18): Frank von Delft - DLS/SGC; Anthony Bradley - Oxford/DLS; Garrett Morris - Oxford; Tim Dudgeon - Informatics Matters; Will Pitt - UCB; Mike Bodkin - Evotec; Ben Allen - e-Therapeutics; Mark Forster - STFC; Nathan Brown - ICR; Martyn Winn - CCP4; Sameer Velankar - PDBe; Jason Cole - CCDC; Andrew Leach - ChEMBL; Oliver Smart - PDBe; Jonathan Hirst - U of Nottingham; Ian Wall - GSK; Paolo Tosco - Cresset; Illian Totrov - CCP5
Videoconference (4): Matthias Rarey - U of Hamburg; Marcus Gastreich - BioSolveIT; Val Gillet - U of Sheffield; Peter Schmidtke - Discngine

InnovateUK grant actions

Agreed Actions:

1. Engage InnovateUK regarding concerns (see below) (by wednesday)
2. Rework work packages and re-present to partners (this week)
3. Complete InnovateUK filing (next two-three weeks)
    a. Business case
    b. Scientific scope (narrower and more precise)

Work packages:

Fragments a core focus - in four different areas with proposed partners

1. Compound design tool - Evotec, e-Therapeutics, ICR, CCDC and Cresset
2. Molecular standardisation / preparation - UCB, ChEMBL, STFC (Mark Forster), CCDC
3. Evaluation / comparison of tools - GSK, CCDC and Cresset
    a. Focus on blind prospective studies

4. Consolidation of studentships - GSK, UCB, Evotec, e-Therapeutics, CCDC

Academic funding - to be discussed - depends on industrial input and work packages

Preparedness to commit

- **Industrial:** Willing to commit in kind (timescale permitting)
- **Academic:** Willing to use framework and commit tools / resource to it

Concerns:

We will aim to have answers to the InnovateUK questions by end of Wednesday.

1. What is fundable?
    a. Compounds
    b. Studentships
    c. Data
    d. Software - code or licenses
2. Signed agreement necessary?
3. Is novelty important?
4. Does data used need to be available?
5. Timescale of the auditing
6. Is European money leverageable? What is UK company?

Key issues for companies to consider?

- How will this grant improve your turnover?
- What in-kind material they can get hands on - and in what timescales?
- Are they able to apply auditing to such things simply?

Companies to include:

1. Astex and Vernalis (fragments)
2. Syngenta (not just medicines)

Longer term

Actions:

1. Approach BBSRC / MRC / WT
2. Funding for meeting of CCP4 style annual meeting, workshops and outreach
3. Europe wide funding possibilities
4. Long-term sustainability proposals for maintaining core code

Concerns:

1. Formal proposal for longevity of CCP CompMedChem
2. Specific technologies (e.g. Docker) - build prop
3. Open-source platform needed (not reliant on one person)

All willing to commit irrelevant of success of InnovateUK grant

Clarifications:

- What is the name?
    - CCP CompMedChem - may not be appropriate for agrochemicals - we will continue investigating other names.
- Is CCP CMC developing our own algorithms?
    - This wouldn't be a core goal of CCP CompMedChem. Rather we would support algorithm developers to move their software into use.
- Why is this different to Knime/Pipeline pilot?
    - CCP CompMedChem is overall technology agnostic. Knime workflows are already able to be submitted as SQUONK nodes. We will focus on any technology that enables tools to be used and interoperated.
- Do we envisage Python / Java / Bash etc at core?
    - No. We aim to be as technology agnostic as possible. Containerisation technologies allow for any software language / library to be used.
- What's going to happen to platform code?
    - We will open-source the workflow tool as part of the InnovateUK funding. This would become a community asset.
- What do you get if you put in money?
    - You get to design the projects - ensure they reach an unmet need for your company.
    - You get first access to any software developed - as you have full access to development code.
- How do we support more interactive tools (e.g. Ipython or SeeSAR)?
    - We aim for these to be part of the project further on.
- Can studentships be part of the contributions?
    - Yes - as part of InnovateUK at least.
- Isn't Docker a major issue?
    - We recognise some concerns with Docker - although most of them are historical. We see it is the simplest way of performing the goal we wish to pursue. If other containerisation technologies are preferable - we are willing to pursue them instead.
- What is are the KPIs of InnovateUK funding?
    - Creation of jobs is a must. Needs to have a clear impact on the bottom line of UK industry.

Outcomes of roundtable:

Will Pitt - UCB:

- Structure preparation
- Fragment-to-hit infrastructure
- *Concerns:* Longevity, technology

Mike Bodkin - Evotec:

- Stretch software budget
- Bring in cutting edge tools
- Curation and collection of available tools
- Interoperability
- Top layer for access
- Translate more targets
- *Concerns:* Cash contrib not possible

Ben Allen - e-Therapeutics:

- Bring together tools
- Experimental work can be part of it
- *Concerns:* Timescale might be challenging

Tim Dudgeon - Informatics Matters:

- Focus on FBDD - need focus for application

Mark Forster - STFC:

- Problem focussed rather than broad area
- Define key business problems / competencies

Nathan Brown - ICR:

- Interested in inputting to use cases
    - hit triage / capturing molecular design ideas
- Use within Med-chem groups
- Teaching tool for comp-chem workshops
- Community use

Martyn Winn - CCP4:

- Previous experience - auditing heavy
- large delays on getting datasets approved

Sameer Velankar - PDBe:

- Molecular standardisation, hard to link databases
- Visualisation contribution
- Ligand validation - especially fragments
- Enriching deposition and establishing the standard for quality
- Mechanism for producing and maintaining datasets

Jason Cole - CCDC:

- Potential to find commitment
- Structure preparation / curation - they have test bed
- Longevity / transferring of PhD projects to use
- Rolling out tools even for CCDC an issue

Andrew Leach - ChEMBL:

- Fragments a good business case
- ChEMBL can support / sustain some resources
- Good opportunity to evaluate methods - big possible opportunity

Oliver Smart - PDBe:

- Selling point - make data live in exploitable format
- Deposition workflow

Jonathan Hirst - U of Nottingham:

- Platform valuable
- Not losing PhD projects
- Off the shelf pipeline
- CCP branding - academic input for longer term vision
- Would want to contribute

Ian Wall - GSK:

- Like the project
- 12th April - tight deadline - level of commitment?
- Ability to assess tools - lots of time currently spent doing this
- Maintenance of open-source tools

Paolo Tosco - Cresset:

- Wider, easier and more productive tools
- Difficult to use HPC
- Too many tools and number of other workflows - where's the novelty?
- Need to define competitive advantage and spell out very clearly

Illian Totrov - CCP5:

- Create a self-sustaining community

Garrett Morris - Oxford:

- Method development
- Teaching