

H2020 EINFRA-5-2015



www.bioexcel.eu

Project Number 675728

D3.2 – Consultancy Proposals

WP3: Consultancy & User Groups



Copyright© 2015-2018 The partners of the BioExcel Consortium



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Document Information

Deliverable Number	D3.2
Deliverable Name	Consultancy Proposals
Due Date	2017-04-30 (PM18)
Deliverable Lead	UEDIN
Authors	Adam Carter (EPCC, UEDIN)
Keywords	Consultancy Proposals
WP	WP3
Nature	Report
Dissemination Level	Public
Final Version Date	2017-04-28
Reviewed by	PMB
MGT Board Approval	2017-04-29

Document History

Partner	Date	Comments	Version
UEDIN	2017-03-29	First draft with content from Google Doc	0.1
UEDIN	2017-04-07	Update	0.2
UEDIN	2017-04-25	Further update with more content from multiple partners	0.3
UEDIN	2017-04-25		0.4
UEDIN	2017-04-26	Updated with further contributions and slightly more emphasis on workflows aspect.	0.5
UEDIN	2017-04-27	Updated with corrections from internal review	0.6
UEDIN	2017-04-28	Final version	0.7

Executive Summary

The partners in BioExcel are engaged in numerous projects, many of which have strong connections to the centre's activities. The three core codes have been in development for many years and have been supported by many different projects over this time. It is the case however, that many of these funding streams are focused on specific scientific applications as opposed to other equally important aspects of research software development. Funding from proposals is necessary to support activities related to BioExcel, but individual projects alone cannot fully address the remit of the CoE, namely preparation for Exascale, improvements to software development processes, extensive support to the community and engaging with industry. BioExcel will support world-leading European software whose development is funded through multiple sources. A CoE is needed to integrate these efforts into a coherent approach to maintain European leadership. If there was no CoE, the community would fragment and there would be a real risk that individual efforts would benefit only a subset of the user community.

BioExcel is in a good position to support work related to the codes that other funding sources do not support directly, such as helping to prepare for Exascale, community building, and to help to provide continuity for code developers as they move between different funding streams.

This document offers examples of project proposals that BioExcel has been involved in, and projects or initiatives that are already underway. These are classified as either focused on the project's core codes or related to BioExcel's activities on workflows and data integration. Examples include proposals for specific code enhancements, such as the GROMEX bid, and involvement in larger proposals such as the EGI/EUDAT/INDIGO: European Open Science Cloud proposal, in which the HADDOCK web portal has been included as a thematic service. In terms of workflows, there are several large projects that involve multiple BioExcel partners including the ELIXIR Interoperability Platform, Multiscale Complex Genomics VRE and WestLife VRE. BioExcel partners are also involved in key development work with Open PHACTS and KNIME.

Industry-related activities to date have focused around the establishment of the Interest Group, "Practical Applications for Industry", which now has 30 signed up members from 23 companies.

Work to explore future business models and sustainability plans is well underway in WP5. Whilst it is not yet known exactly what the balance of future funding for the centre will look like, it is clear that whatever the future funding model for the CoE, it is unlikely that any single line of public funding would fully sustain the work of the centre. It is therefore expected that BioExcel will continue to explore options to apply for funding from multiple sources, and this work has begun under the auspices of Task 3.6. Core BioExcel funding will provide continuity and stability, and serve to multiply the impact of the individual projects which could be, at least partially, co-ordinated by the CoE.

Contents

1	INTRODUCTION	6
2	PROJECTS AND INITIATIVES RELATED TO BIOEXCEL	6
2.1	ACTIVITIES RELATED TO THE PROJECT'S CORE CODES	7
2.1.1	EGI/EUDAT/INDIGO: EUROPEAN OPEN SCIENCE CLOUD	7
2.1.2	USE CASE WORK WITH THE HUMAN BRAIN PROJECT	7
2.1.3	SWEDISH EXASCALE COMPUTING INITIATIVE	7
2.1.4	ADENYLYL CYCLASE-CATALYZED REACTION MECHANISM OF ATP CYCLIZATION	8
2.1.5	USE CASE WORK ON MOLECULAR RECOGNITION	8
2.1.6	ALCHEMICAL FREE ENERGY CALCULATIONS FOR LIGAND DESIGN	8
2.1.7	GROMEX	8
2.1.8	HARDENING AND DEVELOPMENT OF THE GROMACS MOLECULAR SIMULATION ENGINE	8
2.1.9	TORPRED, "IN SILICO GENOME-GUIDED PREDICTION OF TOXICITY AND RESISTANCE"	8
2.1.10	POSTDOCTORAL SCHOLARSHIP FROM CARL TRYGGER FOUNDATION	9
2.1.11	DEEPRANK	9
2.1.12	DYNION RESEARCH NETWORK	9
2.2	RELATED TO WORKFLOWS & DATA INTEGRATION	9
2.2.1	COMMON WORKFLOW LANGUAGE	9
2.2.2	ELIXIR INTEROPERABILITY PLATFORM	10
2.2.3	IMI OPEN PHACTS AND OPEN PHACTS FOUNDATION	10
2.2.4	KNIME	10
2.2.5	MULTISCALE COMPLEX GENOMICS VRE (MUG)	11
2.2.6	COMPSS	11
2.2.7	US FDA BIOCOMPUTE OBJECTS	12
2.2.8	EBI CLOUD PORTAL	12
2.3	PARTNERSHIPS	12
3	INDUSTRY: RELEVANCE AND VALUE	12
4	CONCLUSIONS AND FUTURE PLANS	13

1 Introduction

In this document, we describe our work to date in developing consultancy proposals relevant to the Centre and our plans for expanding on this over the coming months.

The partners in BioExcel are engaged in many projects, many of which have strong connections to BioExcel. The project's core codes have been in development for many years and have been supported by many different projects over this time. It is the case however, that many of these funding streams are focused on specific scientific applications as opposed to other equally important aspects of research software development. Funding from proposals is necessary to support activities related to BioExcel, but individual projects alone cannot fully address the remit of the CoE, namely preparation for Exascale, improvements to software development processes, extensive support to the community and engaging with industry. BioExcel will support world-leading European software whose development is funded through multiple sources. A CoE is needed to integrate these efforts into a coherent approach to maintain European leadership. If there was no CoE, the community would fragment and there would be a real risk that individual efforts would benefit only a subset of the user community.

This document is intended to illustrate ways in which the CoE could fund future activities by offering examples of project proposals that BioExcel has been involved in, and projects and initiatives that are already underway either as part of the work of BioExcel or with the project's partners.

The main part of this document, Section 2, describes these projects and initiatives. In Section 3 we look in particular at industry, and provide a summary of some of the input that has been obtained through the Industry Interest Group to date. A brief description of future plans, and conclusions are offered in Section 4.

BioExcel will continue to explore options to apply for funding from multiple sources as part of Task 3.6, which runs until the end of the project; an update to this document (D3.6) will be published in project month 36.

2 Projects and Initiatives Related to BioExcel

In this section, we describe some projects and proposals that have been made by BioExcel's partners in support of activities in the realm of BioExcel. We have assigned these to broad categories, depending on whether they relate to work on the project's core pilot codes, or whether they are related to workflows and data integration.

2.1 Activities Related to the Project's Core Codes

2.1.1 EGI/EUDAT/INDIGO: European Open Science Cloud

The HADDOCK web portal¹ has been included as a WeNMR² thematic service in the joint EGI³/EUDAT⁴/INDIGO-Datacloud⁵ e-Infra12⁶ proposal “European Open Science Cloud” submitted late March to the EU. The HADDOCK web portal with its 8500+ registered users is in full production, sending >8 million jobs to EGI HTC resources per year. As such it has been recognized as a valuable thematic service in the EOSC proposal. Next to ensuring the operation of the service, our participation in this project should enable the following integration activities:

- Modernization of the HADDOCK portal registration process and integration with the EGI/EUDAT/INDIGO AAI solutions. Currently the service relies on both a local database and a SSO service from the WeNMR VRC.
- Consolidation of the submission machineries building on DIRAC4EGI⁷ and INDIGO solutions
- Integration of distributed data storage solutions for users (e.g. based on EUDAT and INDIGO services).

This should also ensure continuous access to the EGI HTC resources. This bid aims at 36 person months (PMs) of effort to support the operation of the group's web-enabled portals (for HADDOCK, DisVis and PowerFit).

This bid builds on previous successful applications by the HADDOCK team at BioExcel partner UU, including 24 PMs through INDIGO, 8PMs through EGI-engage and 48PMs through the WestLife VRE which funded supported the development of the DisVis and PowerFit portals.

2.1.2 Use Case work with the Human Brain Project

Support for the work of Pilot Use Case on Multiscale Modelling, has been given by the Human Brain Project (HBP) of which BioExcel partner FZJ is recipient (6 PM so far). The Human Brain Project is scheduled to continue for the next seven years. This will allow evaluating and testing the new QM/MM interface in CPMD. The input of HBP could be a useful additional funding stream to support BioExcel's work in this area in the longer term.

2.1.3 Swedish Exascale Computing Initiative

KTH successfully bid for a 5-year grant (15MSEK) to the Foundation for Strategic Research's Research Infrastructure Fellows programme to fund the Swedish Exascale Computing Initiative. A significant part of this bid will support development of GROMACS, looking particularly at making better use of GPUs/accelerators.

¹ <http://milou.science.uu.nl/services/HADDOCK2.2/haddock.php>

² <https://www.wenmr.eu>

³ <https://www.egi.eu>

⁴ <https://www.eudat.eu>

⁵ <https://www.indigo-datacloud.eu>

⁶ <http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/einfra-12-2017.html>

⁷ <https://wiki.egi.eu/wiki/2016-bidding/DIRAC4EGI>

2.1.4 Adenylyl Cyclase-catalyzed reaction mechanism of ATP cyclization

In February, FZJ applied to the JARA-HPC call to obtain time on the CLAIX supercomputing cluster at Rechenzentrum Aachen (Germany). The proposal is still pending at the time of writing. This compute time will also help to support work related to the multiscale modelling Use Case.

2.1.5 Use Case work on Molecular Recognition

The molecular recognition Pilot Use Case is being supported by a €500k (Dutch) eScience grant, which supports 3 years' worth of PostDoc effort and 3 years of an eScience engineer for our collaborators at the VU in Amsterdam who are developing the workflow system into which HADDOCK will be integrated.

On the science side, UU have also secured support from NWO⁸ in a joint TOP-PUNT project with Marc Baldus (also UU, but not directly funded by BioExcel) to work on docking to membrane proteins.

2.1.6 Alchemical free energy calculations for ligand design

This work on extending the PMX framework to support alchemical ligand transformation for inhibitor design was supported until April 2017 by a research grant from Boehringer Ingelheim. As a follow-up, MPG applied successfully for an Alexander von Humboldt fellowship, which was awarded to Matteo Aldeghi to work on protein design via free energy mutation simulations using PMX.

2.1.7 GROMEX

GROMEX⁹, a SPEXXA project, is an example of a project that is currently underway involving two of BioExcel's partners (MPG and KTH). Whilst this project had already begun at the time that BioExcel started, it is an example of one way in which the activities of BioExcel are being complemented by national funding. It is unlikely that the CoE will ever fully support the main development of codes such as GROMACS, but with reliable core funding from BioExcel it will be possible to apply for future similar programs to help support work such as this in BioExcel's areas of interest.

2.1.8 Hardening and Development of the GROMACS Molecular Simulation Engine

As part of the growing international GROMACS community, KTH is part of an NIH grant¹⁰ where they collaborate with US principal investigators to fund work in their teams focused on enhancing the usability of GROMACS. KTH's funded work here is limited to integrating the contributions made by the US PIs, but it is a good example of how better project coordination leads to new funding streams.

2.1.9 ToRPred, "In Silico Genome-guided prediction of toxicity and resistance"

KTH and IRB have together submitted a proposal under the topic sc1-pm-16-2017 (In-silico trials for developing and assessing biomedical products). This project does not directly fund software development, but the project will rely heavily on high-throughput free energy calculations, and will contribute documentation and use cases back to BioExcel.

⁸ <https://www.nwo.nl/en>

⁹ <http://www.sppexa.de/general-information/projects-phase-1.html#GROMEX>

¹⁰ GB10293-151962

2.1.10 Cryo-EM restraint work supported by the Carl Trygger Foundation

To help integrate molecular dynamics simulations with refinement of cryo-EM data and the RELION code, KTH successfully applied for and was granted a two-year postdoctoral scholarship which will support work on combining Molecular dynamics with low-resolution restraints from cryo-EM in GROMACS.

2.1.11 DeepRank

The research group of Alexandre Bonvin at UU, who are partners in BioExcel, has received funding from the Dutch eScience Center for the DeepRank project, which is focused on scoring 3D protein-protein interaction models using deep learning.

Deep neural networks will be trained on data from the Protein Data Bank and the resulting scoring function, DeepRank, will markedly enhance our capability to reliably model biomolecular complexes, assisting the scientific community to gain insights into macromolecular aspects of life. It will be implemented in the HADDOCK modelling platform and freely distributed through GitHub and eStep repositories, ensuring a wide dissemination.

2.1.12 Dynion research network

MPG and FZJ were part of a successful bid to the DFG for the Dynion Research Network. This grant supports 8 groups in Germany working on ion channels and this work is related to the PMX work underway in BioExcel.

2.2 Related to Workflows & Data Integration

2.2.1 Common Workflow Language

One partner in BioExcel (UNIMAN) is active in the community initiative to establish a Common Workflow Language (CWL)¹¹. CWL is a specification for describing analysis workflows and tools in a way that makes them portable and scalable across a variety of software and hardware environments, from workstations to cluster, cloud, and high performance computing (HPC) environments. CWL is designed to meet the needs of data-intensive science, such as Bioinformatics, Medical Imaging, Astronomy, Physics, and Chemistry.

CWL is developed by an informal, multi-vendor working group consisting of organizations and individuals aiming to enable scientists to share data analysis workflows. The CWL project is on Github and follows the Open-Stand.org principles for collaborative open standards development. Founding partners include UNIMAN (through Apache Taverna), Galaxy, and Seven Bridges. KNIME users have requested CWL support.

BioExcel, through UNIMAN, has contributed to the CWL specification, developed the CWL Viewer: the first fully featured CWL visualisation tool and web gallery¹², and is providing the metadata manifest Research Object (RO)¹³ mechanism for describing workflow bundles.

¹¹ <http://www.commonwl.org/>

¹² <https://view.commonwl.org/>

¹³ <http://www.researchobject.org>

CWL+RO is included in the e-Infra12¹⁴ proposal “AI30” submitted late March to the EU.

2.2.2 ELIXIR Interoperability Platform

Three partners in BioExcel (UNIMAN, IRB and BSC) are involved in supporting the ELIXIR Interoperability Platform (EIP). The EIP among other projects, to set a series of recommendation and best practices related to the specification of software APIs, tools and workflows. EIP will promote the Common Workflow Language for the manager-agnostic description of tools and workflows. BioExcel has been chosen as use case for these activities. The common roadmap includes the specification of workflow demonstrators prepared in BioExcel WP2 following ELIXIR recommendations, and the complete and curated registration of BioExcel tools in ELIXIR’s registry bio.tools.

2.2.3 IMI Open PHACTS and Open PHACTS Foundation

One partner in BioExcel (UNIMAN) was involved with the Innovative Medicines Initiative (IMI) funded phase of Open PHACTS and the current Open PHACTS Foundation. The Open PHACTS Discovery Platform has been developed to reduce barriers to drug discovery in industry, academia and for small businesses by integrating and linking together public and pharma data to relate compounds, targets, pathways, diseases and tissues. Data sources include ChEBI, ChEMBL, ChemSpider, ConceptWiki, DisGeNET, DrugBank, Gene Ontology, neXtProt, UniProt and WikiPathways. Access is via the Open PHACTS API or applications including the Open PHACTS Explorer.

BioExcel partner UNIMAN led work on the Platform and its architecture, data loading pipelines into the platform from public and EFPIA datasets, identifier mapping and linkset services, and governance and the Explorer application. We contribute to API design and release management, and developed the platform release management process and the Docker 3rd-party hosted installations (for EFPIA members).

For workflows, UNIMAN developed the first KNIME Node for accessing Open PHACTS. These proved attractive to pharmaceutical companies and UNIMAN is working with the CWL community on KNIME-CWL support for Open PHACTS and outside this project.

BioExcel will benefit from incorporating Open PHACTS KNIME Nodes in BioExcel workflows and provisioning BioExcel workflows as KNIME Nodes into the Open PHACTS community.

2.2.4 KNIME

KNIME is a leading Analytics workflow manager for data-driven innovation. It is an enterprise-grade, open source platform with more than 1000 modules, hundreds of ready-to-run examples, a comprehensive range of integrated tools, and a wide choice of advanced algorithms available. It is a dual product: open

¹⁴ <http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/einfra-12-2017.html>

source so accessing a global community of data scientists, and licensed for rich support, additional features and more specialized KNIME Nodes. KNIME Nodes are wrapped steps in KNIME workflows.

BioExcel is working with KNIME through Open PHACTS (see above), through developing BioExcel KNIME Nodes, and through KNIME-compliant CWL workflows. KNIME and UNIMAN are partners in an INFRAIA-02-2017 proposal “IBISBA” submitted in March 2017. This proposal uses CWL-Compliant KNIME and a CWL workflow repository.

2.2.5 Multiscale Complex Genomics VRE (MuG)

Three partners in BioExcel (EBI, IRB and BSC) are involved in the MuG Virtual Research Environment (VRE) for the multi-scale analysis of chromatin structure. MuG VRE is a H2020 project (Total ca. 3M€). MuG is building its infrastructure on a similar computational infrastructure to that used in BioExcel, and uses the PMES-COMPSs programming model to power its tool infrastructure. BioExcel will benefit of MuG interaction by obtaining use cases for the deployment of genomics, and tools in the field of Nucleic Acids simulation.

2.2.6 West-Life VRE

The Utrecht partner in BioExcel is involved in the West-Life Virtual Research Environment (VRE) project (www.west-life.eu), a H2020 project providing computational and data management solutions for structural biology. West-Life, which build upon the achievement of the WeNMR VRC, provides services for computation and data management to researchers in structural biology, integrating multiple approaches and experimental techniques. It builds on European e-Infrastructure solutions from EGI and EUDAT and links together web services and repositories for structural biology. West-Life and WeNMR are supported by EGI via a [SLA agreement](#) in which eight resource centres have pledged 60 million hours of opportunistic grid computing time, 60 dedicated core on the EGI Federated Cloud and more than 250 TB storage capacity to support the enmr.eu VO. This benefits BioExcel both by connecting HADDOCK to addition services outside the BioExcel pilot applications and by providing the computational resources to the HADDOCK portal which sends >8 million jobs per year to EGI HTC sites.

2.2.7 COMPSs

COMP Superscalar¹⁵ (COMPSs) belongs to the BSC research line on Workflows and Distributed Computing and is being used in WP2 of BioExcel. It is used in production in BSC infrastructures including the MareNostrum supercomputer and is also deployed, in conjunction with PMES, to manage the execution of applications in the IRB Cloud-powering MuG-VRE. BSC participates in other EC funded projects where the COMPSs framework is both adopted to implement scientific applications and extended to be interoperable with new computing platform. As part of the H2020 EUBra-BIGSEA initiative¹⁶, COMPSs has been extended to support the Mesos middleware and the execution of containerized applications. BioExcel users can benefit of these developments by extending the

¹⁵ <https://www.bsc.es/research-and-development/software-and-apps/software-list/comp-superscalar>

¹⁶ <http://www.eubra-bigsea.eu>

possible set of platforms where COMPSs workflows can run. The research on COMPSs is also partly funded through BSC ordinary funding which comes through the Spanish national research framework.

2.2.8 US FDA BioCompute Objects

One partner in BioExcel (UNIMAN) is working with the FDA BioCompute Object initiative. A BioCompute Object (BCO) is a record that includes all software arguments of the executable program, version information, and a reference to all the inputs including the usability domain. BCO is sponsored by the US FDA (Food and Drug Administration) HTS computation and regulation to harmonize HTS analysis, evaluate and validate pipelines and construct novel pipelines through combining multiple BCOs.

BioExcel, through UNIMAN, has participated since Autumn 2016 in this work, working on the BCO specification and bringing on-board CWL and Research Objects. BioExcel gave invited talks (including invited keynote) at the HTS Computational Standards for Regulatory Sciences Workshop at NIH Bethesda March 2017¹⁷. This work includes the Galaxy team.

For BioExcel, the benefit of the cooperation is impact and visibility as well as the technical benefits of working with experts on reproducible and portable regulator-validated HTS pipelines.

2.2.9 EBI Cloud Portal

The EBI's cloud portal that is being used in BioExcel is partially supported by commercial cloud providers with whom EBI has been working and also through the PhonoMeNal project¹⁸.

2.3 Partnerships

We have been exploring several partnerships with different organisations and we have recently entered into partnerships with ELIXIR, MolSSI and Open PHACTS. The partnerships are discussed in D4.4 and D3.3¹⁹, so we do not discuss these further here.

3 Industry: Relevance and Value

Industry includes large companies which generate pharmaceutical treatments of human and veterinary disease (e.g. Bayer), agrichemical products to protect agricultural resources (e.g. Syngenta) and processing of food products by multinational food combines (e.g. Unilever). Most companies in the life science sector exploit modern techniques and methodologies in structural biology which includes biomolecular dynamic modelling and simulation. The practical

¹⁷ https://hive.biochemistry.gwu.edu/htscsrs/workshop_2017

¹⁸ <http://phenomenal-h2020.eu/home/>

¹⁹ BioExcel Deliverables D4.4, *Dissemination report and updated plan*, and D3.3 *Consultancy Modalities and Funding Options*, to be published alongside this document; available from zenodo.org.

application of these techniques for product discovery through making predictions within business processes such as virtual screening and antibody engineering.

BioExcel is positioned to engage with industry to facilitate such successful applications through professional consultancy services. We are starting to do this by formation of an Industry Interest Group which meets monthly to focus on practical applications of computational biomolecular research. It shares published use cases of relevance to industry to understand the impact and benefit of biomolecular modelling and simulations. We expect those aspects which generate most interest will emerge to guide the provision of practical support and advice. This is important for expanding usage of the core applications and developing our sustainability plans. This interest group currently comprises 30 individuals signed up representing 23 companies in industry including AstraZeneca, Janssen (J&J), Syngenta and UCB.

A major challenge for BioExcel is to demonstrate how it adds unique business value over the other competing applications for biomolecular dynamic modelling and simulation. We discuss this further in the accompanying D3.3 Consultancy Modalities document, which categorises those types of consultancy services that would be of interest to paying customers, particularly those from industry.

One key issue is the entry point to BioExcel codes and workflows by industry. From this point of view, it is important that BioExcel pays attention to developing and deploying containers for business that can be safely executed on their platforms of choice and are available through their tools of choice, for example as KNIME Nodes in workflows.

One partner, FZJ, is already working closely with an industry partner in the scope of BioExcel: Fondazione Illy is supporting a 5-years professorship on the topics related to odorant receptors with a 50% contribution. The other 50% is provided by FZJ. Illy supports work on odorant receptors and their cascades. This includes the Adenylate Cyclase enzyme that is currently being studied in BioExcel's Multiscale Modelling use case. This rare opportunity to have a co-funded professorship is proving to be a very effective mechanism for collaboration.

4 Conclusions and Future Plans

It can be seen that BioExcel's partners are involved in a broad range of projects, funded from a variety of sources. It can also be seen that these individual projects, whilst diverse, do not cover the full gamut of what we hope the CoE can become; in particular, they are predominantly focused on single codes, or single applications or platforms (albeit with some examples of integration activities). There are several kinds of added value that a CoE could offer over what is supported in these projects, including better continuity, interdisciplinary discussion and collaboration, more coordinated training programmes and,

importantly, a longer-term vision and approach to problems such as readiness for exascale.

Work to explore future business models and sustainability plans is well underway in WP5. Whilst it is not yet known exactly what the balance of future funding for the centre will look like, it is clear that whatever the future funding model for the CoE, it is unlikely that any single line of public funding would fully sustain the work of the centre. It is therefore expected that BioExcel will continue to explore options to apply for funding from multiple sources, such as those described in this document and this work will continue under the auspices of Task 3.6.

The work in WP3 over the first 18 months has been focused on reaching out the multiple communities with which BioExcel plans to work, as described in D3.1. We have established 7 Interest Groups, delivered 13 webinars, and started to hold our own face-to-face meetings. This kind of work has taken time, but as we move into the second half of the project, increased effort will be put into consultancy proposals to ensure that we take full advantage of the various funding streams available. The range of projects here shows that the Centre's partners are already important players in the area in which the CoE will operate but further work is required to build on this, and to capitalise on the growing momentum of CoE-focused activities.

Partners working in WP3 have also been active in discussions in WP5 and some of the work has fed into the list of funding schemes described in D5.2. This list of funding schemes will be used as the basis for a systematic review of funding opportunities. We will also be exploring and monitoring the extreme scale demonstrators to seek opportunities to engage with any that could eventually support extreme-scale capabilities of our core codes.

Over the next year, we will work to ensure that through WP3 we will make the most of the links we have with the projects described here and make these connections evident on the BioExcel web site. In the longer term, core BioExcel funding could provide continuity and stability and serve to multiply the impact of the individual projects which could be, at least partially, co-ordinated by the CoE.