Innovative Medicines Initiative

IMI-GB-Dec-2011-09.1

2011 4th Call for proposals

Innovative Medicines Initiative

Final version

Table of Contents

GENERAL PRINCIPLES

EU MEDICAL INFORMATION SYSTEM

1. A EUROPEAN MEDICAL INFORMATION FRAMEWORK (EMIF) OF PATIENT-LEVEL DATA TO SUPPORT A WIDE RANGE OF MEDICAL RESEARCH 6

2. ETRIKS: EUROPEAN TRANSLATIONAL INFORMATION & KNOWLEDGE MANAGEMENT SERVICES 20

CHEMISTRY, MANUFACTURING AND CONTROL

3. DELIVERY AND TARGETING MECHANISMS FOR BIOLOGICAL MACROMOLECULES 30

4. IN VIVO PREDICTIVE BIOPHARMACEUTICS TOOLS FOR ORAL DRUG DELIVERY 36

5. SUSTAINABLE CHEMISTRY – DELIVERING MEDICINES FOR THE 21ST CENTURY 44

TECHNOLOGY AND MOLECULAR DISEASE UNDERSTANDING

6. HUMAN INDUCED PLURIPOTENT STEM (HIPS) CELLS FOR DRUG DISCOVERY AND SAFETY ASSESSMENT 54

7. UNDERSTANDING AND OPTIMISING BINDING KINETICS IN DRUG DISCOVERY 65

GENERAL PRINCIPLES

The Innovative Medicines Initiative Joint Undertaking (IMI JU) is a unique pan-European public private partnership aiming to foster collaboration between all relevant stakeholders including large and small biopharmaceutical and healthcare companies, regulators, academia and patients.

The aim of IMI is to support pre-competitive[1] pharmaceutical research and development (R&D) to foster the development of safe and more effective medicines for patients through removing identified bottlenecks in the drug development process, and enhance Europe's competitiveness by ensuring that its biopharmaceutical sector remains a dynamic high-technology sector.

The IMI Research Agenda http://www.imi.europa.eu/content/research-agenda describes the research bottlenecks in the drug development process and identifies four strategic pillars: Predictivity of Safety Evaluation, Predictivity of Efficacy Evaluation, Knowledge Management and Education and Training.

The IMI 2011 Call for proposals will have seven topics covering the three following clusters within Knowledge Management and the Predictivity of Safety Evaluation:

- EU medical information system
- Chemistry, manufacturing and control
- Technology and molecular disease understanding

Before submitting an Expression of Interest, the various Call documents, such as *IMI JU Rules for submission, evaluation and selection of Expressions of Interest*, *Rules for Participation*, *the IMI Intellectual Property Policy*, etc., shall be considered carefully. These documents are published on the IMI website www.imi.europa.eu at the time of the 2011 Call launch.

The size of each consortium should be adapted to the scientific goals and the expected key deliverables.

Synergies and complementarities with other ongoing FP7 activities should be explored in order to avoid overlaps and duplications and to maximise European added value in health research.

DURATION OF THE PROJECTS

The indicative duration of projects is between 4 years (topic 5) and 5 years for the other topics.

[1] In the present context, 'pre-competitive pharmaceutical research and development' should be understood as research on the tools and methodologies used in the drug development process.

FUNDING OF THE PROJECTS

The total available financial contribution from the IMI JU to participants eligible for funding will be up to the amounts that the research-based companies that are members of the European Federation of Pharmaceutical Industries and Associations (EFPIA – www.efpia.org) will contribute as 'in kind'2 contribution (indicated in the respective call topics). The total indicative financial contribution from the IMI JU will be up to EUR 105 million.
The Applicant Consortia shall keep in mind that the budget of each Expression of Interest should be adapted to the scientific goals and the expected key deliverables of the project.

SYNOPSIS OF CALL AND EVALUATION PROCESS

The IMI JU supports research activities following open and competitive Calls for proposals, independent evaluation and the conclusion of Project- and Grant Agreements.

Each topic included in the 2011 Call for proposals is associated with a group of pharmaceutical companies that are members of EFPIA (hereafter called the 'EFPIA Consortia') and which are committed to collaborate with public and private organisations eligible for funding by the IMI JU. The EFPIA members will provide 'in kind' contributions to support their activities within the research projects.

The IMI JU applies a two stage Call process. In the first stage, 'Applicant Consortia' (i.e. formed by academia, small and medium sized enterprises (SMEs), patient organisations, non EFPIA companies, etc.) are invited to submit, to the IMI JU, an Expression of Interest (EoI) in response to a Call topic.

In preparing their EoIs, the Applicant Consortia should carefully consider the research contribution that an EFPIA Consortium will make to a given project as well as the expectations from the Applicant Consortia, as outlined in topic texts below.

Each EoI submitted will be reviewed by independent experts according to predefined evaluation criteria. The Applicant Consortia with the highest ranked EoI will be invited to jointly develop a Full Project Proposal together with the EFPIA Consortium associated to the corresponding topic. The Full Project Proposal will then be subject to a final review by independent experts according to predefined evaluation criteria.

Only Full Project Proposals that have been favourably reviewed in the evaluation process can be selected for funding. These projects will then be invited by the IMI JU to conclude a Grant Agreement governing the relationship between the selected project consortium and the IMI JU.

For full details, applicants should refer to the *IMI JU Rules for submission, evaluation and selection of Expressions of Interest* published on the IMI JU website www.imi.europa.eu at the launch of the 2011 Call.

[2] In kind contribution is e.g. personnel, clinical research, equipment, consumables.

ELIGIBILITY TO PARTICIPATE IN PROJECTS AND TO RECEIVE FUNDING FROM THE IMI JU

Criteria of eligibility to participate in IMI projects and the criteria to receive funding from the IMI JU are specified under the *Rules for participation* published on the IMI JU website www.imi.europa.eu at the launch of the 2011 Call.

The IMI JU financial contribution will be based on the reimbursement of the eligible costs. The following funding rates apply to the legal entities eligible for funding: For research and technological development activities, up to 75% of the eligible costs and for other activities (including management and training activities) up to 100% of the eligible costs charged to the project are eligible for funding. For the indirect costs (overheads)[3],the legal entities eligible for funding may opt for one of the following indirect costs methods: the actual indirect costs; or the simplified method which is a modality of the actual indirect costs for organisations which do not aggregate their indirect costs at a detailed level, but can aggregate them at the level of the legal entity; or a flat rate of 20% of total eligible direct costs (excluding subcontracting costs and the costs of resources made available by third parties which are not used on the premises of the beneficiary).

For full details, Applicant Consortia are invited to refer to the *Rules for Participation* (www.imi.europa.eu).

The research based companies that are members of EFPIA shall not be eligible to receive financial contributions from the IMI JU.

IMI INTELLECTUAL PROPERTY POLICY

The IMI Intellectual Property Policy (IMI IP policy, www.imi.europa.eu) has been developed to be aligned with the objectives of the IMI JU to ensure knowledge creation, together with the swift dissemination and exploitation of knowledge, and fair reward for innovation.

The IMI IP Policy sets out *inter alia* basic principles regarding ownership of Background and Foreground, access rights depending on the entity and the purpose, and dissemination.

In submitting an EoI, the Applicant Consortia fully understand the principals laid out in the IMI IP policy that will apply to all research projects conducted under the IMI JU.

The IP policy does not foresee all details and does not aim to answer to all possible practical situations participants may be faced with. Flexibility is provided for participants to establish the most appropriate agreements (e.g. the project agreement) serving each individual project's objectives, and considering the wider IMI objectives.

Applicant Consortia are invited to read carefully the Guidance Note on the IMI IP Policy (www.imi.europa.eu), whose purpose is to explore ways to handle related issues and

[3] The funding rules regarding indirect costs are subject to the modifications of the IMI model grant agreement to be adopted by decision of the IMI JU Governing Board.

pitfalls that participants may encounter during the preparation, negotiation and completion phases of the Grant Agreement and Project Agreement.

PROJECT AGREEMENT

All participants of a selected IMI project are required to negotiate and sign a Project Agreement between them before the Grant Agreement is signed with the IMI JU.
The Project Agreement is a private agreement which the participants of an IMI project conclude amongst themselves to implement the provisions of the Grant Agreement and to regulate internal issues related to work organisation and objectives for each participant, consortium governance, IP, financial and other matters.

EU MEDICAL INFORMATION SYSTEM

1. A EUROPEAN MEDICAL INFORMATION FRAMEWORK (EMIF) OF PATIENT-LEVEL DATA TO SUPPORT A WIDE RANGE OF MEDICAL RESEARCH

This Call Theme consists of 3 topics:

- Information Framework / Knowledge Management Service Layer
- metabolic complications of obesity
- protective and precipitating markers for the development of Alzheimer's disease (AD) and other dementias.

The EoIs should address one of these three topics. At the second stage, the successful applicant consortium for each topic will merge with the EFPIA consortium to prepare a single Full Project Proposal for the Call Theme.

BACKGROUND

Patient-level health information

The increasing adoption of electronic health records (EHR) and the widespread efforts by European authorities to develop large-scale health information infrastructures are quickly building a wealth of medical information. While this information holds the potential to significantly advance medical and pharmaceutical research, to date, by and large this potential has not been realised due to fragmentation of this information because of the lack of legal and technical standards, cost effective platforms, and sustainable business models for the access and analysis of such data. Thus, even though a considerable amount of relevant patient health information does exist, it typically resides in a variety of systems in a fragmented way in different locations, thereby inhibiting ready and efficient access from a central place. A number of efforts have been made to link a variety of patient health data together, but they are typically small in scale and geographically confined. There is currently no system available that enables researchers to link data on a large scale from patient health record resources with other electronic data sets including, longitudinal and patient research data, survey and administrative data, omics, imaging, social, environmental and economic data.

Furthermore, there is a particular need for patient-level health information for paediatric populations on a larger scale. The European Regulation on Paediatric Medicines was recently set up in order to facilitate the development and boost the availability of medicines specifically adapted and authorised for use in the paediatric population; to improve the availability of information on the use of medicines in children; and to ensure that research into medicines in children is of high quality.

In summary, access to data that are both large-scale and contain detailed health information will open up new opportunities for research by enabling the investigation of research questions that have so far been impossible to address satisfactorily.

The overall vision for the EMIF project is to create a lasting and comprehensive framework to take full advantage of available patient-level data, including a broad network for access to existing patient-level data, a governance model regulating ethics and privacy aspects and a platform for data management and analysis. While starting

with the 2 research topics described in this current call, the framework will extend beyond the current call to calls 5 and 6 (and may extend beyond the IMI programme). Further research call topics will be developed through the identification of drug discovery bottlenecks where access to patient level data is at the core. It is expected that all successful consortium funded under this theme work together as a research cluster, with the EMIF platforms as the horizontal layer connecting activities and ensuring that synergy and efficiencies are maximised. This may include other research methods beyond the extreme phenotype approach.

Specific unmet medical need and research topics

A so-called extreme phenotype approach is a powerful way of investigating disease and phenotypes including assessing risk factors for disease onset, progression and outcomes. The extreme phenotype approach is based on the concept that individuals at the extreme of the distribution of a particular trait have a high probability of having a mono- or oligogenic predisposition to this trait. This has been used in the successful development of innovative diagnostic tests to predict disease, for instance BRCA1/2 sequencing for familial forms of breast cancer. Elucidation of the genetic variants which explain why a certain individual is at the extreme of the distribution is now feasible using breakthrough technologies like exome sequencing and direct comparison of the extremes. This information can then be applied to understand less extreme variations in the phenotype, for diagnostic purposes and to develop innovative therapeutics for the condition under evaluation.

The prevalence of obesity is reaching epidemic proportion in western countries and, at a slower yet steady pace, in developing countries. Because obesity predisposes people to a variety of disabling conditions like osteoarthritis, cancer, depression and diabetes, it represents a major health problem for society. This problem is getting even worse as obesity (and diabetes), are already found in children. Behavioural interventions have shown limited benefit, and so have pharmacological measures to limit weight gain, to the point where the only intervention which has been shown to be effective is surgery. A fraction of obese individuals develop complications. The mechanism underlying this inter-individual variability in the susceptibility to the complications of obesity remains poorly understood. Elucidation of these mechanisms at the molecular level would have major beneficial consequences, as they would make it possible to target interventions to individuals at high risk of complications, and could potentially lead to the development of novel therapies and to the generation of diagnostic tests designed to identify high-risk individuals.

Alzheimer's disease (AD) and other dementias represent a major challenge to western societies. Only symptomatic treatments are available presently. The discovery and the development of innovative therapeutics for AD are hampered by the slow progression of the disease (leading to long development studies), the absence of biomarkers reliably predictive of risk (hence the inclusion of already diseased individuals, at a point of their disease which may have reached irreversibility) and the absence of surrogate markers for efficacy.

NEED FOR PUBLIC-PRIVATE COLLABORATIVE RESEARCH

The development and implementation of a European Medical Information Framework (EMIF) requires a massive effort that spans a number of contributors and experts from pharma industry, healthcare organisations, academia and other third parties throughout Europe. Specifically, input is needed from providers of patient-level health information and experts in the collation, linkage and use of patient-level health information, including

experts in the areas of standards, ethics, semantics, law, etc. For example, in addition to requiring access to data and samples from very large numbers of individuals, access to technologies and resources from both academia and industry is needed as is the provision of innovative diagnostic and therapeutics to move into translation and provide immediate benefits to patients.

The envisaged project output of a) access to a multitude of patient-level health information sources, b) the results of the specific research studies and c) the framework that includes the solutions for data linkage will inform the pharma industry, academia, patient-level health data providers, healthcare professionals, public health organisations, patient organisations, regulatory and legislative bodies and other third parties across Europe.

OVERALL OBJECTIVES

Topic 1

- Develop the Information Framework that:
    a. provides access to the required patient level data in adults and children for the specific research topics below;
    b. is sustainable and scalable to serve as the information base for further research projects that use patient health data following completion of this project;
    c. offers a governance model for all ethical and privacy related aspects.

Topic 2

- Discover predictors of the metabolic complications of adult and paediatric obesity, which shall lead to innovative diagnostic tests, pave the way to novel therapeutics targeted to high-risk individuals, and provide the infrastructure to select individuals for such targeted pharmacological interventions.

Topic 3

- Determine protective and precipitating factors for conversion from pre-dementia cognitive dysfunction to dementia in general as well as conversion from prodromal Alzheimer's disease to typical or atypical Alzheimer's disease. Determine individual susceptibility as well as precipitating factors that cause very elderly patients to convert from biomarker-positive pre-symptomatic stage (asymptomatic at-risk stage) to prodromal, typical or atypical Alzheimer's disease.

POTENTIAL SYNERGIES WITH EXISTING CONSORTIA

- EHR4CR, PROTECT, PharmaCog, NEWMEDS, SUMMIT and the Diabetes personalised medicine call
- Other EU funded projects (e.g. EU-ADR, epSOS, GRIP)
- Any relevant regional/national initiatives

EXPECTED KEY DELIVERABLES FOR EACH TOPIC

Topic 1: Information Framework / Knowledge Management Service Layer

1. A common data platform that provides access to patient medical information from a number of data sources across Europe. This information base will contain

comprehensive and in-depth patient health data comprising clinical, biomarker and other detailed patient information on a number of populations and disease cohorts. The platform also includes a detailed inventory and description of the different data sources and patient populations included in the network.
2. Solutions in the areas of data privacy and ethics, standards and semantic interoperability as well as legal issues and information technology (IT). The solutions are intended to become a standard for patient health data linkage and for access to a combined patient health information base.
3. A business model that governs the use of the project output.
4. This information base will also include paediatric patient level data, covering all age groups from neonatal ages to adolescents.

Research Projects

Topic 2: Metabolic complications of obesity in adults and children

1. A detailed understanding of the inter-individual variability in susceptibility to the metabolic complications of diabetes (i.e. diabetes, dyslipidemia, hypertension and liver steatosis) among adult and paediatric obese individuals from various ethnicities across Europe.
2. The characterisation of the effect of constitutional (age, ethnicity, gender), environmental (lifestyle, diet, exercise) and obesity-specific factors (type of obesity, age of onset, etc...) on susceptibility to metabolic complications.
3. The identification of genetic susceptibility markers for metabolic complications of obesity.
4. The characterisation of epigenetic and metagenomic (microbiome) susceptibility markers for the metabolic complications of obesity.
5. The characterisation of high-risk individuals for targeted pharmacological (using for instance generic, marketed products) and non-pharmacological interventions (including nutriceuticals) with the prospect of having one interventional study running.
6. The development of an algorithm leading to a diagnostic test that would predict high risk for the metabolic complications of obesity.
7. Possibly the identification of novel targets or pathways for future therapeutic interventions.

Topic 3: Protective and precipitating markers for the development of AD and other dementias

1. A detailed understanding of the inter-individual variability in the susceptibility to AD and other dementias among individuals in various ethnicities across Europe.
2. The characterisation of the effect of constitutional (age, ethnicity, gender), environmental (lifestyle, medical and treatment history) and cognitive-specific factors (type of cognitive impairment, age of onset, etc...) on susceptibility to the development of AD and other dementias.
3. The identification of memory test, pathophysiological and topographical susceptibility markers for the development of AD and other dementias.
4. The identification of epigenetic and meta-genetic susceptibility markers for the development of AD and other dementias.
5. The characterisation of high-risk individuals for targeted pharmacological and non-pharmacological interventions.

6. The development of an algorithm leading to a diagnostic test for the identification of mild cognitive impairment (MCI) individuals at high risk for the development of AD and other dementias.
7. Possibly the identification of novel targets or pathways for future therapeutic interventions.

CONSORTIUM

EFPIA PARTICIPANTS

GlaxoSmithKline (co-ordinator), Johnson & Johnson, Servier, Pfizer, Amgen, Boehringer-Ingelheim, NovoNordisk, Orion, Roche, Lundbeck.

Indicative duration of the project

The indicative duration of the project is 5 years. However, it is the intent to expand the project by adding further research projects to the EMIF infrastructure in calls 5 and 6. Furthermore, the aim is also to develop an infrastructure that is sustainable so that EMIF will be a key information base for research projects beyond the life of these particular IMI calls.

Indicative budget for the project

Indicative total in kind contribution from the EFPIA companies is EUR 24 million. It is expected that a further 40 million Euro will be available to support the additional research projects to be launched under calls 5and 6.

Applicant Consortium

(To be selected on the basis of the submitted EoI)

An Applicant Consortium applying for one of the topics of the Call Theme will not be obliged to apply for the other two topics. However, due to the integrated nature of the topics, the successful consortia (one for each topic) are expected to work together to develop one Full Project Proposal.

Indicative expectations from the Applicants, by topic

Topic 1: Information Framework

- Patient-level data sources including adult and paediatric data. (The aim is to include a number of data providers in the applicant consortium, so that some data will be available from the outset and from within the consortium. Further data providers will be included throughout the course of the project, with the exact nature of the relationship with the overall consortium to be determined.)
- Provide evidence of existing excellence in health-related research informatics in a high-quality multidisciplinary research environment to develop and sustain critical mass in research using e-health records.
- Undertake an innovative programme of electronic health data linkage and analysis that aims to lead to improvements in patient or population health. Linkage of datasets including research data, health and socio-economic records using existing and emerging regional and national infrastructures must be core to research activities.
- Collaborate and provide expertise and solutions on the associated issues of data linkage such as data privacy, ethics, data privacy, semantics and IT, data standards and ontology, intellectual property and legal aspects, integration of patients' associations (including parents and children).
- Expertise in observational patient health data quality and analysis

Topics 2 and 3: Research projects

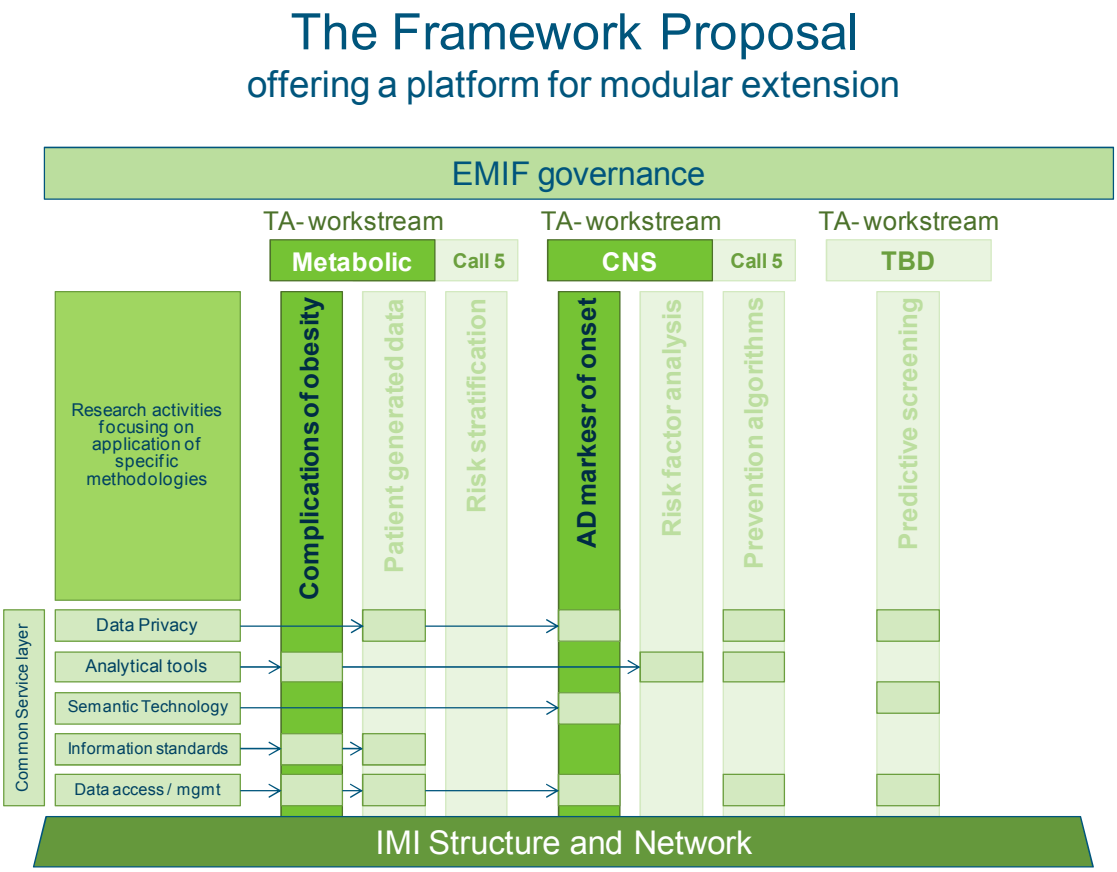
- Access to patients/data pertaining to obesity and/or Alzheimer's
- Research expertise in epidemiology including molecular epidemiology
- Expertise in genetics and utilisation of extreme phenotypes approach
- Analysis of patient-level observational data
- Disease expertise (metabolic, Alzheimer's Disease)
- Expertise in paediatrics, in particular in metabolic diseases including diabetes, and in study design and data analysis.
- Expertise in health outcomes research
- Expertise in metabolomics and other 'omic' sciences
- Imaging expertise and the conduct of imaging tasks
- Expertise in the development of diagnostic tests
- Access to patients if possible
- Interest in and capability to carry out clinical trials

SUGGESTED ARCHITECTURE OF THE FULL PROJECT PROPOSAL

Each Applicant Consortium is expected to address the research objective for the chosen topic and make key contributions to the defined deliverables in synergy with the EFPIA consortium.

The below architecture for the full proposal is a suggestion; different innovative project designs are welcome, if properly justified.

Proposed Work plan
The graph below illustrates the proposed set up and structure of the overall EMIF project.



Governance of the EMIF project

Governance of the interface between the research projects and the development of the information platform:

To ensure smooth collaboration and coordination between the 'horizontal' service layer and the 'vertical' research projects, some guiding principles have been developed. It will be important to further refine these through the work that needs to be done in the work-streams.

Initial guiding principles

- An overarching governing body will be developed to oversee all projects utilising the EMIF platform;
- Research projects have final responsibility for their research programme, approach etc… i.e. the 'framework' is supportive to the research projects;
- Research projects have their own timelines and governance structures to ensure delivery of research objectives;
- There is shared ownership between the 'vertical' research projects and the 'horizontal' framework;
- Research projects should identify their needs for data, but can rely on work done by the horizontal service layer;
- All projects will provide feedback on what they have learnt / insights into the framework – this will enrich new projects;
- Solutions are developed that allow 'two-way' communication with data owners (sources);
- Execution of the project-based research is the overall priority and the framework offers 'services' to the projects but should also do 'research' into further enhancements of the horizontal solutions, e.g. around standards, interoperability, analytical approaches, etc.;
- Having the EMIF framework will allow future research projects to go across the multiple data-sources, thereby leveraging data collected and work done in the various 'vertical' research projects.

Structuring the projects into Work Packages (WPs)

Topic 1: Information Framework (approx. 50% of the total indicative budget)

While starting with a focus on the needs of the initial two research projects, the EMIF knowledge management service layer will gradually grow with the needs of future research projects. That way, the overall vision of a sustainable framework for the support of patient level data research will gradually be realised. Crucially, this project will utilise output and the lessons learned from the IMI project EHR4CR and from paediatric networks.

Given the important and continuous developments in the overall field of healthcare and bio-informatics, it is important to also have an ongoing body of research into these fields. This will ensure that the horizontal knowledge management layer continues to evolve and makes the most recent developments available for the research projects.

Work Package 1 – Governance and business model

- Develop a vision for the future while identifying the growth path and milestones towards the gradual realisation of this vision;
- Define the overall workings of the EMIF framework;
- *Define potential future research projects for inclusion in the platform*;
- Define the business model to turn EMIF into a lasting framework;
- *Define study population data requirements et identify suitable data sources.*

EFPIA Contribution:

- Business model knowledge
- Vision development (Future Sketching)
- Programme management

Work Package 2 – Landscape exploration

- Evaluate the data sources that are available as part of the consortium;
- Identify existing data sources in adults and children that could be used for intended research projects;
- Evaluate the identified data sources, including for data quality and completeness and missing specific information detail;
- Establish contacts and contracts for data access with selected data owners outside the consortium.

EFPIA Contribution:

- Overview of existing data-networks
- Input of data from networks where EFPIA partners have an existing relationship
- Leverage of contacts leading to patient-level data

Work Package 3 – Legal context

- Map out the legal issues around access to and use of patient level data across Europe;
- Define the ethical and legal process to embark on new research projects and specify any ethical and legal issues relevant to the paediatric population'
- Define the 'data life cycle' process, i.e. what need to happen before accessing data, who can use which data for what research purpose and what happens with data after completion of the research.

EFPIA Contribution:

- Legal support
- Experience in working with ethical committees
- Contacts to experts in the domain

Work Package 4 – Architecture and Solution development

- Define architecture for the EMIF platform;
- Development of an environment that allows for access to and integration of diverse and rich patient level data sets, including the linkage of different data sources to achieve comprehensive patient records;
- Ensure EMIF framework can work in alignment with other data platforms such as EHR4CR, eTRIKS and others.

EFPIA Contribution:

- Leverage from other IMI projects
- Expertise in data management and in data standards

Work Package 5 – Data protection, Privacy and security

- Explore how the solution under development for EHR4CR can be reused for EMIF;
- Ensure a privacy robust two-way communication is possible, i.e. from data owner to EMIF and if needed back to data owner.

EFPIA Contribution:

- Leverage from other IMI projects

Work Package 6 – Data preparation and integration

- Develop methodologies and solutions for data access (data storage options to be evaluated);
- Identify data standards required for a smooth integration of data in a way that (over time) cross-therapy area research becomes possible;
- Identify leverage opportunities and gaps with the solution under development for EHR4CR.

EFPIA Contribution:

- Expertise in data management and in data standards
- IT architecture expertise

Work Package 7 – Analytical methodologies

- Identify potential analytical solutions that will support the research projects;
- Work with research project leads to identify statistical and data visualisation needs.

EFPIA Contribution:

- Expertise in analysis of longitudinal patient records
- Access to experts in the field
- Analytical platforms that could be embedded in EMIF framework

Topics 2 and 3: Research projects (approx. 50% of the total indicative budget)

For both topics the same work packages are anticipated with adaptation to the disease area.

Topic 2 Metabolic complications of obesity	Topic 3 Protective and precipitating markers for the development of Alzheimer's disease and other dementia
Work Package 1 : Define study population data requirements and identify suitable data sources	
• Identify the type of data required or that is desirable to address the objectives of this research project. In addition to capturing demographic and the typical clinical data, special attention will be paid to obtaining information including longitudinal information on anthropometric parameters (such as BMI, waist-to-hip ratio, resting blood pressure), biochemical parameters (such as HbA1C or glucose/insulin/C-peptide or creatinine levels) and imaging data (such as fat content of the liver, fundoscopy, echocardiography data if available). • Work in partnership with the Information Framework team to do the following. o Identify and evaluate suitable data sources for adults and children including information quality and data gaps / missing details. Such data sources include patient level data from routine clinical practice including patient registries, and experimental data. o Assess the scope and feasibility of collating and collecting additional patient level health information and carry out the additional data collection as appropriate. o Gain access to the different data sources from a central platform. EFPIA Contribution: disease expertise and information requirements identification and evaluation of patient-level observational data sources epidemiologic research expertise expertise in paediatrics	• Identify the type of data required or that is desirable to address the objectives of this research project. In addition to capturing demographic and the typical clinical data, special attention will be paid to obtaining information on cognitive parameters (such as episodic or working memory loss, MMS-E) and other clinical parameters (such as resting blood pressure, BMI, cardio-vascular or diabetic comorbidity), socio-demographic parameters (such as education level, familial status), biochemical parameters (such as systemic inflammatory markers) and imaging data (such as structural MRI) if available. • Work in partnership with the Information Framework team to do the following. o Identify and evaluate suitable data sources including information quality and data gaps / missing details. Such data sources include patient level data from routine clinical practice including patient registries, and experimental data. o Assess the scope and feasibility of collating and collecting additional patient level health information and carry out the additional data collection as appropriate. o Gain access to the different data sources from a central platform. EFPIA Contribution: disease expertise and information requirements identification and evaluation of patient-level observational data sources epidemiologic research expertise

<table>
<tr><th colspan="2">Work Package 2 : Characterise the study population</th></tr>
<tr><td>Focus on obesity and metabolic complications

• Describe the distribution of patient demographics, medical history and other relevant characteristics.
• Identify the obese subpopulation in the adult and paediatric population from birth to adolescence and describe the distribution of the specific anthropometric measures to characterise the heterogeneity of obesity in the subpopulation.
• Describe the distribution of metabolic complications in the obese subpopulation
• Perform a detailed analysis of the inter-individual variability in the susceptibility to diabetes and liver steatosis among adult and paediatric obese individuals, identify constitutional factors (age, gender, ethnicity), environmental factors (diet, sedentarity) and obesity-related factors (fat distribution, obesity shape, age of onset, family history....) associated with these complications.

EFPIA Contribution:
Disease and other scientific expertise
Epidemiologic research expertise including study design and statistical methodology
Data management and analysis</td><td>Focus on mild cognitive impairment or prodromal signs of AD in elderly patients, presymptomatic markers in very elderly patients

• Describe the distribution of patient demographics, medical history and other relevant characteristics.
• Identify the MCI/prodromal AD subpopulation, describe the distribution of the specific measures to characterise the heterogeneity of cognitive impairment, alteration of instrumental activities in daily living and comorbidities in the subpopulation
• Investigate the presence of impairment on specified memory tests, pathophysiological (amyloid β_{42}, total Tau, Phospho-tau in CSF, PET amyloid tracer uptake) and topographical (^{18}FDG PET, structural MRI) biomarkers and analyse their distribution in the elderly and very elderly subpopulations in centres where these markers can be measured.

EFPIA Contribution:
Disease and other scientific expertise
Epidemiologic research expertise including study design and statistical methodology
Data management and analysis</td></tr>
<tr><th colspan="2">Work Package 3 : Identify extreme phenotypes</th></tr>
<tr><td>Selection of individuals at the extreme of the distribution for a particular trait. Trait may be HbA1C or fasting glucose and liver fat content (if available) or liver function tests. Age can also be used to identify extreme phenotypes. Individuals from one extreme will be matched with individuals of same age, gender, ethnicity, BMI and other confounding factors identified in WP2 to select opposite extremes. In addition, people will be selected based on access to biological samples (genomic DNA), capacity to contact and collect additional data/samples.</td><td>The probability of subsequent development of clinical AD can be substantially increased by recruiting asymptomatic individuals with a combination of risk factors (e.g., older age, APOE ε4 genotype, increased retention of amyloid radioligand in the brain, and evidence of age-related decreased volume of the hippocampus) but without achieving certainty that AD will develop or by when.
• Perform a detailed analysis of the inter-individual variability in the susceptibility to rapid or slow conversion from prodromal or pre-symptomatic AD to symptomatic</td></tr>
</table>

<table>
<tr><td>EFPIA Contribution:
Disease and other scientific expertise
Epidemiologic research expertise including study design and statistical methodology
Data management and analysis</td><td>predementia stage or AD respectively.
• Identify constitutional factors, environmental factors, biological factors (e.g. blood CRP, cytokines, CSF amyloid β_{42}, total Tau, Phospho-tau) and genetic factors (e.g. APOEε_4) and analyse their distribution in the relevant population
• Select individuals at the extreme of the distribution for a particular trait, symptom cluster or marker.
• Individuals from one extreme will be matched with individuals of same age, gender, ethnicity, BMI and other confounding factors to select opposite extremes. In addition, people will be selected based on access to biological samples, imaging equipment, capacity to contact and collect additional relevant data/samples.

EFPIA Contribution:
Disease and other scientific expertise
Epidemiologic research expertise including study design and statistical methodology
Data management and analysis</td></tr>
<tr><td colspan="2">Work Package 4 : Characterise extreme phenotypes</td></tr>
<tr><td>Patients at the extremes of the distribution will be compared using an extreme-discordant case-control design. Comparison will include :
- omics:
o Genomic: possibly exome sequencing of the extremes, possibly GWAS;
o Meta-genomics : possibly comparison of the metagenomic profile of selected extremes (tbd);
o Meta-bolomics and transcriptomics.
- Clinical chemistry biomarkers: (tbd) needs to be extensive as people are selected already based on extreme metabolic phenotypes.

EFPIA Contribution:
Genetic and other scientific expertise
Study design and statistical methodology</td><td>Patients at the extremes of the distribution will be compared using an extreme-discordant case-control design. Comparison will include:
- Genomic : possibly exome sequencing of the extremes, possibly GWAS (tbd);
- Meta-genomics : possibly comparison of the metagenomic profile of selected extremes (tbd);
- Clinical chemistry biomarkers: needs to be complicated as people are selected already based on extreme phenotypes;
- Imaging biomarkers, when available.

EFPIA Contribution:
Genetic and other scientific expertise
Study design and statistical methodology
Access to ‘omic’ technology platforms
Data management and analysis</td></tr>
</table>

Access to 'omic' technology platforms Data management and analysis	

Work Package 5 : Validate in general population in adults and children
Exploratory markers which will have been identified in WP4 will be tested specifically in additional extreme cases and controls and in the general populations to characterise predictive performances. EFPIA Contribution: Scientific expertise Study design and statistical methodology Access to 'omic' technology platforms Data management and analysis

Work Package 6: Identify and select individuals for pharmacological or non-pharmacological interventions	
with high risk for metabolic complications	**With high risk for conversion to AD**
EFPIA Contribution: Scientific expertise Planning, designing and conducting clinical trials including statistical methodology and data management and analysis as well as interpreting the results	

Work Package 7: Project management and communication
The WP should cover all aspects of project management and coordination, including the dissemination and communication strategy. The results of the project will be disseminated in the form of publications, meeting presentations, and via the consortium's website. EFPIA Contribution: Project coordination and management Communication strategy Dissemination of results

Glossary:

AD: Alzheimer disease; EHR: electronic health record; GRIP: Global Research in Paediatrics; GWAs: Genome Wide Associations; TA: therapeutic area;

2. ETRIKS: EUROPEAN TRANSLATIONAL INFORMATION & KNOWLEDGE MANAGEMENT SERVICES

BACKGROUND

Translational Research (TR) in drug discovery and development requires the management and sharing of knowledge between pre-clinical and clinical activities. In order to answer simple questions, integrated data and data services from diverse domains is required. For example:

- What is the correlation between animal models and human data?
- What is the best biomarker strategy for a given compound?
- What is the best indication for a given compound?
- How could patients be stratified based on clinical data?
- Do clinical data support a target of interest?

As a result many pharmaceutical companies have complex TR Knowledge Management (KM) processes, standards and services in place internally. However, no such collaborative KM service currently exists to support cross-institutional studies such as those found in public-private partnerships like IMI. This gap impairs TR investments in Europe and inhibits the pharmaceutical sector's productivity. Currently, every pre-competitive translational study requires bespoke data management and analysis investments. This results in unnecessary and significant overheads, complex intellectual property, and delays the sharing of existing translational information and know-how.
Johnson & Johnson (J&J) has decided to move tranSMART, a translational medicine platform, into the public domain as a pre-competitively sharable asset. Apart from being used within J&J, the platform has also been tested in US academic medical centres and the IMI (Call 1) U-BIOPRED project with clinical trial/omic data.
However software alone is not sufficient to change TR KM, there is a need for:

- A KM service supporting projects using the software platform: managing TR content long term, curating new & archive content and managing code sharing thus driving platform innovation;
- Open Standards to enable information sharing and system interoperability achieved through community standards research, agreement and adoption;
- development of trust to enable sharing of content, both open sharing of study data as well as limited access through an honest neutral broker;
- active community research in TR analysis and methodology.

The suggested eTRIKS consortium will make the wealth of data generated during TR investments (such as in many of the IMI projects) accessible in a standardised way to the global translational research community. Furthermore, as demonstrated in the bioinformatics and cheminformatics domains, having access to standardised TR data and content stimulates research, innovation and novel approaches in data insight mining.

Expected benefits are:

- improved efficiency: by reducing costs and overheads arising from bespoke solutions, for cross organisation TR;
- stable legacy: enabling IMI TR data security (time and policy) and accessibility;
- high-value resource: stable repository of curated and annotated translational studies;
- improved understanding: through maximising the utility of individual studies; through gaining new scientific insights to support and accelerate medicines development; by fulfilling the ethical responsibility to extract most value for contributing patients and by permitting combined, cross study analyses;
- improved data sharing and interpretation: by developing and supporting independent, agreed and stable public-private standards; by developing and providing common interfaces reducing the threshold for data access to researchers and system interoperability;
- improved analytics: through fostering the research in and development of TR informatics and analysis methods;
- strengthened community of TR informatics and KM professionals.

NEED FOR PUBLIC-PRIVATE COLLABORATIVE RESEARCH

The challenges involved in building and supporting a collaborative translational platform and associated content should be addressed through a cross-pharmaceutical Public Private Partnership (PPP) approach and requires active informatics and KM research in the context of professional level KM services.
Specifically, EFPIA industry know-how, experience, service delivery and problem understanding needs to be combined with academic/SME expertise in methodology and standards development as well as community engagement.
As part of the project eTRIKS needs to develop a plan for becoming a sustainable platform. It is expected that this might take the form of a Research Infrastructure in the context of ESFRI or the formation of an institute of Translational Research Informatics.

OVERALL OBJECTIVES

The establishment and maintenance of a stable consortium to develop and provide the necessary informatics and KM support for TR activities. The focus will be initially across IMI projects, but with an objective to provide services for all EU translational collaborations.

The overall primary objectives of the project are:

- single access point to standardised, user friendly TR study information;
- sustainable, interoperable, collaborative TR platform, based on open, agreed standards;
- development of an active TR analytics and informatics community.

This will be delivered through a number of activities.

- The hosting and establishment of TR KM support and services based initially on the tranSMART platform. The platform is to include data curation / data management, hosting, Quality Control (QC), training, etc.
- Research & development of the platform. For instance; exploring KM solutions for the management of different data types, developing information standards,

establishing interoperability interfaces with related systems and activities globally, and researching Graphical User Interface (GUI) design.

- Innovation through research into TR informatics and analytics approaches to enable biomarker discovery, drug response, patient stratification, disease mechanism understanding, etc.
- Access to relevant legacy translational study information, further enabling knowledge exploitation activities.

Long term, the expectation is that this investment will seed a stable, sustainable platform and service to support academic and commercial TR activities across Europe. This will require additional funding, beyond this initial call: one of the consortium deliverables will be establishing sustainable funding models, for example through additional PPP funding, direct commercial sponsorship, academic grant funding and/or a not-for-profit charge mechanism to supported projects. It is also expected that the platform will be extended to support a wide range of modalities covering not only efficacy but safety TR investigations, possibly requiring further specific investments, possibly in the form of future IMI calls.
Key milestones for review of these long term objectives include:

- Year 1 Milestone: Proposed IMI Call 5 Safety Data Sharing topic, with extensions to the eTRIKS framework specifically to support cross-Pharma/academic pre-clinical and clinical safety data sharing.
- Year 2 Milestone: Assessment of eTRIKS value post 2 years. Review of impact for Euro spent, along with assessment of other existing KM Service infrastructures and overall need for long term, pre-competitive TR KM Services, not only for IMI but other EU PPP. May require an up-scaling of eTRIKS funding in year 3 to support potential increased demand if proving value for money. Publication of proposal for eTRIKS long term (5+ years) future including closure/restructuring, continuation as is (with sustainable funding models) and/or partnerships/ engagement (or merger) with other infrastructure initiatives (EU & US).
- Year 4 Milestone: Repeat of year 2 assessment to confirm value and future planning.

POTENTIAL SYNERGIES WITH EXISTING CONSORTIA

Two IMI projects in Europe (U-BIOPRED, OncoTrack) as well as The Sage Bionetworks/Genetic Alliance-led Clinical Trial Comparator Arm Partnership (CTCAP) in the US are currently implementing the platform. Furthermore, the platform is being considered for implementation by other IMI project consortia (SAFE-T, PharmaCog, eTOX).
It is expected that during the lifetime of the CTCAP consortium, a large amount of tranSMART development work will be undertaken. Therefore, co-ordination and synergies to be gained by coordinated EU- and US-based translational research platforms will allow access to a large body of translational data.
Similarly this project should be closely coordinated with ESFRI infrastructure investment proposals (EATRIS, ECRIN, ELIXIR, BBMRI), which are part of the European Community BioMedical Sciences (BMS) research infrastructures strategy.

Synergies with existing consortia beyond the ones mentioned above:
1) DDMore, EHR4CR and Open PHACTS for semantics standards for common concepts/entities;
2) EMIF, with respect to coordination of the interface between clinical and patient record information and semantic standards.

EXPECTED KEY DELIVERABLES

- Open TR KM and analysis platform (evolved tranSMART).
- Hosting and management of the tranSMART production platform.
- Research and development of the platform and associated services, including the identification and adoption of existing assets.
- Service wrapper around tranSMART providing data management services to EU TR studies and investments:
  - curation support for active TR trial data loading & hosting (clinical, omic, genetic data types are a minimum);
  - curation services for historic TR data;
  - development of guidelines and best practices for data curation;
  - quality control processes and services.
- A high-value, stable repository of curated and annotated translational studies (current and historic).
  - Public and restricted content.
- Query and analytical tool research and development.
- Support for internal mirroring of the platform and data export facilities.
- Training.
- Independent, published and adopted TR KM standards (semantic and data exchange standards).
- Governance policies and structure regarding data, security, standards.

CONSORTIUM

EFPIA PARTICIPANTS

AstraZeneca (co-ordinator), Johnson & Johnson, GlaxoSmithKline, Roche, Pfizer, Sanofi-Aventis, Bayer, Lundbeck, Merck-Serono, Eli Lilly.

Indicative duration of the project
The indicative duration of the project is 5 years.

Indicative budget for the project
Indicative total in kind contribution from the EFPIA companies is EUR 9.880.000 .

Applicant Consortium
(To be selected on the basis of the submitted expression of interest)

Given that the focus of this call is on both TR KM research and service provision, the consortium should have a small number of participants (3-5) with a focus on a hub model whereby the bulk of funding goes to a single institute/organisation. The expectation is that the hub organisation will host, research and develop the core standards and services required with additional funding going to 'spoke' organisations to research/develop analytics methodology, standards, visualisations, etc. to be deployed alongside the core platform. This centralisation is desired given the challenge involved in the coordination

and delivery of the required services and infrastructure. However, alternative proposals to deliver the same effect will be considered.

The Applicant Consortium is expected to have demonstrated ability in a number of key areas.

- **TR KM Service provision** – operating a data service in the area of translational, medical or bioinformatics. Specifically, success in setting up robust database infrastructure, establishing policies around data access, providing user outreach and training services, managing complex multi-discipline user base interactions. Combinations of project managers, database administrators, Extract, Transfer & Load (ETL), release and configuration engineers, software engineers, QC engineers and trainers are expected to be available for these tasks.
- **TR KM research** - experience and expertise in developing novel methods and application of data mining and knowledge management techniques in translational informatics (bioinformatics and medical informatics), development of new methods to mine the translational data and continuously disseminate and publish the results and train the scientific community.
- **Platform provision and maintenance** - experience and staff to establish and maintain an appropriate large-scale hardware platform to manage all translational data. The size of the data resource is expected to grow into the 100s of TB requiring robust processes and services. The consortium shall provide all necessary secure facilities and hardware infrastructures.
- **Data warehouse and ETL** – experience in applying best practices necessary for successful establishment of data warehouse and specific experience in large scale ETL activities.
- **Software engineering** - expertise and experience in developing and maintaining professional software engineering practices. All the code developed during the run of the project will have to conform to software industry best practices and it will be the responsibility of the Applicant Consortium to adhere to those and also to enforce them across the community. Best practices shall include project management, adherence to professional code writing standards, issue tracking, code repositories, QC/QA, documentation, etc.
- **Governance/policy** – experience in setting up and maintaining robust data governance and best practice policies around a large scale data warehouse including data stewardship, access, provenance, governance and security.
- **Curation** – experience in manual cleansing and curation of biomedical data including molecular profiling experiments. Initially, most of the clinical and *in vitro/in vivo* data is expected to be curated manually so robust processes including QA/QC are paramount. Later, it is expected that automation of these tasks shall take place so experience in automating cleansing and curation shall be desired. The resource to be established will be as good as the data available through it. Therefore, it is paramount that the curation service establish guidelines and best practices for the provision of cleansed, standardised data of the highest quality to the community through the duration of this Call.
- **Clinical data management** - experience in developing database strategies and ontologies as applied to clinical data management and in applying and developing novel methods for analysing clinical data. It is expected that novel methodologies for managing and analysing/mining clinical data need to be researched and/or implemented during the duration of the project.
- **Omic/genetic data management** - experience in database strategies and ontologies applied to omic data management and applying and developing novel methods for analysing omic data, including gene expression, protein profiling, SNP, CNV, sequencing (DNAseq, RNAseq, CHIPseq, etc.), proteomics, metabolomics, lipidomics, etc. It is expected that novel methodologies for managing and

analysing/mining omic data need to be researched and/or implemented during the duration of the project

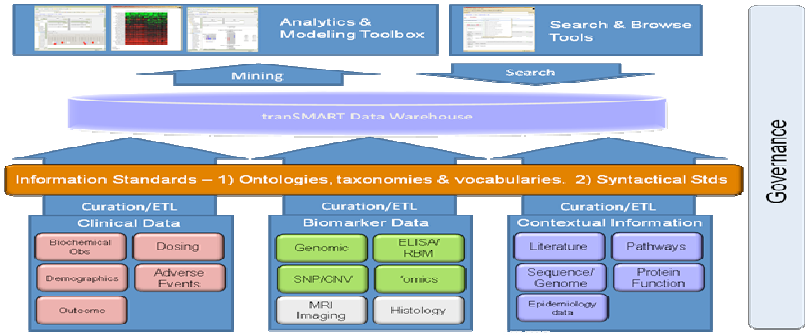
- **Image data management** - experience in database strategies and ontologies applied to image data management and applying and developing novel methods for analysing image data, including histology, MRI, etc. It is expected that novel methodologies for managing and analysing/mining image data need to be researched and/or implemented during the duration of the project.
- **Visualisation/GUI design** – experience in developing novel methods and algorithms and robust software engineering expertise for data access graphical user interfaces and high content/high volume interactive visualisation. Novel GUI paradigms have to enable easy access to the data so that the widest scientific audience can extract value from the translational research resource.
- **Standards development** – expertise in developing and implementing ontologies as well as scientific standing to be a champion in leading standardisation efforts in particular for the task of defining and championing an *in vitro* and *in vivo* ontology. Moreover, a translational ontology/dictionary shall be established to manage translation between clinical trials/clinical observation ontologies (e.g.: CDISC SDTM) and *in vitro* and *in vivo* ontologies.
- **Patient advocacy and policy** – the high value data managed and accessed through this data resource is expected to be of high interest for patient advocacy and policy groups. Moreover, clinical data and associated molecular profiling/biomarker data is governed by complicated regulations in different countries. Experience and expertise to develop appropriate policies and work with government regulators and patient advocacy groups to advance the enhancement of the regulatory environment for enabling improved pre-competitive sharing of translational data.
- **Community reputation and outreach** – this resource will be the first EU-wide translational research knowledge management system. Responsibility, standing and reputation to effectively advocate this resource in the scientific community, organise meetings and trainings and play the role of change agent.
- **Virtual translational research platform coordination** - interaction and coordination with US-based development activities based on tranSMART as well as the constituent open source software packages (i2b2, GenePattern, etc.). It is expected that code bases will have to be merged/integrated/updated on a regular basis so that maximal leveraging of the world-wide investment in this translational research platform can be achieved.
- **Software integrator experience** - experience in developing software in a virtual community mode so best practices as applied to community software engineering (source code version control, issue tracking, QC'ing and integrating community source code) and providing help, tutorials, and software engineering training on the system. It is expected that software developed through other community efforts as it relates to tranSMART will be available for enhancing this platform just as the development capabilities in eTRIKS may benefit US-based or tranSMART development efforts outside IMI. Leadership and participation in community-wide scientific software engineering exchange.

SUGGESTED ARCHITECTURE OF THE FULL PROJECT PROPOSAL

The Applicant Consortium is expected to address all objectives and make key contributions to the defined deliverables in synergy with the EFPIA participants.

The below architecture for the full project proposal is a suggestion; different solutions are welcome, if properly justified.

TranSMART is a translational medicine data warehouse with associated processes implemented within J&J using open source components such as i2b2 and GenePattern. It integrates internal, public and in-licensed data from non-/pre-clinical and clinical studies and enables querying and mining of subject level phenotypic and genotypic data in a secure and scalable environment. It employs public and private dictionaries and ontologies to support user-friendly hypothesis testing workflows.



The adoption of tranSMART is the key differentiating element of this Call Topic. By starting from a proven platform, the consortium can rapidly deploy a platform to address many of the needs of a cross-institute, study-centric KM platform to support translational research in a matter of months as opposed to years if starting from scratch. What is required from this call is a consortium able to build on this success and deliver a mature, evolving TR KM service and knowledge exploitation platform.

It is expected that the tranSMART platform will rapidly evolve (research) once deployed into multiple studies with further requirements for enhancements and modifications fed back into the core service. These requirements will need to be triaged, managed and delivered by the hosting consortium, for example replacing components for existing services, improving search and analytics services, improving export functionality, platform security, platform portability etc. In addition a key area for evolution will be translational information standards adoption in the system to ensure content and system integration and interoperability. A close relationship with existing standards bodies in this domain is essential.

Data types to be handled include, but are not limited to:

- clinical data: demographics, medical history, clinical endpoints (efficacy and safety), clinical lab results, pharmacokinetic data;
- *in vivo, ex vivo*;
- biomarker data: mRNA, RBM – rules based medicine, immunohistochemistry, ELISA, proteomics, candidate SNPs, SNPChp, metabolomics, flow cytometry, whole genome sequence, miRNA, CTCs, epigenetics, lipidomics, imaging (MRI, PET, histology);
- contextual knowledge: literature, patents, pathways, reference genomes, protein function.

The following is an example of SMART Objectives considered for a team of 10-12 FTE which should help to understand the project.

- TranSMART installed and operational within 3 months of initiation
- Year 1: governance and prioritisation board established
- 10 historic translational studies curated first year
- 15 historic translational studies per subsequent year
- 10 active translational studies supported in year 1
- 20+ active translational studies supported per subsequent year
- Year 2: translational standards publication and adoption
- Year 2: platform mirrored in more than 2 other institutes/organisations
- Year 2: platform interoperable with other key global translational platforms
- 5 significant enhancements (2 FTE month + amount of effort) per year
- Year 3-4: long term sustainability plan published and agreed
- Year 4-5: long term funding secured.

Work Package 1- TR KM Study Service Delivery

This work package is to provide support to IMI and other translational studies in providing data and information management services in the form of content upload, management and curation. In the first generation of the service this may be labour intensive while the upload standards are developed. A key measure of success will be the number of translational studies supported by the service in a given period.

- delivery of supported, stable, standardised services;
- system functionality (GUI & Application Programming Interface (API));
- data services: e.g. curation support;
- help services: user support;
- GUI and technical training.

EFPIA Contribution*:*

- Transfer of most current tranSMART code and associated processes (J&J)
- Training on tranSMART system and processes – installation/maintenance, software engineering, ETL, curation and end-user trainings (J&J)
- Database/IT technology expertise
- Curation and ETL expertise
- Pre-competitive curated data (not counted as in-kind)

Work Package 2 - Translational KM Platform Research

- Development of tranSMART core architecture
- Development/modularisation of data management modules, especially missing data components (as determined by trials being supported): notably imaging data, histopath data types;
- Coordination of the output of a federated development community (beyond the Call).

EFPIA Contribution:

- Informatics expertise (clinical and pre-clinical)
- Translational / Clinical Science expertise
- Some tools, methods, standards (all)

Work Package 3 - Translational Information Standards Research/Coordination:

Given this service will be on the 'sharp end' of translational information standards use it is expected to play a pivotal role in both standards development where there are recognised gaps, and adoption/evolution where there are existing standards, both for syntactic and semantic standards.

- Development of community agreed information and data standards, including semantic standards. So allowing system interoperability, modularisation of components and data integration.

EFPIA Contribution:

- Standards expertise
- Translational / clinical science expertise
- Some tools, methods, standards (all)

Work Package 4 - Translational Analytics Research

- Development of TR analytics; algorithms, GUIs, workflows.

EFPIA Contribution:

- Informatics expertise (clinical & pre-clinical)
- Translational / Clinical Science expertise

Work Package 5 - Governance and Business model

- Develop and establish a governance group and process to oversee:
  - service and development demand from TR studies prioritisation, e.g. prioritisation of historic content curation, community code and data sharing coordination and adoption, standards development and adoption;
  - information module integration, possibly the output from future IMI KM calls.
  - development and execution of legal and security frameworks: ensuring privacy compliance as well data security models for limited access data;
  - coordination of platform and content versioning across tranSMART instances.
- Develop a sustainable business model, exploring options for funding; both long-term as well as how to scale the service beyond the capacity enabled within the Call core

funding. If successful, expected demand for support will outstrip core funding so will need alternative funding models, such as charge-for-service, for content curation/hosting etc.

EFPIA Contribution:

- Data governance expertise
- Legal, compliance and regulatory expertise

Work Package 6 - Community Engagement & Outreach

- Communication and awareness activities to promote adoption and community platform build.
- Fostering interaction between the traditionally distinct pre-clinical informatics (typically bio) and clinical informatics communities to enable TR KM innovation and problem solving.

EFPIA Contribution:

- To be determined

Glossary: TR: Translational Research; KM: Knowledge Management; GUI: Graphical User Interface; QC: Quality Control; EMIF: European Medical Information Framework (Call 2011-4): ETL: Extract, Transform and Load; API: Application Programming Interface; PPP: Public Private Partnership.

CHEMISTRY, MANUFACTURING AND CONTROL

3. DELIVERY AND TARGETING MECHANISMS FOR BIOLOGICAL MACROMOLECULES

BACKGROUND

Therapeutic modalities based on macromolecules of biological origin such as proteins, peptides and oligonucleotides have a huge pharmacological potential due to their highly selective mode of action and their potential activity against targets that are considered 'non-druggable' by more traditional small organic molecules. However, the widespread application of many potential macromolecular therapeutics has been very limited due to pharmacokinetic and drug disposition limitations at both the tissue and cellular level. Therefore, their success in developing future innovative medicines will heavily depend on improvements in both chemistries and delivery technologies that can address these limitations. For instance, biological macromolecular drugs with intracellular targets will only be effective after systemic administration if they avoid degradation, hepatic/reticuloendothelial system uptake and renal excretion, traverse the microvascular endothelium, cross target cell membranes and escape degradation in the endosome-lysosome system. Subverting these barriers has contributed to the requirement for large doses and low therapeutic margins observed in both animal experiments and clinical trials. As a result, biological macromolecular drugs may have issues of therapeutic ratio that could be mitigated by improvements in selective delivery.

NEED FOR PUBLIC-PRIVATE COLLABORATIVE RESEARCH

Innovative strategies for chemical stabilisation and delivery of biological macromolecules can only be developed and tested with the availability of broad expertise and appropriate resources. This can only be delivered by a cross-functional/cross-institutional consortium of academic, SME, regulatory and industrial scientists to foster a better understanding of:

- drug development (e.g. of biological macromolecules in a pre-clinical or clinical experimental medicine setting);
- molecular and cellular biology of cellular uptake mechanisms of macromolecules;
- protein and nucleic acid chemistry, e.g. for conjugation with targeting molecules;
- manufacturing and characterisation of biological macromolecules;
- nanotechnologies.

There are also patient interest groups, particularly for rare diseases, who would press for and encourage this approach to the delivery problem.

Potential Impact on the Pharmaceutical Industry

Improving therapeutics by providing a step change in delivery solutions to tailored macromolecular drugs of biological origin, including oligonucleotides, will pave the way for new treatments in areas of major medical need, e.g. cancer, diabetes, Alzheimer's disease and rare genetic diseases. It is anticipated that solving the delivery problem will afford biological macromolecules a similar or potentially greater level of impact than small molecules for the treatment of diseases.

OVERALL OBJECTIVES

- Eliminate delivery and targeting bottlenecks for developing novel innovative medicines based on biological macromolecules such as proteins/peptides and oligonucleotides.
- Focus on improved understanding of exploitable transport and targeting systems for biological macromolecules of interest to relevant tissues, cells and intracellular targets, e.g. vascular endothelial cells, extravascular tissues, and intravascular target cells.
- Develop nanocarriers to deliver biomacromolecular drugs to and across epithelial barriers:
    - blood brain barrier (BBB);
    - air-blood barrier;
    - blood-retina barrier;
    - skin barrier;
    - intestinal barrier.
- Develop nanocarriers for oral uptake of biomacromolecular drugs.
- Synthesise and test *in vitro* and *in vivo* a range of nanocarriers for the various biological therapeutic modalities, and select a lead approach using a selection of delivery approaches.

POTENTIAL SYNERGIES WITH EXISTING CONSORTIA

To our knowledge there are no obvious synergies with existing IMI proposals.

EXPECTED KEY DELIVERABLES

- detailed elucidation of cellular uptake mechanisms;
- understanding of variables which determine distribution of biological macromolecules;
- organ and tissue specific nanotechnology-based delivery methods for biological macromolecules; and
- delivery strategies for biological macromolecules via local (e.g. inhalation and oral uptake) and injectable application routes.

CONSORTIUM

EFPIA PARTICIPANTS

Sanofi-Aventis(co-ordinator), GlaxoSmithKline, Astra Zeneca, NovoNordisk.

The EFPIA participants will contribute with biological macromolecules as tool compounds, as needed, e.g. proteins, antibodies, oligonucleotides. They will test the novel nanocarriers in pharmacological studies which will be performed in appropriate animal models. For statistical analysis the EFPIA participants will make available their infrastructures and their expertise. In case there are novel nanocarriers that reach the level of clinical testing, resources and expertise for the design of prospective clinical trials will be provided. Finally, the EFPIA participants will take care of consortium and project management.

Indicative duration of the project
The indicative duration of the project is 5 years.

Indicative budget for the project
Indicative total in kind contribution from the EFPIA companies is EUR 9 million.

Applicant Consortium
(To be selected on the basis of the submitted expression of interest)

The Applicant Consortium is expected to provide both pre-clinical and clinical expertise and the ability to carry out interdisciplinary and inter-sectorial work and to cover the following critical fields:

- molecular biology, biochemistry and chemistry of biological macromolecules (peptides, proteins, oligonucleotides);
- molecular mechanisms of transport of biological macromolecules across biological barriers and cellular membranes;
- mechanisms of intracellular trafficking of biological macromolecules;
- imaging technologies for monitoring and quantification of cellular uptake and intracellular trafficking of biological macromolecules;
- cell and animal biology for the generation of *in vitro* and *in vivo* models for monitoring cellular uptake and intracellular trafficking of biological macromolecules;
- nanotechnology and chemistry for the generation of novel nanocarriers supporting uptake of biological macromolecules across biological barriers and cellular membranes;
- nanotechnology for the generation of novel nanocarriers supporting oral uptake of biological macromolecules;
- novel chemistries for linking biological macromolecules to targeting ligands for targeting to target tissues and cells, and novel chemistries for targeting nanocarriers using biological macromolecules as payloads;
- expertise for manufacturing of nanocarriers and biological macromolecules;
- data management and integration expertise;
- professional project management expertise.

SUGGESTED ARCHITECTURE OF THE FULL PROJECT PROPOSAL

The Applicant Consortium is expected to address all the research objectives and make a key contribution to the defined deliverables in synergy with the EFPIA consortium.

We envision the following six key areas where detailed work packages will be defined within the framework of the overall proposal:

The below architecture for the full proposal is a suggestion; different innovative project designs are welcome, if properly justified.

Work Package 1 - Mechanisms of uptake of biomacromolecules across biological barriers

- Identify the uptake processes across biological barriers *in vivo* for both normal and diseased states:
    - identify and characterise the uptake and transport functions across epithelial barriers;
    - determine cross species homology;
    - identify suitable *in vivo* models.

EFPIA Contribution*:*

Supply of biological macromolecules as tool compounds, as needed, e.g. proteins, antibodies, oligonucleotides.

Work Package 2 - Mechanisms of uptake of biomacromolecules across cellular membranes

- Identify suitable *in vitro* models of transport:
    - develop culture conditions for target cells that mimic the *in vivo* situation;
    - identify suitable cell lines / primary cells to model uptake, transport and escape of biological macromolecules into and across cellular compartments;
    - characterise cells for relevant uptake, transport and endosomal escape functions.
- Develop a methodology to accurately measure payload and delivery vehicle concentrations:
    - develop techniques to sample sub-cellular compartments;
    - develop sensitive quantification methodologies (i.e. dual photon microscopy, BioAFM and improved optical and spectroscopic imaging techniques for all payloads and carriers).

EFPIA Contribution:

Supply of biological macromolecules as tool compounds, as needed, e.g. proteins, antibodies, oligonucleotides.

Work Package 3 - Novel approaches for delivery of biomacromolecules across biological barriers and cellular membranes

- Design and optimise delivery vehicles for biological macromolecules:
  - Define 'ideal' properties of nanotechnology-derived agents and conjugation technologies;
  - investigate new materials to fulfil requirements of nanocarriers for crossing biological barriers and cellular membranes.
- Investigate novel approaches for targeting biomacromolecules to disease tissue and/or target cells, either by
  - targeting nanocarriers containing biomacromolecular payload; or
  - directly conjugating biomacromolecules with targeting ligands.
- Investigate new materials for enabling oral uptake and enterocytic delivery of biomacromolecules.
- Demonstrate the *in vitro* effect of delivery approaches:
  - administer novel nanocarriers in carefully controlled experiments with 'standard' payloads in *in vitro* models;
  - determine the functional characteristics of novel materials that enable enhanced uptake.
- Demonstrate the *in vivo* effect of delivery approaches:
  - conduct tolerability and efficacy studies in pre-clinical animal models to confirm *in vitro* findings;
  - determine effects of the route of administration and dosing regime on target penetration.
- Mature nano-formulations and conjugation technologies for potential clinical application.

EFPIA Contribution:

Supply of biological macromolecules as tool compounds, as needed, e.g. proteins, antibodies, oligonucleotides; testing of novel nanocarriers in appropriate disease animal models; pharmacological studies using nanocarriers in appropriate disease animal models.

Work Package 4 (Optional) - Clinical testing

- Demonstrate efficacy of delivery system in the clinic.
- Perform low dose studies to demonstrate efficacy in healthy volunteers (HVs) and patients.

EFPIA Contribution*:*

Design of prospective clinical trial(s): Resources and expertise for trial design.

Work Package 5 - Database and analysis

This work package area should include:

- data analysis and interpretation, including statistics and bioinformatics, of the data generated in work packages 1 to 7;
- mathematically derived, physiologically-based systems models, which integrate multiple data types obtained from *in vitro* and *in vivo* assays with data on compound uptake and clearance *in vivo*, and/or other parameters;

- modelling that builds on approaches used currently to predict drug doses in man and drug-drug interactions may be most productive.

EFPIA Contribution:

Infrastructure and expertise for statistical analysis

Work Package 6 - Project Management

This work package area will address the strategy and implementation of the project management encouraging regular meetings and interaction between sub-groups and teams, to coordinate the work effort.

- Each work package should have a work package leader who will be responsible for ensuring the deliverables are addressed.
- Each work package will encompass different workloads throughout the duration of the project. Progress will be monitored through regular project meetings involving all participants.

EFPIA Contribution:

Overall consortium and project management

4. IN VIVO PREDICTIVE BIOPHARMACEUTICS TOOLS FOR ORAL DRUG DELIVERY

BACKGROUND

Whilst oral administration is the preferred route for the overwhelming majority of drug treatments, the development of oral dosage forms to provide robust performance in patients is very much an empirical process. The lack of validated and predictive assays for oral drug products means that performance can only be verified by *in vivo* testing in animals or clinical studies. This approach is obviously resource intensive both in terms of time and cost. There is a clear opportunity to radically improve this critical area for drug research through the development of new in vitro tools which are truly reflective of product performance in patients. Integrating such tools with validated *in silico* models would allow the accurate prediction of product performance and provide a transformative change in the way oral dosage form development is conducted. The development of new approaches for characterising dosage form performance and building predictive models to 'change the game' regarding development of oral drug delivery products are at the core of this proposal.

Pharmaceutical product design during drug development is a critical step in the translation of pre-clinical research to successful human trials and products on the market. Modern high throughput drug discovery approaches yield highly potent and selective molecules but it is clear from experience across the industry that molecules identified as hits from such screens are becoming increasingly difficult to formulate due to poor molecular properties such as hydrophobicity and low aqueous solubility. In response to this, formulation science and technology has delivered a number of diverse approaches such as nano-sized materials and stabilised amorphous systems to overcome the obstacles of low solubility and dissolution rate limitations which would otherwise impede drug development.

However, current *in vitro* characterisation tools, such as conventional dissolution systems, simply do not replicate the rapidly changing dynamic environment of the gut lumen well enough. Additionally, the physiological variability and complexity observed on meal ingestion cannot be accurately reproduced by current methods. The result is that oral formulation development is often empirical in nature with many *in vivo* studies in preclinical species required to confirm formulation performance. However the resemblance between animal models and humans with respect to critical factors is not fully understood. Therefore it is important to define those areas and models where truly predictive data are generated that can add value over current tools.

The lack of predictive biopharmaceutical tools also limits the ability to fully utilise the power of biomodelling software to predict formulation performance. The use of predictive tools to understand and evaluate how changes in drug product processing or manufacture impact clinical performance is a key component of the Quality by Design (QbD) concept. There is a clear opportunity to embed the use of better or novel predictive tools in a regulatory framework and reduce the significant economic burden of empirical testing when developing the QbD design space for commercialisation of drug products. Drug development costs and time can also be reduced through the use of the Biopharmaceutical Classification System (BCS) to facilitate the replacement of *in vivo* bioequivalence studies with *in vitro* testing for a limited number of drugs (BCS class I). It would be desirable to further extend the BCS-based biowaiver provision through a much better understanding of the bio-relevance of predictive tools, especially in combination with the enhanced knowledge generated by applying QbD.

When considered together, there is a substantial need to improve *in vitro* methodology to meet these requirements and provide a game-change in our ability to perform predictive *in vitro* and *in silico* tests.

Need for public-private collaborative research

Combining industry experience and resources with academic expertise is essential to improve our understanding of the biopharmaceutics of oral drug delivery and to develop predictive tools which can be used to drive a more efficient product development process and deliver significant clinical benefits through rational product design. A pre-competitive public-private partnership initiative will bring together expertise across multi-disciplinary areas and provide the resources for pooling information on formulation characterisation and clinical product performance to create a world-leading knowledgebase for biopharmaceutics. Industry has historically attempted to this within the confines of single companies or in limited collaborations with academic groups with limited success. In addition, the significant advances in bio-simulation software in the last 10 years now allows the combination of much more *in vitro* test results derived from drug substance and drug product as well as pharmacokinetic and pharmacodynamic *in vitro* and *in vivo* outcomes. This initiative will create a focal point for pre-competitive biopharmaceutics data-sharing and the assembly of shared data would provide an unprecedented resource for industry and academic expertise. Such a resource will be used to drive the development of next generation *in vitro* and *in silico* tools allowing the validation of these new tools, as well as currently available tools, with a large and diverse database of active pharmaceutical ingredients and formulation approaches.

OVERALL OBJECTIVES

The overall objectives of this proposal are:

- define the critical physicochemical, formulation and physiological factors that determine product performance following oral administration of a dosage form;
- develop both experimental and theoretical models which can be used to robustly predict the *in vivo* performance of formulated drug products;
- fully leverage industrial knowledge and experience through pooling existing physicochemical, *in vitro* characterisation, preclinical and clinical data to assess the reliability of currently available prediction methods and to underpin the development of new modelling and simulation tools to improve the accuracy of *in silico* approaches.

POTENTIAL SYNERGIES WITH EXISTING CONSORTIA

This initiative is aligned with a recently submitted FP7 Marie Curie Initial training network proposal titled '*In vivo* Predictive Pharmaceutics training Network' (FP7-PEOPLE-2011-ITN). This proposal is pending review and it is not known if funding will be granted at the time of this application. If successful, the FP7 network proposal would be focused on creating and training a cohort of new researchers in the field of biopharmaceutics. The strategic direction of the proposed FP7 research topics is very much aligned with the longer term research focus of the current proposal.

EXPECTED KEY DELIVERABLES

- A more efficient and data-driven development process for challenging compounds delivered through an improved biopharmaceutics 'developability' assessment in the compound selection phase and less trial and error in formulation development to provide the required target performance profile.
- Replacement, reduction and refinement of animal experimentation in accordance with 3R initiative through the use of validated predictive tools and models for oral drug absorption.
- Definition of clinically-relevant pharmaceutical quality criteria in the context of Quality by Design
- Extend opportunities for biowaivers (*in vitro* based bioequivalence) and reduced need for expensive and time consuming clinical studies.

CONSORTIUM

EFPIA PARTICIPANTS

AstraZeneca (co-ordinator), Pfizer, Novartis, GlaxoSmithKline, Boeringer Ingelheim, Sanofi-Aventis, Bayer, Novo Nordisk, Lundbeck, Orion.

The contribution from companies will consist of in-kind and in some cases cash contributions. Participants will contribute FTEs for the coordination of the consortium, intellectual input and experimental work. The level of this contribution is likely to be around 0.5 to 1.5 FTEs/year over 5 years per company.

Participants will contribute data from clinical and preclinical bioavailability studies including formulation studies, dose ranging studies, absorption studies; food interaction studies including *in vitro* dissolution characteristics of clinical formulations; physico-chemical and *in vitro* biopharmaceutical characteristics of drug compounds included in data bases.

Specialised experimental tools for drug form/formulation physical, physico-chemical and biopharm characterisations will be provided by participants, along with *in silico* tools and data management/analysis tools.

EFPIA participants will provide expertise and experimental support for specialised experimental tools; supplies of model drug compounds for experimental investigations; as well as general biopharmaceutics, pharmacokinetics and pharmaceutics expertise.

Coordination, active participation and input will be provided. Participants will also provide direct funding, supervision and laboratory support to research students, post-doctoral fellows within the scope of the current research programme.

Indicative duration of the project
The indicative duration of the project is 5 years.

Indicative budget for the project
Indicative total in kind contribution from the EFPIA companies is EUR 7.600.000 million.

Applicant Consortium

(to be selected on the basis of the submitted expression of interest)
The Applicant Consortium is expected to be a multi-disciplinary body and to enable effective communication between key stakeholder groups (academia; SME technology service providers and regulatory agencies). The consortium is expected to have an established expertise in the following scientific areas relevant to the proposal:

- an understanding of the variability in physiology of the GI tract, and its influence on oral drug absorption;
- a proven track record in delivery of innovative approaches for the measurement and understanding of complex processes influencing drug dissolution, solubility and precipitation in the GI tract and its *in vitro* prediction;
- a proven track record and publication history in the measurement and understanding of membrane transport and metabolism during the absorption and first-pass processes, and the application of in-vitro drug permeability, metabolism and transporter methods to understand and predict *in vivo* behaviour;
- the application of physiologically based integrated *in silico* absorption modelling and to understand oral absorption;
- mathematics and statistics expertise to build and evaluate quantitative models;
- *in vivo* imaging of events in GI tract and access to other specialised *in vivo* tools such as *in vivo* absorption study tools;
- skills in the interpretation of *in vivo* pharmacokinetic data including expertise in population pharmacokinetics tools.

The applicant consortium will also provide expertise to consolidate data provided by the industrial pharmaceutical companies, to enable the gathering of larger data sets on drugs, formulations and pharmacokinetic parameters (e.g. from early phase human clinical studies), and then to interrogate the data for trends and a wider understanding of factors influencing optimal formulation selection.
Participation of SMEs and companies is considered relevant for evaluation and validation of specialised *in vitro* and *in silico* tools.
It is desirable to include representatives from the European Medicines Agency at a coordination level to facilitate implementation of regulatory opportunities delivered by the programme. A more active level of participation may also be possible. A coordinating role for science-focused industry bodies such as EUFEPS and APS may also be appropriate.

SUGGESTED ARCHITECTURE OF THE FULL PROJECT PROPOSAL

The Applicant Consortium is expected to address all the research objectives and make key contribution on the defined deliverables in synergy with the EFPIA consortium.

The below architecture for the full proposal is a suggestion, different innovative project designs are welcome, if properly justified.

Work Package 1 - Physico-chemical tools

At pre-clinical stage when new molecules are developed, biopharmaceutics assessment giving guidance both to compound selection and early formulation strategy needs to be performed on limited data primarily of physico-chemical nature.

This work package aims to:

- link physiochemical properties with clinical performance, possibly extending into formulation space for bio-enhanced formulation approaches;
- establishing 'rules' relating to developability and suitability for different formulation approaches based on simple molecular/physico-chemical compound properties.

EFPIA Contribution:

- Data from clinical and preclinical bioavailability studies including formulation and corresponding physico-chemical and in vitro biopharmaceutical characteristics of drug compounds.
- Specialised experimental tools for drug form/formulation physical, physico-chemical (including high-throughput/automated capabilities) and biopharm characterisations. *In silico* tools and data management/analysis tools.
- Supplies of model drug compounds and formulations for experimental investigations

Work Package 2 – *In vitro* tools

The ability of formulation scientists to identify an optimal drug delivery approach or predict the impact of manufacturing changes on formulation performance are hampered by the basic nature of the in vitro characterisation tools available. This work package aims to understand the fundamental underlying processes such as oral drug absorption, the performance of current in vitro tests, and to design and develop novel methodologies.

2A. Dissolution

- High quality *in vivo* validations of existing dissolution methods to identify compound/formulation 'space' with acceptable predictability and gaps requiring additional development – e.g. BCS class II acids and bases with high solubility at some gastrointestinal pH; BCS class I and III formulations with slower dissolution profiles than current regulatory requirements for biowaivers; regional variations of pH/solubilisation in small intestine; importance of buffer capacity/components for acids and bases;
- Improved understanding and development/definition of predictive methods for challenging aspects of dissolution like intestinal drug supersaturation/precipitation,

influence of hydrodynamics, food effects, colonic dissolution, drug particle size including the nano-scale and amorphous drug formulations.

- Pharmaceutical validation of more bio-similar but complex dissolution equipment e.g. TIM-1 by TNO, Dynamic Gastric Model by Norwich Food Institute or others (e.g. Heigoldt et al, *Eur J Pharm Biopharm*, in press).

2B Permeability

- Definition of testing strategy to obtain quantitative predictions of human intestinal drug permeability required for accurate absorption predictions by *in silico, in vitro* and *in vivo* tools. Will include aspects such as regional variations, influence of drug ionisation/microclimate pH, active transporters, and unstirred water layer/mucus. The impact of formulation type and vehicles used could also be evaluated.

2C Intestinal stability

- Define and establish correlation vs *in vivo* data of *in vitro* test medium that provide *in vivo* relevant esterase/amidase activity by pancreatic enzymes (USP media too low activity, AstraZeneca data to be published) for testing of drugs and sensitive excipients;
- Refine method and establish correlation vs in vivo data for in vitro method based on anaerobial faecal inoculums for determination of drug/excipient degradation in the colon mediated by bacteria. Should also include aspect of regional variation in the colon.

2C Other

- Understand and predict effects of solubility enhancing formulation principles beyond dissolution effects –e.g. nanoparticles, lipid formulations, formulations using amorphous drug, unstirred water layer effects of colloidal particles.

EFPIA Contribution:

- Data from clinical and preclinical bioavailability studies including formulation studies, dose ranging studies, absorption studies, food interaction studies including *in vitro* dissolution characteristics of clinical formulations. Physico-chemical and *in vitro* biopharmaceutical characteristics of drug compounds included in data bases.
- Specialised experimental tools for drug form/formulation physical, physico-chemical (incl. high-throughput/automated capabilities) and biopharm characterisations. *In silico* tools and data management/analysis tools.
- Pre-clinical experimental *in vivo* models also including intestinal fistulation model for local sampling and drug administration.
- Supplies of model drug compounds and formulations for experimental investigations

Work Package 3 – *In vivo* tools /understanding

Due to the limitations of current *in vi*tro methods, *in vivo* testing is often required to confirm formulation performance. However we do not fully understand how *in vitro* methods reflect the *in vivo* situation and how well *in vivo* models translate to human experience. This work package aims to allow a thorough assessment of the current tools to be made and allow the development of improved test systems which better reflect the performance of oral formulations in humans.

- Understanding of colonic drug absorption and limitations - extending database on human regional absorption studies (Tannergren et al, *mol Pharm* 2009) and analyses vs drug molecular, physico-chemical and biopharmaceutics properties. Another output from the database evaluation should be to determine the benefit of different diagnostics to quantify oral regional drug absorption and how data from such studies relate to data obtained from standard formulation administrations to improve interpretation of regional absorption data.
- Improved physiological input into absorption prediction computer models – fluid volumes, effective intestinal surfaces, mucus etc.
- Elucidation of population pharmacokinetics modelling to predict influence of pharmaceutical factors on drug absorption in patient population to define clinically-relevant pharmaceutical quality criteria
- Establish guidelines for rational use of animal models based on physiological resemblance with man and *in vivo* bioavailability data bases.
- Potentially (if possible to accommodate within budget) create understanding for predicting influence of patient variability – impact of genetic inter-subject differences, e.g. enzyme mutations, age, disease and ethnic factors.

EFPIA Contribution:

- Data from clinical and preclinical bioavailability studies. Physico-chemical and *in vitro* biopharmaceutical characteristics of drug compounds included in data bases.
- Specialised experimental tools for drug form/formulation physical, physico-chemical (incl. high-throughput/automated capabilities) and biopharm characterisations. *In silico* tools and data management/analysis tools.
- Pre-clinical experimental in vivo models also including intestinal fistulation model for local sampling and drug administration.
- Supplies of model drug compounds and formulations for experimental investigations

Work Package 4 - Integrating different models including complex *in silico* tools

The development of the biopharmaceutical classification system 15-20 years ago is one of the major advances in oral biopharmaceutics. Originally, it was developed as basis for regulatory acceptance of biowaivers (*in vitro* based bioequivalence), but has also found multiple applications in the industry as a risk assessment tool (Lennernäs, Abrahamsson *J Pharmacy & Pharmacol.* (2005) 57; 273-285). However, recent findings indicate that the BCS classification is a too conservative or crude approach as exemplified by Muenster et. al. (*Eur J Pharm Biopharm*, Feb 16, ePub ahead of print). Thus, there is a great need to further develop and expand applications of the Biopharmaceutical Classification System, both in a regulatory context as well as an industrial aid, but this has not happened due to lack of focused efforts integrating recent scientific advancements in the area and gathering sufficiently large industrial data bases.

Another important area of progress in recent years has been the introduction of systems biopharmaceutics software like SymcypTM, GastroPlusTM and TIMpKTM integrating the different biopharmaceutical and physiological factors to predict the time course of drug absorption into the system circulation. However the accuracy of predictions is often limited by the relevance of the *in vitro* information used as well as major gaps in the algorithms used. These knowledge gaps prevent the implementation of a more effective,

model based product development. This work package aims to improve *in silico* tools by providing better quality and more data (legacy and novel) from all partners. The focus of the work package will be:

- extended/refined BCS and validation/refinement of Developability Classification System (DCS) (Butler, Dressman, *J Pharm Sci* 2010);
- detailed *in vivo* validation of mechanism/physiologically based commercial prediction software including both human and animal data also including toxicological study exposures in the latter case;
- establishing 'rules' relating to developability and suitability for different formulation approaches based on simple molecular/physico-chemical compound properties.

EFPIA Contribution:

- Data from clinical and preclinical bioavailability studies including formulation studies, dose ranging studies, absorption studies, food interaction studies including *in vitro* dissolution characteristics of clinical formulations. Physico-chemical and *in vitro* biopharmaceutical characteristics of drug compounds included in data bases.

[1] Constable, D.J.C., Dunn, P.J., Hayler, J.D., Humphrey, G.R., Leazer, Jr., J.L., Linderman, R.J., Lorenz, K., Manley, J., Pearlman, B.A., Wells, A., Zaksh, A. and Zhang, T.Y. *Green Chem*., **2007**, *9*, 411–420
[1] Trost, B.M. *Science* **1991**, *254*, 1471-1477

5. SUSTAINABLE CHEMISTRY – DELIVERING MEDICINES FOR THE 21ST CENTURY

BACKGROUND

Both society and the pharmaceutical industry derive substantial benefit from the invention of medicines that allow patients to live longer, healthier, and more productive lives. In addition, pharmaceutical companies are committed to finding novel ways to bring additional value such as developing key medicines with minimum impact on the environment. The concept of Sustainable Chemistry provides an ideal framework upon which to develop synthetic capabilities necessary to develop medicines of the 21st century. The successful development of this emerging scientific field will require substantial innovation and investment and therefore represents an excellent opportunity for substantial industry to academia collaborations.

Due to their molecular complexity and development constraints, medicines are often produced with low mass efficiencies (e.g., less than 1 kg of active ingredient produced per 100 kg input materials). Low efficiencies exist because it is still common in pharmaceutical chemistry to use stoichiometric rather than catalytic quantities of reagents, utilize bulky protecting groups to ensure chemical selectivity, and consume large amounts of solvent to facilitate processing. Such products often require large facilities and have long lead times to produce due to the inefficient use of materials. The end result of these inefficiencies is medicines which are costly to develop and commercially produce. A large environmental footprint from the treatment of waste materials also results.

While material consumption, waste treatment, and costs are obvious 'sustainability' concerns for the pharmaceutical industry, another less apparent concern is the emerging scarcity of precious metals currently used to synthesize most medicines. Ubiquitous catalysts such as platinum are being depleted at alarming rates, with some estimates indicating depletion past economic viability within 30 years. Such precious metal catalysts have been key to improving mass efficiency over the past 50 years, and their disappearance could result in substantial efficiency losses unless substitutes are found.

In order to ensure the sustainable delivery of drugs to patients in Europe (especially as the population ages) and ensure European R&D and manufacturing competitiveness, it is critical to develop innovations in chemical synthesis and manufacture. Methodologies which offer 'green' alternatives, such as through synthetic biology, improved and sustainable catalytic methods, and technological solutions such as flow chemistry, may offer significant improvements but require a long term commitment and investment to develop.

NEED FOR PUBLIC-PRIVATE COLLABORATIVE RESEARCH

The discovery of green and sustainable synthesis methodologies is a long-term (> 3 years) endeavour and currently the expertise predominantly lies with academic researchers. Historically, Pharma R&D have routinely implemented methodologies developed in the public sector rather than driven the 'in-house' discovery of fundamentally new transformations. To increase the efficiency and impact of this research on European medicines development it is essential to combine the skills and

expertise that reside in the public sector with the knowledge of the most problematic reactions that reside within industry.

Public-private collaboration provides European universities, institutes, and industries with an opportunity to develop an integrated, long-lasting, and leading position in 'green' research and development, while providing researchers with the skills and training required to meet future industry challenges. With sustainability being a key focus of many EU member country and world economies, this sustainable chemistry initiative will develop a key talent base important to economic development, while delivering innovative solutions to strengthen the EU pharmaceutical industry.

OVERALL OBJECTIVES

The overall aims of this project are two-fold:

- The discovery of new synthetic methodologies for the development and manufacture of small molecule medicines which are demonstrably more sustainable than existing methods. Methods should be derived from the following approaches:
    - novel organic and organometallic catalysis
    - process intensification / flow chemistry
    - biocatalysis
    - synthetic biology
- The exemplification of sustainable chemistry principles into PhD and Post-Doctoral medicinal and process chemist education and training. This exemplification will hopefully lead to broader uptake by the European academic community following the lead of EFPIA consortium members.

POTENTIAL SYNERGIES WITH EXISTING CONSORTIA

The proposed Sustainable Chemistry project addresses areas that are complementary to several existing Green Chemistry efforts globally. A few examples are provided. The ACS Green Chemistry Institute Pharmaceutical Roundtable (link), consists of pharmaceutical companies interested in promoting green chemistry and green engineering in pharmaceutical manufacturing. This group has played a key role in identifying the industry's chemistry and engineering challenges, and has directly funded modest sums towards resolving these challenges. Through the ties to this group by participating EFPIA members in the IMI project, broader industry input to consortium goals can be obtained, and non-proprietary activities of the roundtable can be shared with the consortium. Through both the ACS GCIPR and the Green Chemistry Network (www.greenchemistrynetwork.org/), efforts into educational outreach have been made which can be leveraged for the contribution to Work Package 5.

EXPECTED KEY DELIVERABLES

- An evaluation of pharmaceutical synthetic methodologies which identifies synthetic approaches that pose the most significant sustainability challenges for medicine development and manufacture post 2020.
- Synthetic methodologies (derived from work package 1 and agreed between consortia and EFPIA members) which are clearly superior to the current state, as

indicated by the 12 principles of green chemistry and engineering through common 'green' metrics (e.g.,process mass intensity, E-factor, atom economy).

  - Exemplification through each of the work package areas should be delivered: organic/organometallic catalysis & flow chemistry, biocatalysis, and synthetic biology

- Recommendations for graduate training and education programs to ensure that future generation of medicinal and process chemists understand the methods and application of 'Green and Sustainable Chemistry'.

CONSORTIUM

EFPIA PARTICIPANTS

GlaxoSmithKline (co-ordinator), Sanofi-Aventis, Johnson & Johnson, Orion, Bayer.

The contribution from companies will be entirely in kind. Participants will contribute FTEs for coordination of the consortium, intellectual input and experimental work. The level of this contribution varies across companies and across work packages, but will be approximately 13 FTE/annum for 4 years across participants.
Participants will contribute a list of appropriate substrates and selected tool compounds which will be made available for chemical and biological transformation studies. Research will be conducted in parallel with the consortium, with industry partners performing synergistic and coordinated work on the same transformations being assessed by the consortium. For biological work packages, scale-up facilities for enzyme and whole cell systems will be made available to support the research. Computational and analytical assay support will also be provided to support biocatalysis and synthetic biology work. To support the educational work program, the synthesis of molecules in development will be shared, as well as 'green metrics' on these. Educational materials will be prepared and shared with consortium members. Participants would host post doctoral workers and students in industrial laboratories to provide access to technology and equipment where required.

The indicative duration of the project
The indicative duration of the project is 4 years.

The indicative budget for the project
Indicative total in kind contribution from the EFPIA companies is EUR 9.700.000 million.

Applicant Consortium
(to be selected on the basis of the submitted expression of interest)

The Consortium should include the necessary expertise to:

- Study chemical synthesis and processes through conventional catalysis, biocatalysis, synthetic biological approaches and technological methodologies as flow chemistry, with the output being the discovery of new methodologies
- Calculate green metrics for studied synthetic approaches to demonstrate their improvement over conventional methods
- Incorporate sustainable chemistry concepts into educational and training programs for medicinal and process chemists

SUGGESTED ARCHITECTURE OF THE FULL PROJECT PROPOSAL

The Applicant Consortium is expected to address all the research objectives and make key contribution on the defined deliverables in synergy with the EFPIA consortium.

The below architecture for the full proposal is a suggestion, different innovative project designs are welcome, if properly justified.

Work Package 1- Problem Identification.

This work package involves a detailed understanding of the transformations and synthetic methodologies of concern beyond 2020. The consortium researchers will provide an update to the 2007 paper published by the Green Chemistry Institute Pharmaceutical Roundtable,[i] considering also the potential scarcity of key materials and the possibility to have alternatives or new ways of producing them. This work package should precede the others as it will set the priority methodologies for industry/consortium focus.

The researchers will also look at the key reactions from a future sustainability and cost perspective and understand the following:

- transformations which are highly dependent on scarce metals as catalysts
- transformations which currently have low Process Mass Intensity, E-factors, low atom efficiencies, or otherwise have significant deficiencies relative to the 12 principles of green chemistry and engineering
- transformations which could benefit from the use of more benign solvents or processing conditions
- synthetic targets which have the potential to be produced using synthetic biology methods
- opportunities for flow chemistry to access transformations that are efficient but currently not practiced by Pharma

The output will be

- A selection of reaction types and synthetic methodologies which should be further studied in work packages 2-4. Consortia will work with EFPIA members to complete the evaluation and to obtain agreement on areas for WP 2-4 focus. For each transformation or set of transformations, an assignment should be made to the appropriate work package.
- An up-to-date evaluation of what synthetic biology has already achieved for replacement of plant extraction or semi-synthesis for Active Pharmaceutical Ingredients (as well as intermediates and drug substances), including an identification of the limiting factors for the development of synthetic biology to additional products.
- It is expected that the majority of work associated with this work package will be completed in the first year of the project and set the priorities for the remaining work packages.

EFPIA Contribution:

- Input and data from industrial medicinal and process chemists on their current synthetic challenges

- Definition of industry used green metrics and criteria for the evaluation of synthetic methods
- With consortium members, a mapping of synthetic methodology approaches to WP 2-4 (e.g., amidation might be best studied by conventional catalysis, but hydroxylation may target a biocatalytic approach).
- Synthetic biology approaches for small molecules (intermediates or APIs) and challenges of EFPIA member companies
- EFPIA partners will provide resource to assist in review and prioritization process

Work Package 2 - Green catalysis: Organo/Organometallic Catalysis and Flow Chemistry/Intensification

The ideal of synthetic sequences with minimal number of reaction steps can result from combining appropriate catalytic processes with more atom efficient use of starting materials.[ii] A number of transformations have been identified as bottlenecks in current synthetic methodology:

1. Catalytic amide formation to improve atom economy
2. Aliphatic and aromatic selective C-H activation by catalysis
   a) C-C, C-O, C-N formation
3. Asymmetric synthesis for quaternary stereo centers
   b) quaternary carbon center
   c) tertiary alcohols
   d) amination of disubstituted carbonyl compounds
4. Chiral amine synthesis (N-centered chemistry)
5. Alcohol activation for nucleophilic substitution
6. Green fluorination (selectivity/reagents)

The goal is to develop new greener catalytic methods enabling synthesis of analogues and/or complex molecules with high atom economy. Such processes will be based around organo/organometallic catalysts or combinations thereof and be safe to operate and robust towards scale up. Kinetic understanding is desirable to facilitate development and optimisation of target processes. Minimisation of catalyst consumption should be considered and appropriate measures pursued such as recovery/reuse using techniques such as supported catalysts or scavenging.

An alternative method for shortening synthetic routes can be the use of flow chemistry. The utilization of flow chemistry can enable the use of conditions and thus reactions that aren't possible in conventional batch processes. In some instances, reagents can be eliminated and thermal forces can instead be used to drive reactions to completion. In addition, some reactions can be made more sustainable in flow environments through the use of less solvent, the use of greener solvents, and improved selectivity over batch processes. This work package may deliver flow or process intensification solutions as well as catalytic solutions to address chemical transformation challenges.

The consortium will deliver new catalytic and process intensification methodologies for reactions agreed between EFPIA members and the consortium in work package 1. These will be a subset of the reactions listed above, and will be jointly worked on with EFPIA members.

EFPIA Contribution:

- Identification of useful substrates for reactions identified in work package 1

- Industrial point of contact to assist in guiding/peer reviewing research
- Flow chemistry experience at member companies and access to equipment
- In-kind R&D on selected transformation at EFPIA member's site

Work Package 3 - Cost effective, sustainable access to medicines using biocatalysis

The widespread application of large scale biocatalysis is restricted because of the limited number of commercially available synthetically useful enzymes, and the availability of suitable starting points for an enzyme catalysed reaction upon which to evolve a more specific catalyst. This work package seeks to develop new enzymes capable of accepting a broad range of substrates for selected synthetic transformations which are particularly poor from an atom economy point of view, or the use of precious natural resources.

Preferred reactions types for biotransformation investigation are:

(i) Ester to Amide. As key objective for sustainable chemistry involving two stage process of hydrolysis to the carboxylic acid followed by activation of the acid and subsequent reaction with a suitable amine, where the activation of the acid is frequently non-green (as an acyl chloride; an anhydride or imidazolyl etc.).

(ii) Redox – Reactions. Currently the most direct and sustainable processes for oxidations/reductions are not selected for lack of appropriate green reagents. As a result, processes have more steps to achieve the correct oxidation state relying heavily on toxic or expensive metals such as osmium, chromium, platinum and palladium and halogenated solvents which incur significant safety concerns.

(iii) Alcohol activation. Activation can be achieved by multiple approaches (e.g. Mitsunobu reactions; the use of sulfonyl chlorides or by substitution with halides) but these approaches are generally not considered green and can give rise to genotoxic materials. A whole cell approach would avoid these issues while allowing substitution of alcohols through activation processes such as phosphorylation or acetylation (using cell machinery to provide the necessary ATP or AcCoA).

(iv) Fluorination. Fluorination of small molecules is becoming an increasingly important green chemistry target. Approximately 45% of all new small molecule APIs contain a fluorine atom. The methodologies for introducing fluorine involve either metalation followed by treatment with an electrophilic fluorine source (e.g. Selectfluor) or reaction of alcohols or carboxylic acids with DAST, SF_4 etc. These reagents are hazardous and relatively expensive.

(v) **Hydroxylation.** Classic chemical approaches for introducing hydroxyl groups selectively into mostly aromatic molecules often use inefficient methodologies with protecting and blocking groups. Selective enzymatic (di) hydroxylation methods have the potential to be much more efficient.

The deliverables from this work package will be:

- Develop for selected reaction types enzymatic systems that allow for a fast access to the respective biotransformation comprising
    - Enzymatic systems with high promiscuity as a starting point for enzyme development
    - Determine the range of suitable substrates
    - Identification of suitable host organisms for genetic modification to allow expression and optimization of the desired enzymes while allowing utilisation of the required co-factors from within the cell.

    - Optimization of the co-factor systems where applicable (e.g. CYP450)
- Enzymatic systems (isolated enzymes and whole cell biotransformations) for desired reactions with
    - Acceptable stability of the reaction system
    - Optimized specificity to meet industrial requirements

EFPIA Contribution:

- Collection of biocatalysis successes and challenges across industry partners
- Assistance in the selection of suitable substrates for selected reactions from Work Package 1
- Industrial point of contact to assist in guiding/peer reviewing research
- In-kind R&D on selected transformation at EFPIA member's site

Work Package 4 - Synthetic Biology

In order to develop more sustainable chemistry it is possible to envisage assembling part of, or even a complete synthetic scheme using enzyme catalysed reactions in sequence. Biological organisms do this routinely as part of both primary and secondary metabolism and are able to start with simple building blocks and make molecules as diverse and complex as steroids, porphyrins, amino acids, antibiotics, pigments, etc. With the advent of modern 'High Throughput' molecular biology it is possible to modify the enzymes and regulatory elements constituting a biochemical pathway, both individually and in combination, such that multiple versions of the pathway could be created and catalogued. Such modified pathways could potentially accept alternative starting molecules (feedstocks) and/or generate different metabolic products, with the resulting molecules capable of further elaboration into diverse chemical space (a library of novel compounds) or even use as drug substances in their own right.

It is clearly undesirable to seek to engineer unique hosts for each new biosynthetic pathway and preferable to take advantage of established and emerging techniques to express individual enzymes and even entire biosynthetic pathways in alternative biological hosts. A small set of well defined and characterised, general purpose, free to use, open access organisms as expression systems would allow much faster and wider uptake of this growing area of science.

This work package would aim to prepare new biotechnological routes for the pharmaceutical and fine chemical industries.

The outputs of this work package will be:

- A small number of biological host organisms with the following characteristics:
    - Non-pathogenic, suitable and predictable for scale-up of cell cultivation and downstream processing operations using standard methods and equipment
    - Forming genetically stable recombinant systems capable of developing high biomass and/or product concentrations
    - Generic systems suitable for expression of entire pathways or individual enzymes, in the latter case including both intracellular and protein secreting systems
    - Utilising a range of low-cost, sustainable raw materials as substrates

- To demonstrate the technology, a range of mutants around a single biochemical pathway to catalyse the multistep biosynthesis of heterocycles of synthetic interest from either a single starting point or multiple starting points

    - Defined relationships between mutations in the pathway and resulting changes in substrate specificities and products and side-products formed
    - Cloned genes encoding individual enzymes and regulators for further optimisation
- Definition of the limiting factors for development of the technology :

    - Expression/ modification of complex protein structures (P450, PKS, …)
    - Location of component enzymes to different subcellular compartments
    - Co-factor biosynthesis and regeneration
    - Modeling of synthetic biochemical pathways in micro-organisms
    - Restrictions of available host systems (yeasts, fungi, bacteria, algae, other systems?)
    - Regulatory procedures and constraints to change from an extractive or chemical process to a synthetic biology process for API manufacture
    - Identification of desirable metabolic pathways specific to plants or animals

EFPIA Contribution:

- Identification of candidate organisms as starting point for generic expression systems
- Identification of suitable genetic modifications to starting organisms to give generic host organisms.
- Support for assay development and screening, computational protein structural input, molecular biology
- Support for Fermentation scale-up and provision of industry capability.
- Identification of useful substrates and products for selected synthetic biochemistry demonstration
- In-kind R&D at EFPIA member site working towards similar problems in industry
- Industrial point of contact to assist in guiding/peer reviewing research

Work Package 5 - Medicinal and Process Chemist Education

Educating and motivating medicinal and process chemists to adopt Sustainable Chemistry methodologies represents one of the key deliverables of this IMI proposal.

The vast majority of energy, chemical reagent and solvent usage occur during scale up and manufacturing of medicines rather than during the research phase. However, the synthetic routes defined early in the research phase are often adopted as the scale up route and once a route becomes embedded into a clinical program, e.g. impurity profile defined and toxicity trials performed, there are significant time and financial pressures to stay with that route. Thus the initial route choices made by the medicinal chemists can play an important role in defining how sustainable future scale up and manufacturing routes are.

However, Green and Sustainable Chemistry is not currently a major objective for medicinal and process chemists as they are focused on the complex and expensive task of designing novel drug candidates with increasingly stringent selection criteria (potency, selectivity, acceptable toxicity profile, developability, synthetic accessibility).

This work package aims to support medicinal and process chemists by demonstrating that early adoption of sustainable chemistry methodologies has the potential to replace increasingly scare reagents and processes with sustainable alternatives.

The approach that this IMI application will use is to build partnerships with leading green chemistry academic groups to achieve the following:

1. Innovative optimization of sustainable routes to new drug candidates as part of postgraduate training
2. Knowledge transfer: learning from successful drug case histories
3. Recommendations for integration of sustainable chemistry principles and case histories into post-graduate training
4. Development of advanced sustainability metric tool that can be used in early development by academia and industry alike to select most optimal chemical and technological synthetic pathways for production of specified API components

To support these deliverables, the outputs of this work package will be:

- Establish a network of innovative expert Chemistry Departments (up to three) to partner with industry to build the capability to optimize the sustainability of newly identified drug candidates (up to 6 per year). Such new drug candidates are extremely valuable assets to industry and inclusion of these into work package 5 maximizes the chances of the IMI initiative having high commercial impact.
- Deliver new and greener, more sustainable syntheses on each newly identified drug candidate to deliver an improved synthetic methodology as part of post-graduate training. Common green metrics should be used to determine program success. Each work program would run for up to 12 months, requiring up to 2 academic FTEs (Post Docs) and 0.5 industrial FTES.
- A Review and published analysis of synthetic routes used in manufacturing of current top 25 drugs. Through targeted publication and dissemination of this information, current and future medicinal chemists will be better informed as to the major drivers of environmental impact of current drug manufacturing and thus be aware of opportunities to improve sustainability through chemistry innovation.
- A Review and publication of the environmental fate of top 50 drugs (half life, environmental metabolites and routes of metabolism)
- Build and publish data comparing human and environmental "metabolism" of top 25 drugs
    - Are there opportunities to design in environmental "instability" into next generation?
    - Develop structure activity predictive tools for discriminating between human and environmental deactivation of drugs
- Development of advanced sustainability metric tool and establishment of network of academics and industry that takes continued ownership of this tool to ensure further optimization and roll-out within pharma&academic community
- Based on all of the above, the consortium should provide recommendations for the incorporation of green and sustainable chemistry principles into postgraduate training and education for medicinal and process chemistry

EFPIA Contribution:

- Disclosure of synthesis of top 25 drugs along with in-house green metric and toxicological data, and other data needed to construct databases
- Green chemistry educational material custom focused on consortium effort
- The supply of new drug candidates (up to 6 per year) which consortium will work on in parallel with industry to develop sustainable syntheses.
- Contributions to publications

Glossary

ACS GCIPR – American Chemical Society Green Chemistry Institute Pharmaceutical Roundtable

MI- Mass Intensity

PMI – Process Mass Intensity

Useful Reference

1. David J. C. Constable, Peter J. Dunn, John D. Hayler, Guy R. Humphrey, Johnnie L. Leazer, Jr., Russell J. Linderman, Kurt Lorenz, Julie Manley, Bruce A. Pearlman, Andrew Wells, Aleksey Zaksh and Tony Y. Zhang **Key green chemistry research areas—a perspective from pharmaceutical manufacturers** *Green Chem*., 2007, 9, 411–420

TECHNOLOGY AND MOLECULAR DISEASE UNDERSTANDING

6. HUMAN INDUCED PLURIPOTENT STEM (HIPS) CELLS FOR DRUG DISCOVERY AND SAFETY ASSESSMENT

BACKGROUND

Scientific advances and interest in stem cell research have developed very rapidly in the past few years with increasing impacts on drug discovery and development. This has led to a deeper understanding of the enormous potential of stem cells and their cellular derivatives in modelling human conditions in cell culture.

Embryonic stem cells are characterised by their ability to both self-renew and to differentiate to all somatic cell types that constitute an organism. The reprogramming of somatic cells to an embryonic-like stage by transfecting fibroblasts with four transcription factors enabled the derivation of induced pluripotent stem cells (iPS) starting from adult tissues. More recently, the generation of iPS cells from patients opened the field for improved *in vitro* systems for disease modelling, drug discovery, and safety assessments.

This project based on human iPS cells will thus be able to mimic patient responses and enable the development of personalised medicines by population-relevant testing of efficacy and toxicity. Patient-derived human iPS-derived cells and genotype specific patient-derived human iPS-derived cells will be unique tools to address disease biology and responses to drug treatment for diseases for which there are no animal models available. With this, the outcome of our efforts would facilitate a paradigm shift in drug discovery and safety assessment.

There are some consortia working on the use of stem cells for research, mainly based on cells from embryos (hESC) and some addressing adult-derived (iPS) cells. The uniqueness of the current proposal lies mainly on its main focus being patient-derived iPS cells, the corresponding differentiated cell types, and their application in drug discovery and safety assessment. The three main areas of focus of our research will be neurodegenerative/neurodysfunctional diseases, diabetes, and safety assessment. This project will thus be able to mimic patient responses in a selection of relevant tissue/cell types and enable the development of personalised medicines by population-relevant testing of efficacy and toxicity. Also, iPS-derived cells will be unique tools to address disease biology and responses to drug treatment for diseases for which there are no animal models available. With this, the outcome of our efforts would facilitate a paradigm shift in drug discovery and safety assessment.

Application to drug discovery and safety assessment

The current paradigm in drug discovery and development resides firmly on much better early prediction of the 'winner' in the race towards a new medicine. The improved model systems with human iPS derived cell types provide the opportunity to optimise drug development by increasing the correlation between therapeutic mode of action, safety assessment and clinical trials. The use of more relevant human cells in preclinical assays could lead to a safer drug profile and thus reduce the attrition rate and lower the costs of drug development.

Furthermore, such assays make it possible to assess compound efficacy prior to preclinical development and enable the repositioning of drug candidate molecules for novel diseases. As we are focusing on genetic-based diseases, iPS cells are the preferred option to enable stratification of patients *in vitro*. Phenotypically diverse, healthy donor iPSC-derived model cells should be employed both as reference controls and where the genetic basics of pathophysiology are still contentious. Human iPS cells will show a degree of heterogeneity, but this would also be the case with other stem cells (e.g. hESCs). The variability is not negligible, but recapitulation of genetic variants in hESC would be inefficient, time-consuming and sometimes impossible. Resident stem cells, another putative source of pluripotent cells, are not a preferred option due to issues in accessibility as their collection often requires invasive procedures.

Assays with iPS-derived human cells have the potential to reduce the risk of adverse events following drug launch, decreasing also the health risks for the volunteers and patients. It will also be possible to build a bank of iPS derived organotypic cell lines from individuals with genotypes and phenotypes known to affect drug metabolism and disposition important for drug exposure and toxicity. Moreover, and from a drug discovery perspective, patient-derived iPS cells will help to understand the genetic contribution to diseases and to unravel the underlying mechanisms. It will also be possible to stratify patients based on their iPS-derived organotypic cells as models for organ responses (including e.g. liver, cardiac, CNS, kidney, pancreas, and gut).

An important component necessary for the successful use of iPS cells in drug discovery and safety is the thorough characterisation of the iPS cell lines and of their differentiated derivatives, together with the establishment of appropriate biological and cellular assays. It is also envisioned that the application of more relevant human assay systems will lead to a reduction in the use of animals for laboratory research.

To fully capitalise on these promises, the project proposes a better structured and easier access to well characterised, physiologically relevant cells and the development of a standard package of preclinical drug safety assays available to European pharmaceutical companies and academia.

Problem Statement

Drug safety (preclinical and clinical toxicity) accounts for the failure of approximately 40% of all molecules in the development pipelines of European (many global) pharmaceutical companies, while lack of efficacy is the second leading cause of attrition for both pharma and academia. Thus, pharma and academia share a common interest to develop improved *in vitro* biological systems, in particular those relevant to man, i.e. composed of human cells.

Currently, *in vitro* compound screening relies on cell lines or primary cells derived from, at best, human origin, but mostly from animal sources or recombinant cell lines expressing the specific pharmacological target of interest. Often compound effects identified with such assays cannot be reproduced under more physiological conditions or *in vivo*. The approach is also hampered by the fact that human primary cells are available only in limited amounts and demonstrate high batch-to-batch variations.

Stem cells, and their cellular derivatives, can be produced in sufficient amounts and offer the opportunity to develop more physiological and thus more relevant assays. Moreover, a sustained source of specific human cell types, through iPS cells, will help overcome the batch-to-batch variability observed thus far. Further, the ability to phenotype differentiated cell types derived from patient specific iPS will improve the ability to link disease properties back to the physiology of defined cells. This issue is particularly acute for diseases of the central nervous system. While it is possible to obtain cells from

biopsies in diseases like diabetes, cardiovascular disease and cancer, this cannot be done for CNS diseases. As a consequence, understanding of the molecular basis of diseases like depression, bipolar disorders, and schizophrenia have lagged behind. Similarly, diabetes is a major area of interest due to the large impact on public health and the variety of organs and cell types involved in its pathophysiology (pancreas, adipose tissue, liver, etc.) and ancillary organ damage such as diabetic nephropathy. Linking the iPS cells back to the genotype of patients will aid in the exploration of the genetic link between patient and disease, and will allow phenotyping of disease relevant cell types for which the genetic makeup is known.

Besides the immediate scientific advancement, this project will ensure the sustainable application of the acquired knowledge by fostering the implementation of a centre for the long-term maintenance of the cell lines and standardised assay protocols beyond the timeframe funded.

NEED FOR PUBLIC-PRIVATE COLLABORATIVE RESEARCH

The use of iPS cells in drug discovery requires several pre-requisites which can only be tackled by a collaboration including all stakeholders. These aspects include the following.

- Establish a bio-bank for iPS cell lines including quality control assessment that fulfils pre-defined scientific requisites. This will imply strong links with hospitals, patient organisations, and pharmaceutical industry with access to human samples.
- Establish and make accessible iPS cell lines from different ethnicities and patients with defined genotypes/phenotypes affecting drug efficacy and side effects. This may include the generation of new iPS lines or the purchasing of suitable, commercially available specimens.
- Establish standardised biological assays, including differentiation protocols, using stem derived cell types addressing disease biology, response to treatment and safety assessment.
- Develop strong communicative and collaborative links with other consortia addressing similar questions, to avoid duplication of effort and enable scientific exchange to accelerate progress.

Hence, the pharmaceutical industry and academia are the main drivers of the scientific effort, but the inclusion of patient groups to address ethical questions, of SMEs to access cutting edge technology (possibly including the provision of existing iPS cell lines) in this fast developing field, of hospitals with access to clinical samples and records, and of regulators to ultimately facilitate the acceptance of biological systems and assays are pivotal for significant progress and success of this endeavour.

OVERALL OBJECTIVES

Human iPS lines and their cellular derivatives have opened up new avenues in drug discovery and development. To build on these advances, it is important to further align and support the field with the following EU-wide strategy/

- Facilitate access to a defined number of well characterised, genetically diverse iPS-derived cell types in large scale for pharma and academia in Europe.
    - Focus on relevant cell types representative of genetically determined disease phenotypes such as diabetes, neurodegenerative disorders, QT-related genotypes, and of cell types of major interest for safety

assessment in four major organs i.e. liver (e.g. hepatocytes, stellate cells, Kupffer cells), kidney (podocytes, epithelial tubular cells, fibroblasts); heart (cardiomyocytes) and gut (epithelial cells and enteroendocrine cells). Depending on outcome, other tissue types such as vasculature (smooth muscle and endothelial cells) and skeletal muscle may be also included.

- Focus selection of patient populations on diseases either wholly or partially genetically-driven, with major focus on neuro-dysfunctional disorders and diabetes. Diseases of the central nervous system include autism, Parkinson's disease, schizophrenia, depression, and Alzheimer's disease. Diabetes, Type 1 and Type 2, including the investigation of a variety of organs/cell types, including beta cells, adipocytes, sensory neurons, liver cells, entero-endocrine cells and skeletal muscle cells
- Set up a framework to
    - optimise cell-based assays predictive of preclinical or clinical toxicity for liabilities in target organs to be defined;
    - optimise assays predictive for the understanding of disease biology and of compound efficacy for a panel of diseases defined above;
    - optimise protocols for the appropriate culture, differentiation, expansion and maintenance of stem cell derived cell-types, in particular multicellular, organotypic 3-dimensional cell cultures.

Foster the implementation of a central test facility, accessible to European academia and industrial members of EFPIA, for the maintenance of the biobank beyond the funded period of this project, the maintenance of standardised protocols and the pre-competitive assessment of small molecules in iPS cell derived assays of drug toxicity and efficacy. EFPIA consortium partners and Applicant Consortium partners should profit from such a central facility with respect to training purposes and technology transfer. To ensure that ethical and legal requirements are fulfilled and include discussions with relevant stake-holders (e.g. pertinent patient organisations and ethical committees), an **Ethical Committee** will be formed to provide advice on ethical issues to the Steering Committee and Work Package leaders for decision making. The Ethical Committee shall advise the participants on compliance with European and national regulation and laws in all relevant project activities, including clinical trials and animal studies. An **IP Committee** will provide advice on IP issues to the Steering Committee and Work Package leaders for decision making. The IP Committee shall assist the Steering Committee and Work Package leaders in ensuring the application of the Intellectual Property sections of the Project Agreement. The IP committee shall assist participants in defining the most appropriate strategy for sharing and utilising existing technology IP, particularly around hIPSC.

POTENTIAL SYNERGIES WITH EXISTING CONSORTIA

Synergies with existing consortia are possible and will be sought and implemented. A dedicated Work Package will be in charge of coordinating the scientific and technological exchange with other consortia. In particular, exchange of protocols for differentiation and maintenance of cell lines, and for the scalability of protocols might be of great advantage to this and other consortia (e.g. SC4SM, ScreenTox, and other IMI projects). Also, exchange on suitable quality criteria for defined cell types is a possible asset. In addition, and in terms of appropriate biological and cellular assays to be applied, exchanges with consortia focusing on biomarkers for specific diseases or toxicities (e.g. PSTC and PredTox) could be advantageous, since the putative biomarkers might be suitable endpoints for testing specific iPS cultures, while the data obtained from such cultures may be supportive for the biomarker qualification.

EXPECTED KEY DELIVERABLES

A. Definition and selection of suitable patient populations

Definition of patient populations of interest (e.g. diabetes, neuro-dysfunctional diseases, pain, heart failure, progressive lung diseases) for the derivation of iPS lines, and definition of protocols to generate differentiated cell lines thereof to investigate disease biology and assess drug efficacy and safety. Definition of specific phenotypes/genotypes of interest, e.g. poor, rapid and ultra-rapid metabolisers based on drug metabolising enzymes known to affect drug exposure and toxicity. SMART: Within the disease defined in the research proposed in the call, specific patient sources and patient populations will be defined and recruited. Timeline: Defined by the 1st. year of the project and recruitment throughout the duration of the project.

B. Bio-bank and Central test facility

a. Establishment of a biobank of a defined number of iPS lines and stem cell-derived cell types in Europe with a standardised nomenclature. Establishment of a self-funding, validated facility, accessible to European academia and industrial members of EFPIA, for the maintenance of the biobank and the distribution of the established cell lines; this biobank may be formed by coordination of existing activities and infrastructures if available from the Applicant Consortium or other consortium initiatives. SMART: A biobank for storage of patient derived iPS cells will be in place either as an independent organisation, as part of the Applicant Consortium or as a collaboration with existing biobanks. Timeline: Defined by the 2nd year of the project, implemented by the 4th year.

b. Characterisation of iPS-derived cell types in terms of: cell donor origin (fully anonymous patient record required); phenotype (healthy, diseased, susceptibility for certain toxicity (e.g. HLA diversity), response to drug treatment, age, gender, etc.), genotype of patient of origin, original tissue source, etc. and establish a 'score card' for iPS characteristics such as epigenetic signatures, pluripotency status, gene expression profiles etc. Epigenetic changes and stability (also upon passaging, etc.) will be taken into consideration and be carefully assessed during cell characterisation. SMART: for each iPS cell banked, appropriate characterisation regarding phenotype, genotype, score card, etc will be documented and reported. Timeline: parallel to the generation of iPS cells.

c. Provision of a large scale source of 20-100 human cell lines (iPS) per disease and cell type to ensure a broad coverage of the expected phenotypic and genotypic variability among patients. Include healthy donor-derived iPS lines as reference controls and where the genetic basics of pathophysiology are still contentious. The preferred iPSC technology will be based on non-integrating systems. Available commercial cell lines or fee for service generation of iPSC may be considered as a possibility to obtain the necessary cell lines. hESC are currently the 'gold standard' and therefore will be used to some (small) extent as comparators.

d. Scaling of the cell culture protocols to prepare, from iPS cells, terminally differentiated cell types at large scale, of defined quality and reasonable costs. Compare results from different starting tissues (e.g. skin, blood) whenever appropriate. Provision of human Embryonic Stem Cell (hESC) derived cell types as positive controls. SMART: at least 20 human iPS-types will be generated and characterised during the project; each of these types will be applied to several

patients. Timelines: 10 iPS-types will be finalised by the 2nd year, and approximately 10 each subsequent year. Major effort will be placed to attain scalability or 2 to 3 different cell types (3rd year). This will be complemented by efforts to adapt validated assays (see below) to suitably available commercial HTS platforms and milieu.

e. Establishment of a self-funding, validated facility, accessible to European academia and industrial members of EFPIA for

(i) the pre-competitive assessment of small molecules in iPS cell derived assays of drug toxicity and efficacy;

(ii) training purposes and technology transfer from the central facility towards the consortium partners.

SMART: A central resource for assay performance should be established, either as an independent organisation or as part of the applicant consortium. Timeline: Defined by the 2nd year of the project, implemented by the 4th year.

C. Drug efficacy

a. Assay development for iPS-derived cell types representing a panel of diseases defined in 'A', e.g. sensory neurones for pain, and cardiomyocytes for heart failure. SMART: Preliminary panel of assays will be defined and agreed upon in the 1st year of the project. Additional assays and endpoint may be developed throughout the project period.

b. Validation of assays to measure cellular responses to target-mediated drug effects (efficacy). Emphasis is placed on high reproducibility, high throughput and low costs. SMART: Preliminary panel of assays defined in point C.a will be validated in years 3 and 4 of the project.

c. Implementation of biological assays in 384-well format for exploratory (mechanistic) investigations. Validation of these assays using panels of compounds provided by the pharma partners. SMART: Depending on assay performance defined in point C.b, miniaturisation and HTS approaches will be undertaken in year 5 (and possibly beyond).

d. Evaluation of the impact of donor phenotype of the iPS lines (e.g. old/young; healthy/diseased) on the outcome of the assays and the overall assessment. SMART: Assay outcome of different donor phenotypes will be assessed in year 5.

D. Safety assessment

a. Assay development with iPS-derived cell types to predict preclinical or clinical hepatotoxicity and cardiotoxicity. Emphasis is placed on high reproducibility, high throughput and low costs. Use of these assays to assess safety liabilities of compounds in the development pipelines of pharma and academic discovery organisations. SMART: Preliminary panel of assays will be defined and agreed upon in the 1st year and continuously updated throughout the duration of the project. In the 2nd and 3rd year, some of these assays will be validated using model compounds with known toxic liabilities.

b. Assay development with iPS-derived cell types to address specific toxicity target organs (e.g.cardiomyocytes, hepatocytes, kidney tissue, etc.) using appropriate cell types and patient. SMART: Assays will be performed in cells derived from specific patient populations in the 2nd and 3rd years. The results will be used to correlate donor phenotype with outcome.

E. Scalability and improved culture conditions

a. Evaluation of organotypic 3-dimensional cell cultures for a specified number of iPS-derived cell types to better mimic tissue responses. SMART: Improved 3D protocols in terms of scalability, longevity of the cultures, possibility of multicellular culture systems will be explored on the 5th year (and beyond).

b. Definition of most appropriate biomaterials for such culture systems. SMART: Improved biomaterials for scalability, longevity of the cultures, possibility of multicellular culture systems will be explored on the 5th year (and beyond).

c. Comparison of the performance of 3D cell cultures with conventional monolayer cell culture conditions. SMART: Comparison of performance of 'simple' and complex cell culture conditions in terms of assay performance will be explored on the 5th year (and beyond).

CONSORTIUM

EFPIA PARTICIPANTS

Roche (co-ordinator), Astra Zeneca, Boehringer-Ingelheim, Sanofi-Aventis, Johnson & Johnson, Eli Lilly, Merck Serono, Novo Nordisk, Orion, Pfizer.

The EFPIA participants will contribute:
Legal and patenting expertise. This will include the participation of legal and IP experts to facilitate discussions among the partners, with other stake holders (e.g. IP owners, patient organisations, regulatory authorities).
Scientific background on disease biology, disease targets for selected diseases. The pharma industry is in a unique position to provide excellent expertise in diseases and therapeutic interventions. This will be pivotal to define the specific lines of investigation (e.g. cell types, patient population, etc.)

Scientific background and pre-existing knowledge of toxicology issues. Due to their long experience in the drug development process, EFPIA members have access to a wealth of data that will help define the toxicological questions to be addressed. Hence, the prior knowledge of target organ and pre-clinical (animal) data will be the base for the assay development strategy in terms of toxicological evaluation.

Clinical expertise from the companies, samples from clinical trials (regarding safety and pharmacology outcome)

Expertise in stem cell biology to support and guide academic collaborators and SMEs. In case of purchasing of cell lines or available licenses, this may include payment of licenses and fees.

For the formation of the biobank/central facility, support in discussions with business experts and possibly some experienced venture capitalist to assess viability.

Experimental support for the characterisation of iPS cells and cell lines regarding their phenotype and genotype. This will include functional cellular assays/pharmacological profiling (e.g. second messenger quantification, signalling pathway assays, electrophysiology, cell growth, secretion of specific functional markers), cell marker detection for characterisation (e.g. immunocytochemistry, FACS analysis, ELISA),

molecular biology techniques (e.g. gene expression, gene sequencing, methylation/copy number variation).

Indicative duration of the project
The indicative duration of the project is 5 years. Funding beyond this period may be warranted.

Indicative budget for the project
Indicative total in kind contribution from the EFPIA companies is EUR 26 million.

Applicant Consortium
(To be selected on the basis of the submitted expression of interest)

The Applicant Consortium is expected to contribute with high level academic scientists, SMEs, patient groups, and representatives from health and regulatory authorities.
In addition, specific expertise in the following areas is necessary.

- Expertise in data analysis and database for the storage, analysis and querying of the collected data. This should include inventory capabilities, capturing of patient records, data analysis and storage of high throughput molecular biology techniques and of cellular and biological assays.
- Profound experience in cell culture systems. In particular, experience with differentiation inducing biomolecules including extracellular matrices is required. Experience with 3D organotypic bioreactors and material sciences, including the influence of matrix on the cell cultures.
- Broad experience in cellular assay development and validation. Including miniaturisation of assays, use of different assay formats, and robotics. Endpoints should include cellular assays (e.g. enzymatic activities), molecular assays (e.g. genomics), and receptor-mediated cellular effects.
- Experience in biobanking of cells and data management. Existing infrastructure of biobanking. Specific experience in data and knowledge management. Databases and queries that can be combined with existing resources (in collaborative efforts).
- Experience in quality control of cell lines, including state-of-the art genotypic and phenotypic analyses.
- Experience in the standardised production of several iPS lines and their cellular derivatives.
- Clinical expertise to access patient material (tissues) and of patient records relevant to characterise the tissues. Knowledge of disease pathophysiology and clinical outcome.
- Knowledge of specific disease- and/or toxicity-related pathways.

SUGGESTED ARCHITECTURE OF THE FULL PROJECT PROPOSAL

The Applicant Consortium is expected to address all the research objectives and make key contribution on the defined deliverables in synergy with the EFPIA consortium.

The architecture below for the full proposal is a suggestion; different innovative project designs are welcome, if properly justified.

Both applicant consortium and EFPIA participants will work in a synergistic manner. The Applicant Consortium is expected to provide expertise complimentary or supplementary (in terms of resources) to the in kind contribution provided by EFPIA participants to address the research objectives. The preliminary work package structure for the Full Project Proposal described here should serve as guidance but can be adapted or modified during the preparation of the Full Project Proposal in order to enable innovative approaches that might not been taken into account at this point in time. We propose to subdivide the activities into nine work packages as follows:

Work Package 1 - Project management

This work package will address strategy, implementation and communication within the project (across WPs) and with external stake holders such as advisory boards, IMI, other consortia, etc.

- Each WP will have a WP-leader and a co-leader, preferably one representative from EFPIA and one from the Applicant Consortium.
- The Consortium's Executive Committee will be formed by one official representative or each member, entitled to vote.
- The Consortium's Steering Committee will be composed of the WP-leaders.
- The consortium will select a Scientific Advisory Board including independent scientist and stake holders such as representatives from Regulatory Authorities, Patient Organisations or other public-private partnerships working on similar research areas.
- The consortium will establish an Ethical Committee to oversee patient selection criteria, ensure uniform approaches to eliciting patient consent and reside over full compliance with ethical frameworks presented by the individual Member State countries where clinical samples will be collected.
- The consortium will establish an IP Advisory Board, which will coordinate various aspects of the legal issues around use of technology patents in iPSC generation.
- Work packages will work independently towards the assigned goals. Progress will be monitored in regular project meetings with the Steering Committee.

EFPIA Contribution:

Project management, inclusive fees and grants for external support

Work package 2 - Sample/patient collection

This work package will enable access to samples from defined patients and healthy donors to generate iPS cells, focusing on populations defined in the proposal.

- The work package will ensure that all ethical and legal requirements for the use of human material are fulfilled.
- The work package will pursue necessary contacts with stake holders such as hospitals (if not part of the applicant consortium) and patient organizations.

- The work package will identify patient populations that are suited to fulfil the envisioned research, in close collaboration with WP6-8.
- This work package will collect and make available patient-derived data from hospitals or included clinical trials

EFPIA Contribution:

Clinical expertise for the selection of patient populations and access to samples from selected clinical trials if appropriate.

Work package 3 - Establishment/viability of biobank

This work package will be in charge of establishing a biobank, either as a nodal, de-localised structure, or by a centralised facility.

- The work package will evaluate existing possibilities, either from the applicant consortium or from other private or public organisations.
- The work package will ensure feasibility and sustainability of the selected solution.

EFPIA Contribution:

Scientific guidance, business and legal advice on feasibility and sustainability.

Work package 4 - Data management

This work package will work on appropriate data collection, querying and statistical analysis of data generated in WP2, WP6-8.

- The work package will evaluate options on data storage and database architecture that may already exist, either from the applicant consortium or from other organisations.
- Tracking of information on cell origin (supporting the biobank defined in WP3).
- Statistical data analysis and model generation assessing prediction of pharmacological or toxicological effects on defined cellular assays based on iPS and generated in WP6-8.
- Analysis of data (protein and gene expression, genotyping, functional assays) generated in WP6-8 during the characterisation of cell systems.

EFPIA Contribution:

Generation of data sets in an appropriate format and statistical evaluation of predictive models.

Work package 5 - Communication with other consortia

This work package will ensure that duplication of efforts in the research field of stem cells (iPS) will be minimised and possible exchanges pursued.

- Members of this work package will actively seek contact with identified consortia and initiatives and report back to the Steering Committee to be able to optimise synergies. They will foster exchange from scientific knowledge (through sharing protocols) or material (such as appropriate cell lines) as necessary.
- Collaborations and joint efforts with other consortia could be implemented as training, joint seminars, etc.

EFPIA Contribution:

Crucial point of contact for initiatives involving EFPIA sponsorship. Active participation in training exchanges either as hosts or guests from the institutions of the applicant consortium.

Work package 6-8 - Application of iPS for the study of diabetes (WP6), CNS and neurodysfunction (WP7) and toxicology (WP8)

These three work packages will address the core of the proposed research plan and generate the majority of the data. They may be organised as sub-packages as needed. For each of these WPs, the following activities are envisaged.

- Strong link with WP2 for the selection of appropriate patients and healthy donors for iPS.
- Acquisition through own research or through purchasing of iPS cell lines appropriate for the proposed research.
- Generation of tissue-specific mature cell populations to be used as tools for drug discovery and safety assessment.
- Characterisation of the obtained cell types based on molecular, cellular or functional markers/pharmacological profiling of endogenous proteins with reference compounds
- Evaluation of cell responses using existing or newly developed assays for efficacy and safety.
- Training and technology transfer from the central facility towards the consortium partners.

EFPIA Contribution:

Clinical expertise to guide the research. Experimental support for the characterization of cell types with assays that require complex technologies (e.g. genotyping, HCI, gene expression analysis, etc). Possibly fees and grants for purchasing of cell lines.

Work package 9 - Assay development and validation and scale-up

This WP will work in close collaboration with WP6-8 as it will develop assays to be used in the cells generated in these WPs.

- Cellular assays will be established and SOPs made available to the consortium for dissemination and implementation.
- Results from defined cellular assays will be validated (technical validation and biological qualification) in the appropriate setting.
- Validated assays will be used as readouts in WP6-8.
- When appropriately validated, assays will be scaled up and transferred to appropriated HTS platforms.
- Training and technology transfer from the central facility towards the consortium partners.

EFPIA Contribution:

Compound supply (commercially available or proprietary compounds) to serve the validation of the assays. Experimental support in automation and throughput as necessary.

7. UNDERSTANDING AND OPTIMISING BINDING KINETICS IN DRUG DISCOVERY

BACKGROUND

This proposal seeks to improve our understanding of the interactions of small molecules with protein targets. This is fundamental to molecular drug discovery and improved capability in this area has the potential to significantly impact attrition in development and therefore bring more effective drugs to patients.

This proposal focuses on the kinetics of binding but could be part of a much larger ('Think Big') proposal based on understanding the critical factors that drive molecular interactions and how they correlate to processes involved in drug action.

Understanding the kinetics of binding is of increasing importance in drug discovery.[4] Recent reports suggest that compounds with slow off-rates (candesartan – angiotensinII AT1, maraviroc – CCR5, tiotropium – muscarinic M3, montelukast – CysLT1, celecoxib – COX2) are more likely to succeed in development. However, it would seem that such properties are normally found serendipitously and not characterised until late in the compounds' development or as part of retrospective analysis of drug effects post-launch.

An analysis from Swinney[5,6] revealed that for drugs approved by the FDA between 2001 and 2004, 34% had non-equilibrium kinetics and 31% were known to induce conformational changes in proteins. Whilst the data are not available, it is almost certain that this represents a significant enrichment compared to all compounds progressed into development in the corresponding period.

Conversely, a study of cyclooxygenase inhibitors suggests that rapid off rates are a means of minimising mechanism-based side effects.[7] Hence it is reasonable to conclude that greater consideration of optimal kinetics at the time of clinical candidate selection will lead to reduced attrition during development and that it will be possible to differentiate future drugs on the basis of their kinetics. Slow off-rates are desirable in the absence of mechanism based toxicity to ensure maximum target engagement and enhanced specificity resulting in greater safety margins and reduced adverse events. Rapid off-rates are desirable where there is mechanism-based toxicity as a means of minimising these effects.

Given the potential value of optimal binding kinetics, it is perhaps surprising that there are few reports of systematic design, screening and analysis of structure-kinetic relationships (SKR) and their translation to *in vivo* outcomes. Ideally, we would like this information to be generated in a timely manner such that it can inform rational design of compounds and maximise understanding of kinetic properties that influence both efficacy and safety of future medicines.

[4] Copeland, R. A. et al. *Nature Reviews Drug Discov*. 2007, 5, 730
[5] Swinney, D. C. *Current Topics in Med. Chem.* 2006, 6, 461
[6] Swinney, D. C. *Nature Reviews Drug Discov*. 2004, 3, 801
[7] Swinney, D. C. *Lett. Drug Design Discov*. 2006, 3, 569

Problem Statement

Currently, despite the evidence for the importance of optimising binding kinetics of drug molecules, it is given inadequate consideration in discovery. There are three major reasons for this:

1. Our understanding of molecular interactions that lead to changes in kinetic profile is very limited. It is likely that conformational changes and different types of ligand-protein interaction both play a role in defining the relative energies. To our knowledge, attempts to quantify these processes have been rather narrow in scope and insufficient to allow the derivation of useful guidelines that could be used to drive drug design.

2. It is currently challenging to generate *in vitro* kinetic data with sufficient throughput and adequate timescales to inform drug discovery projects. This is particularly limiting for membrane bound proteins. In addition, there is a lack of standardised technology formats and methodologies coupled with a clear understanding of their scope and limitations based on detailed historical data-sets.

3. The impacts of *in vitro* kinetic changes on *in vivo* pharmacological effects (i.e. PKPD relationships) are not well understood. In some cases they appear to translate directly but do not in others. There is a requirement to better understand how these differences arise. In addition there is a need for validated mathematical models to predict the *in vivo* effects of altered kinetics from *in vitro* data and drug.

NEED FOR PUBLIC-PRIVATE COLLABORATIVE RESEARCH

Innovative strategies to better understand and define molecular, cellular and physiological processes that govern altered target occupancy can only be delivered through the application of a broad ranging private-public partnership such as IMI. Although this area of science has been investigated for many years, much of the expertise lies either in academic networks or within individual drug discovery projects across the private sector, many of which have been closed for strategic or technical reasons. It is only by marrying these datasets, knowledge and expertise with a strategically focused set of questions that the potential benefits of these isolated investments can be realised.

This proposal will bring together leading European academic and industrial communities from structural, biophysical, pharmacological (both *in vitro* and *in vivo*) and chemical fields. The industrial partners involved in this proposal agree to provide tool compounds, assay reagents and *in vivo* models in relation to the proposed work across a number of soluble and membrane-associated targets (to be defined as part of the formal formation of the consortium). Hence by collaborating, much more robust studies could be performed systematically that address this concept through to a pre-clinical pharmacological level whilst mitigating against the significant resources required for any one laboratory or company to prosecute on an individual basis.

From our discussions over a number of years across the pharmaceutical sector and from feedback with academic collaborators, we would anticipate there being a number of potential academic collaborators who would be interested in different aspects of this work.

OVERALL OBJECTIVES

The proposed project would consist of three parts:

- **Understanding binding kinetics at a molecular level**

We would seek to study molecular interactions in order to determine the factors responsible for changes in kinetic behaviour. For instance, we would investigate conformational re-arrangements associated with altered kinetic properties of compounds (both on- and off-rates) to establish the types of small molecules that are capable of bringing this about with a view to incorporating this knowledge into future compound design. Ideally this would be informed by X-ray crystal structures, NMR studies etc. where targets are amenable but other techniques such as mutagenesis could also be employed for membrane targets. Larger consistent data sets of targets from different protein families are necessary to derive patterns of protein-ligand interactions which give rise to different binding kinetics.

Such patterns provide the basis to develop novel molecular descriptors and/or scoring functions, which could either be used to analyse protein structures or to develop ligand-based structure-property models. In combination with the analysis of the conformational states of proteins, e.g. by molecular dynamics simulations, kinetic constants could be calculated computationally, offering new opportunities in predicting the behaviour of compounds. The objective of this would ultimately be to establish how computational or experimental approaches could be used to predictably modify kinetics in novel targets. A recent example illustrating the potential for these methods to work in synergy has been disclosed for the beta-2 adrenergic receptor.[8]

It is possible that the above effects are always target specific and that there will be no generically applicable rules. In contrast, it is reasonable to postulate that different molecular interactions (e.g. hydrogen bonding, charge, lipophilic) should have different effects on kinetic components by having different ranges of activation energies. This would be more likely to be transferable. By understanding these effects further, it is likely that we will be able to generate rules that allow the manipulation of kinetics in a more predictable fashion. A recent study of dopamine D2 antagonists shows that such analyses are indeed possible[9]. As part of this, the enthalpy / entropy signatures of compounds showing differing kinetics could be studied to examine, for instance, how protein conformational changes manifest themselves thermodynamically.

- **Assay technologies**

The widely used current assay technologies such as Biacore show great promise[10], but their applicability to membrane bound enzymes and receptors is less well established (although some results are encouraging) and clearly this limits their scope for targets of current interest. In addition there are a number of emerging technology platforms such as solution-based SPR (Pharmadiagnostics, Corning EPIC, SRU Bind, ENSPIRE, Resonance Waveguide) that are being evaluated independently by a limited number of investigators (usually pharma-based due to the considerable financial implications). We would like to assess the feasibility of developing and optimising high(er) throughput methods of determining these parameters for membrane bound receptors and enzymes. It is important that these be suitable for the generation of robust SKR in a timely manner. It is possible that observed kinetic effects *in vitro* are attributable to assay conditions and not transferred to the natural systems, hence, in order for these data to

[8] Kobilka, B. K. et al. *Nature*, 2011, 469, 7329
[9] Tresadern, G. et al. *Bioorg. Med. Chem*. 2011, 19, 2231
[10] Andersson, K. et al. *Expert Opin. Drug. Discov*. 2006, 5, 1

be meaningful, we would need to investigate these effects more generally and any higher throughput method would need to be validated against a physiologically relevant dataset. There is also a need to perform experiments across multiple cellular backgrounds in order to provide confidence that this is a target-specific phenomenon.
This proposal would include an evaluation of new and existing assay technologies using a consistent dataset to validate the methods and ultimately derive confidence that driving drug discovery efforts against any such technology is meaningful. This last aspect of the work would be carried out by the industrial partners.

- **Understanding *in vivo* translation**

In vivo efficacy and duration of action are affected by multiple parameters in addition to drug-target interactions. There are a number of pharmacological models available in order to understand target-specific binding kinetics ranging across cellular fragments, isolated enzyme preparations, whole cells, primary culture, *ex-vivo* tissue preparations and ultimately *in vivo* PD models. We would intend to study a range of compounds with different receptor kinetics across these platforms in order to better define the real benefit that such properties impart and determine appropriate considerations for scaling to man. This may include the development of appropriate in vivo models.
Of course, target engagement (i.e. whether a compound exhibits a pharmacological response) is not just a function of interaction with the target but also encompasses compound dose, systemic exposure (pharmacokinetics) and the physiological reaction of the system to the drug. Therefore robust modelling of pharmacokinetic and pharmacodynamic relationships together with *in vitro* and *in vivo* DMPK support would be vital. This input could be provided by the industrial partners.

POTENTIAL SYNERGIES WITH EXISTING CONSORTIA

To our knowledge there are no obvious synergies with existing IMI proposals, although as stated in the introduction, we envisage that this will synergise with future proposals in the field of molecular interactions.

EXPECTED KEY DELIVERABLES

In the timescales of this project, we envisage delivering:

- guidelines for understanding the molecular phenomena that allow manipulation of kinetics by small molecule design;
- technology evaluation using agreed benchmark tool compounds and molecular systems;
- improved methods and recommendations for obtaining high(er) throughput kinetic measurements;
- robust, predictive PKPD kinetic modelling paradigms that will improve understanding of how *in vitro* kinetic effects translate into *in vivo* efficacy and duration of action;
- enhanced data-sharing within the framework of drug binding and kinetics.

In the long term, it is our aspiration that this will result in reduced attrition in drug development.

CONSORTIUM

EFPIA PARTICIPANTS

Astra Zeneca (co-ordinator), Roche, Novartis, GlaxoSmithKline, Sanofi-Aventis, Bayer, Merck.
The contribution from companies will be entirely in kind. Participants will contribute FTEs for coordination of the consortium, intellectual input and experimental work. The level of this contribution is likely to be around 1 to 2 FTEs per company.

Participants will contribute a list of appropriate targets / systems of interest and selected tool compounds which will be made available for studies spanning a range of different target classes and compound structures to allow coordinated and consistent studies throughout the consortium. These would be entirely precompetitive and so would be restricted to targets no longer of commercial interest and anti-targets to enable free and full information sharing and unrestricted publication possibilities. This will include *in vivo* probe compounds.

Participants will also contribute computational support (molecular modelling & dynamics, systems biology); coordination, active participation and input; and host post doctoral workers and students in industrial laboratories to provide access to technology and assays where required.

Indicative duration of the project
The indicative duration of the project is 5 years.

Indicative budget for the project
Indicative total in kind contribution from the EFPIA companies is EUR 7.400.000 million.

Applicant Consortium
(To be selected on the basis of the submitted expression of interest)
The Applicant Consortium is expected to bring forward proposals targeted at any of the three areas outlined: molecular understanding, assay technology and link to *in vivo*.

We would envisage applications from investigators specialised in various aspects of this and the role of the consortium will be to coordinate studies and ensure knowledge is shared.

Molecular understanding

Applicants would be expected to have the capability to provide methods of studying systems which allow different components of binding events to be uncoupled to assess individual contributions to on and off rates to allow the analysis which could lead to general considerations for optimisation of these parameters. We would also envisage that there will be expertise that allows the prediction of protein conformational changes which may be tested with the proposed systems to assess their future predictive value.

Assay technology

Current methods for assessing binding kinetics are promising but also have their limitations. As part of this consortium we will conduct a robust assessment and comparison of the various methods. Applicants would be expected to bring forward proposals for developing new methodologies that address the identified short comings. We would particularly encourage the development of methods that are applicable to membrane proteins.

Link to *in vivo*

We would anticipate applicant proposals capitalising on expertise in the fields of pharmacology that allow the study of binding events in open systems such as cells, isolated tissues and ultimately *in vivo* models such that the additional phenomena governing receptor occupancy in these more complex systems can be understood. Particular emphasis would be placed on comparing systems which translate well from isolated enzyme and those that do not.

These expectations are not intended to be prescriptive. It is understood that this field offers great potential for innovation and for additional expertise. Hence additional creative proposals targeted at the broad area would also be welcomed.
The IMI JU financial support to this project will approximately match the cash equivalent of the EFPIA companies' in-kind contribution.

SUGGESTED ARCHITECTURE OF THE FULL PROJECT PROPOSAL

The Applicant Consortium is expected to address all the research objectives and make a key contribution to the defined deliverables in synergy with the EFPIA consortium.

The below architecture for the full proposal is a suggestion; different innovative project designs are welcome, if properly justified.
It is anticipated that this will be achieved through collaboration and individual applicant need not address the deliverables in their entirety.

Work Package1 – Molecular understanding

Applicants would use proteins and ligands provided by the EFPIA consortium to study binding events in the defined systems. Studies would be coordinated between individual applicants to allow cross comparison of data.

EFPIA Contribution:

The EFPIA participants will provide protein and ligands along with corresponding kinetic data to enable design of subsequent studies. In addition, additional data and access to equipment would be provided as required.

Work Package 2 - Assay technology

Applicants would develop novel methods for measurement of kinetic data addressing the shortfalls with existing methods. These new methods would be benchmarked against existing methods using the compounds provided to allow a fair and robust assessment of their value.

EFPIA Contribution:

The EFPIA participants will provide protein and ligands along with corresponding kinetic data for benchmarking. Additional systems for study could also be provided where required.

Work Package 3 – Link to *in vivo*

Applicants would use tool compounds provided to study the translation of their effects from isolated protein systems into cellular and more complex environments more relevant to the *in vivo* situation. Emphasis would be given to the study of systems which do not appear to translate in order to establish the reasons behind this.

EFPIA Contribution:

The EFPIA participants will provide protein and ligands along with corresponding kinetic data and *in vivo* data where available. Access to *in vivo* models for generation of further data and PK/PD modelling expertise could also be provided where required.

---

[i] Constable, D.J.C., Dunn, P.J., Hayler, J.D., Humphrey, G.R., Leazer, Jr., J.L., Linderman, R.J., Lorenz, K., Manley, J., Pearlman, B.A., Wells, A., Zaksh, A. and Zhang, T.Y. *Green Chem.*, **2007**, *9*, 411–420

[ii] Trost, B.M. *Science* **1991**, *254*, 1471-1477