1
Administrative Optimization of Proteomics Networks for Drug Development

André van Hall and Michael Hamacher

Abstract

Administrative structures are gaining more and more importance in the complex world of modern science. This article will define the terms administration and networking, describing the aims and tasks of project management. The analysis of neurodegenerative diseases with proteomics technologies will be looked at from the administrative point of view with a focus on the different phases of strategy development, human resources, project control and networking. The realization of these tasks is illustrated by short presentations of a national funded network, the German Human Brain Proteome Project (HBPP) within the National Genome Research Network (NGFN), as well as of the international Brain Proteome Project of the Human Proteome Organisation (HUPO BPP).

1.1
Introduction

In modern science, the importance of administration has increased steadily over the last few decades. Nevertheless, administrative work and its influence on the success of projects as well as on financial aspects (e.g., refunding) are still undervalued in the academic field. While industry recognized the importance of organizational aspects long ago, positions responsible for administrative tasks within scientific research groups (excluding administrative departments of the universities themselves) are rare. The number of operative relative to administrative personnel is still much higher in academia than in companies (at least in Europe). As a consequence, these tasks are often done by the coordinator of a given project or one of his coworkers, who are often overloaded with work, sometimes unmotivated and mostly untrained in this field. A picture of the typical administrative research scientist as being exhausted by research, teaching and organization is emerging. In addition, staff turnover in these positions is often high, resulting in loss of knowledge, lack of continuity, and commonly, a lack of

Proteomics in Drug Research
Edited by M. Hamacher, K. Marcus, K. Stühler, A. van Hall, B. Warscheid, H. E. Meyer

ISBN: 3-527-31226-9

perception as to where responsibility lies. At the same time, such positions could be extremely important for the overall success of the group, e.g., in the crosslinking of basic research and commercialization.

Owing to the increasing complexity of modern science, e.g., international networking and large consortia, and the urgent need to present scientific research to the public and to the governmental project management/advisory board, a department-spanning administration should be implemented. Most research efforts in the health sciences are extremely complex and are difficult to explain to nonscientists, which often leads to misunderstanding, antipathy or even hostility from the public (e.g., see stem cell discussion, gene technologies, etc.). As the last 20 years have clearly shown, the support of a common administrative staff leads to the scientific personnel being relieved of additional work to which they are not suited, to an optimization of the scientific output (increasing added value) and to a broader acceptance in society. The need for management expansion has also been recognized by the European Union and its advisory councils, as expressed by Ernst-Ludwig Winnacker, the president of the European Heads of Research Council, in an interview with *The Scientist*: "The networks of excellence are big enterprises that require a great deal of management, and these have not been appreciated by scientists as much as the smaller, short-term programs that are less complex to manage and that facilitate work with smaller partners." (*The Scientist* online, 25 August 2004: http://www.biomedcentral.com/). The reasons for this development will be shown in the next paragraphs.

1.2 Tasks and Aims of Administration

The following chapter will present a short overview about modern scientific administration, mainly focusing on the academic side of research. To set a common starting point of what "administration" is about, the following definition is used:

- The act or process of administering, especially the management of a government or large institution.
- The group of people who manage or direct an institution.

Simultaneously with the increasing complexity of life science, the tasks and aims of the administration have steadily grown and evolved to a much more active management role. Originally mainly involving finances and human resources, these tasks have been joined by numerous other duties and responsibilities. Many projects demand large groups or consortia resulting in network systems (see below), thus making the organization and the feedback of teamwork as well as facilitation of the flow of information within a network an essential part of work. Additionally, interactions between the network and other national as well as international research projects, research institutions and private enterprises have to be handled. This includes so-called lobby work, the discussion with and

convincing of policy makers e.g., within the European Union, to support the kind of research one advocates as the most promising approach.

Further tasks required of an administration are the composition of progress reports/business plans and final reports on schedule, the organization and calling of coordination meetings, the coordination and active participation in public relations (conferences, seminars, TV, radio, journals, etc.). This includes the planning and realization of training courses concerning technologies and topics provided by the consortium members, and the publication of the subproject results obtained at the respective time points. Moreover, (existing) homepages should be improved and optimized steadily, so that they serve not only as an information platform, but also as an interchange and communication portal.

Taken together, the administration has to

- build up a network offering fast and efficient information flow;
- elaborate business plans, evaluate the progress of subprojects and co-ordinate efforts;
- implement infrastructures (see evaluation, Section 1.4);
- serve as a central contact and administration point (added value);
- increase public knowledge and acceptance of proteomics;
- implement a bioinformatics infrastructure that will serve as a basis for further data base projects.

The aims of the administration – particularly in universities – are obviously to optimize processes and workflows within the respective department or network. Though implementation of controlling and monitoring could be hard to adopt in academia (in regard to the strong group autonomy), both processes are inevitably mandatory, especially in times of decreasing budgets and funding, as a consequence of which some US universities have started to gather discarded or not-required high-tech equipment from local departments and offer it to all other groups for free, avoiding unnecessary investments and expenses.

There are several other domains that have to be carefully considered when aiming at successful projects, most notably in human resources, where the generation of job specifications and the consequent identification of adequate coworkers should not be underestimated. Qualified and motivated employees who fit into the group structure are the basic requirement for planning, performing and finishing work packages in a defined schedule. These have to be generated carefully and in regard to several questions, e.g., medical need, potential return of investment, proof of concept and commercialization.

Commercialization was more or less been ignored in academia until the 1980s, when more and more scientists came to the opinion that research and marketing do not necessarily exclude each other. Several processes around the world now show the increased importance of marketing. No application within the EU can be submitted to obtain grants without presenting utilization strategies. Scotland started a Proof of Concept Fund in 1999 to advance promising ideas from university to readiness for marketing (www.scottish-enterprise.com/proofofconceptfund).

More than 140 projects have already been funded with €36 million, resulting in six existing and ten planned spin-offs. In Germany, universities and research organizations, e.g., the Fraunhofer Gesellschaft have implemented utilization departments specializing in regard to patents, licensing, consortia contracts, etc. The Ruhr University in Bochum, Germany, for example founded the Research and Collecting Society "RUBITEC – Society for Innovation and Technology" in March 1998 (http://www.ruhr-uni-bochum.de/rubitec/start.htm), consulting the numerous groups at the campus. The National Institutes of Health has elaborated a complex organization structure including the Office of Technology Transfer (http://ott.od.nih.gov) dealing with 341 invention disclosures and US$ 53.7 million in royalties in 2004. These centers offer competent help in realizing products and patents, but leave the initial efforts to the research groups. Scientists have to inform themselves about possible strategies and have to evaluate the putative success. An administrative coworker assuming this time-consuming job will function as a bridge between the groups and the central transfer departments. Thus, taken together, the optimal realization of these tasks will lead to the relief of the operative coworker, enabling the researcher to concentrate on the actual scientific work, to shorten the time from idea to output, and to commercialize his output successfully.

As already mentioned above, research efforts are more and more bundled in consortia and networks. Owing to the importance of this circumstance, it is necessary to discuss some theoretical aspects of networking and the consequences resulting from its nature.

1.3 Networking

Networks are an organizational structure with at least two independent entities being in a repetitive, long-lasting exchange/interaction status (see also Burt, 1980). Owing to the independency of the entities the network is more or less bound together by social relations, according to one or more motivations:

- Necessity: interaction is initiated by law or regulatory prescription.
- Asymmetry: to gain influence and control over the partner/its resources.
- Reciprocity: to achieve bilateral aims and interests.
- Efficiency: to gain higher input/output-ratio by utilize synergistic effects.
- Stability: to reduce/absorb/predict uncertainties.
- Legitimization: to gain or improve reputation, image or prestige.

The process of composing and inspiring a network can be divided into seven phases:

- Self-analysis: what is the goal?
- Specification: which resources are missing?
- Preselection: who offers the lacking resources?

- Partner analysis: does the new partner fit in the overall concept?
- Definition of goals: what do the partners expect from each other?
- Process modeling: how can the goals be reached?
- Realization.

Industry is again on the cutting edge in establishing strategic alliances or regional clusters. In Switzerland more than 80% of all biotechnology companies are concentrated in the four regions Basel (Biovalley), Zurich (MedNet), Lake Geneva (bioalps) and Tessin (biopolo) (Veraguth, 2004), profiting from the "big pharma" industry that offers potential financiers, manpower and licensees.

In academia, the factors asymmetry, reciprocity, legitimization and efficiency probably have to be considered as the main motivation for building up networks. Nevertheless, most cooperative enterprises follow from personal relationships or historically derived projects that have been performed in the group several years ago. The need for combining synergistic resources is often unseen, sometimes hampered by ignorance of which potential partners are working in the same field or could offer complementary techniques. The identification of key players and potent partners is therefore an essential task in organizing a powerful network.

In addition to this selection mode, the management has to deal with regulation between the partners as well as between the consortium and external entities, with allocation concerning the access of given resources and with evaluation in regard to the output (profit, innovation, proof of concept). Problems within networks often evolve from the opportunistic behavior of one or more partners or due to different strategic targets, thus demanding complex agreements and interaction/communication right from the beginning to generate confidence between the partners. Throughout the whole project, several quality control steps concerning work packages, finances, etc., have to be performed.

1.4 Evaluation of Biomarkers

In general, the struggle for understanding and fighting e.g., neurodegenerative diseases, is intended to find either drug targets involved in the pathological processes or diagnostic markers that allow sensitive identification of disease stages (Zolg and Langen, 2004). Diagnostic markers can be subdivided into:

- Screening markers: allow indication of the transit from health to disease [e.g., maternal serum invasive trophoblast antigen for Down syndrome during the second trimester (Palomaki et al., 2004)].
- Prognostic markers: allow prediction of the disease process [e.g., survivin expression in pancreatic cancer patients (Kami et al., 2004)].
- Stratification markers: allow prediction of the response to a medication strategy [NQO1 genotype in adenocarcinoma of upper gastrointestinal tract (Sarbia et al., 2003)].

- Efficacy markers: allow monitoring of the efficacy of a given drug treatment [serum CYFRA 21-1 (cytokeratin-19 fragments) in breast cancer (Nakata et al., 2004)].

Before starting research, several questions have to be answered in a detailed business plan when aiming at a successful utilization concept in industry (Zolg and Langen, 2004), e.g.,

- Do competitive markers already exist on the market?
- Will the marker be easily accepted in the market?
- Will the marker cover/exceed the research costs?

As academic research usually is much more philanthropic than industrial, these considerations are normally secondary for scientists in universities. Nevertheless, it is highly advisable to elaborate a business-plan-like approach concept dealing with pros and cons, work packages and possible contingency plans to increase efficacy and output.

The interconnection between the basic research and commercialization is structured most efficiently in an innovation process organized with clear stage gate decisions (see Chapter 17). An estimated 50% of all life science companies are using this structure. Here, product ideas originating from the research will be judged by a decision board in regard to economically relevant features (e.g., market need, competition, etc.). People and know-how will be transferred in several stages to the commercialization branch. This milestone-oriented process will be reviewed constantly by a board. After passing all criteria including concept, market attractiveness, competitive market position, competitive technology position, reward, and risk, the project will go into the next stage of the innovation process with clear planning for milestones and budgets. This phased project planning was developed by NASA in the early period of crewed spaceflights and propagated by product development experts such as R. G. Cooper (Cooper, 2001). Work is divided into sequential phases avoiding overlapping activities, but as every gate has to be carefully evaluated, it is inherent to the process that there will be a relatively long time from the idea to the market.

To bypass this, the so called bounding box approach (management by exceptions) can be implemented: prior to the beginning of a project, all internal and external factors are surveyed (budget, profit margin, schedule, etc.) and boundaries are fixed in which the project is regarded as on-track. If these boundaries are crossed, a decision board has to reevaluate the work. As the team is free to move within the boundaries, time-consuming evaluation processes are minimized, as is time to market. Alternatively, the well known project risk management can be chosen. Risk management is a process of thinking systematically about all possible undesirable outcomes before they happen and setting up procedures that will avoid them, minimize their impact, or cope with their impact. Thus, risk assessment and risk control are two important concepts that have to be kept in mind.

1.5
A Network for Proteomics in Drug Development

The concepts described so far, *administration*, *networking* and *bio markers* represent fundamental cornerstones for considering how to establish a scientific program for drug development within the field of proteome analysis.

The identification of bio markers by proteomics and proteomics-associated technologies is the key approach for drug development on the protein level. It is obvious that a higher number of identified proteins will increase the chances of finding relevant markers regarding a specific pathogenic question. After validation these marker proteins can then be used as starting-point for a drug development process.

To meet this challenge, it is necessary to combine a wide range of technologies including "classical" proteomics, e.g., 2D-PAGE and mass spectrometry, new proteomics approaches like multidimensional chromatography, and technologies for transcriptional analysis. As very few institutes provide all these applicable approaches, the reasonable procedure is to combine groups with outstanding expertise in the different fields to form a network of excellence.

In Figure 1.1, a possible structure for such a consortium is shown. First, general considerations lead to a hypothesis which comprises an approach for understanding pathogenic mechanisms of a specific disease. Based on this hypothesis, the appropriate tissue as well as the suitable model organism must be defined, and providers of the relevant samples found. For networking reasons, integration

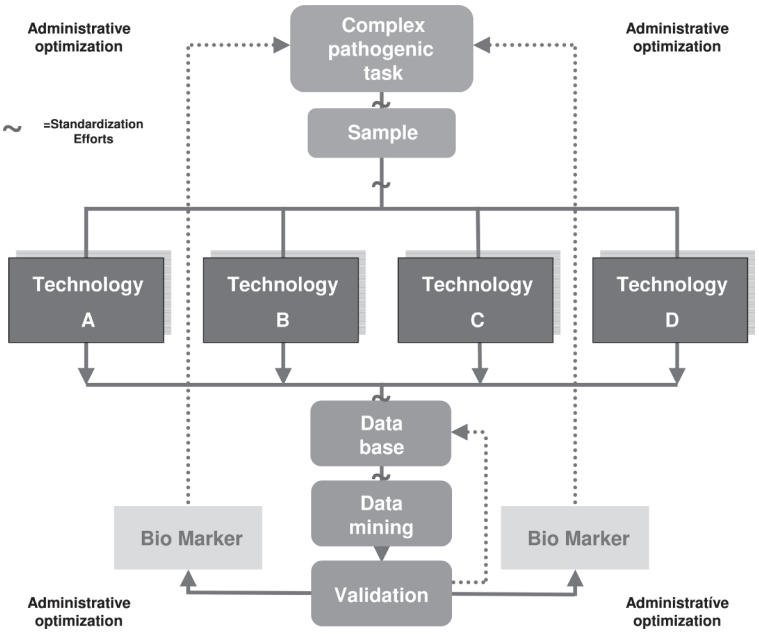

Figure 1.1 Structured workflow within a consortium for disease-oriented proteome analysis.

of the tissue providers into the consortium is recommended. The standardized samples will then be distributed to the single-technology partners within the network. As already mentioned, the range of this established technology portfolio is crucial for the possible impact of the entire consortium on the pathogenic relevance. Thereafter, the generated data will be incorporated into the project data base and reanalyzed by data mining experts (see Chapter 2). Using this process, proteins will be identified as bio markers to provide potential drug targets. Another crucial step is to validate the candidate proteins. Here, technologies for analyzing protein functions and protein–protein interactions are the instrument of choice. The comparison with the initial hypothesis will then hopefully lead to a feasible clinical approach for drug development.

Within the entire workflow described above the standardization of sample preparation [implementation of standard operating procedures (SOPs)], analyzing procedures and data handling will assure the comparability of results within the network as well as with results outside the consortium. Thus, the standardization is essential for efficient networking.

In establishing a scientific network, it is indispensable to bring together experts in the required fields. Recapitulated, partners for the following tasks have to be identified: tissue provision, technology-based analysis, pathology, data management, data reanalysis, and validation. Owing to the complexity of such an accumulation of heterogeneous partners that are also locally separated, implementation of a goal-oriented coordination is necessary. In the following, the realization of the described network structure will be illustrated by presenting German initiatives in proteomics networking in both the national and international environments.

1.6 Realization of Administrative Networking: the Brain Proteome Projects

The need for large international collaborations is obvious when analyzing the human proteome. The reasons are manifold, e.g., the low abundance of the majority of most cellular proteins (10% of all genes probably encode for 90% of all prevalent proteins), the absence of suitable high-throughput techniques for increasing sensitivity [polymerase chain reaction (PCR) equivalent for proteins] as well as the enormous number of protein species as the consequence of differential splicing, posttranslational modifications, etc. (Humphery-Smith, 2004). In addition, most diseases might not be monogenetic, but may be caused by multiple genes, modifier genes, the genetic background, etc. As a consequence, the most promising and synergistic approach is the analysis of the protein complements via transcriptome, proteome and toponome profiling.

One of the most striking tasks to start with is standardization (Meyer et al., 2003; see Chapter 2). Although it may not be feasible to elaborate fixed SOPs for all imaginable setups and questions, the key parameters of each experiment have to be annotated at least, so that possible differences can be traced back to variable

steps in the chain of work (Hamacher and Meyer, 2005a). Some elements of standardization cannot be realized employing human material. Each of us is supplied with a diverse set of genes (polymorphisms) and has undergone a different history within his lifespan, entailing varying proteomes. This might be solved via studying numerous human samples and statistical methods. In general, single groups or technical approaches are not sufficient to overcome the complexity of this challenge or to describe a given (disease) status properly. Instead, the simultaneous efforts of numerous, but standardized working groups are essential for this huge challenge and to develop a knowledge base of the normal human proteome (Hanash, 2004a,b). Two mainly academic examples will demonstrate the attempt to understand and to ease/cure the malfunction of the diseased brain, namely a national funded consortium as well as an international, voluntarily driven project (Klose et al., 2004).

1.6.1
National Genome Research Network: the Human Brain Proteome Project

In 2001, the German Federal Ministry of Education and Research (BMBF) initiated the National Genome Research Network (NGFN) as a nation-wide multidisciplinary platform network aiming at the analysis of common human diseases, as well as aging. Within the NGFN the so-called Human Brain Proteome Project (HBPP) focuses on the analysis of the human brain in health and disease. The concept is based on three consecutive steps:

- Elaborating and establishing the necessary technology platforms: HBPP1 (2001–2004).
- Proteome analysis of Alzheimer's and Parkinson's diseases: HBPP2 (2004–2007).
- Validation of target proteins and analysis of disease mechanisms: HBPP3 (planned for 2007–2010).

The HBPP1 has been funded for a period of three years with approximately. €10.5 million (2001–2004). In this project 12 partners formed a strategic consortium, consisting of nine academic groups and two companies (Marcus et al., 2003, 2004) The main focus was on the improvement of proteomics-related technologies on the basis of brain analysis. One aim of the consortium was the characterization of the human and mouse brain proteomes in regard to the identification of proteins, generating mRNA profiles, studying protein/protein interactions and validating possible targets. Data gained was used to compare mouse models and relevant human tissues for neurodegenerative diseases. To achieve these aims, the essential technological methods had to be improved and new technologies identified.

The interest of the consortium in developing and testing new tools for proteome analysis was directed to solutions for particular technical problems concerning sample preparation, the 2D PAGE system, protein quantification, and the development of UniClone sets (nonredundant cDNA expression library) from the adult human brain to be used for creating clinically relevant biochips. These

techniques were intended to be combined to develop a fully integrated Proteomic Workstation in which samples are prepared and processed automatically, e.g., by establishing 2D/3D biochips on which the samples are immobilized for further analyses. To overcome the large number of data sets that were generated by the different groups, the bioinformatics activities were expanded, e.g., a build-up of the project data base in which all data files provided by the project partners will be stored. Owing to the annotated information, for instance the link variation in protein expression to particular genes, hopefully it will be possible to elucidate the regulatory network acting between the genome and the proteome. This will create new insights for drug development concerning neurodegenerative diseases like chorea Huntington's, Parkinson's and Alzheimer's diseases or multiple sclerosis, also resulting in marketable technology products in these fields.

In the second funding period of NGFN the Systematic Methodical Platform (SMP) Human Brain Proteome Project 2 continues the work in a new formation of nine academic partners and one company. The aim of HBPP2 is to optimize developed technology and gain knowledge that, once applied, enables the development of new strategies for the diagnosis and treatment of neurodegenerative diseases. To achieve this goal, HBPP2 has gathered a critical mass of interdisciplinary German research groups with extensive experience, an unprecedented research infrastructure, a global science network within the Human Proteome Organisation (HUPO) and a solid record in clinical and preclinical work encompassing human genetics, cell biology, animal models, molecular biology, and biochemistry, thus encompassing the integration of large-scale functional genomic and proteomics approaches.

In this second funding period of the HBPP the three main goals are:

- The advancement of already established technologies: based on advances in technology achieved within NGFN1, HBPP2 will further advance its technology platform for the planned scientific program. Proteomics technologies (large 2D-PAGE, multidimensional chromatography, mass spectrometry), toponomics and functional assays such as cellular overexpression, pharmacological inhibition, RNAi and optical methods such as green fluorescent protein (GFP)-labeling, immunofluorescence and fluorescence resonance energy transfer FRET/bioluminescence resonance energy transfer (BRET) will be employed to analyze the functional implications of gene mutations selected in collaborations with the clinical partners.
- Investigating neurodegenerative diseases: HBPP2 will emphasize on applications of genomic and proteomics technology. A focus will be systematic analysis of proteins in human/mouse brain and nervous-system-related proteins in bodily fluids under normal and pathological conditions. Alzheimer's and Parkinson's diseases will be studied on the basis of human material and selected mouse models. The technology portfolio provided by the HBPP2 offers a conceptually novel opportunity to understand disease mechanisms in that it attempts to progress from current reductionist approaches to an integrated understanding of biological systems.

- Networking within NGFN-2: tight collaborations with clinical groups will allow the performance of clinically relevant proteomics studies or protein analyses that offer the most advanced proteomics technologies to the NGFN. In addition, HBPP2 is open to cooperation with other systematic-methodical platforms, namely bioinformatics, RNAi, and mammalian models. Data obtained in the project will be collected in a new type of database. Standards, SOPs and software for data management and integration will be developed. Together, these tools will form the basis for an efficient analysis and the generation of knowledge on the fundamental biological processes in normal and disease-affected brain.

Taking the current phase 2 of the HBPP as an example, the workflow within the consortium is shown in Figure 1.2. Starting with considerations about pathogenesis of neurodegenerative diseases the consortium will be provided with mouse, ape and human brain tissue. The data sets derived from the different available technologies will be incorporated into the project data base which will be presumably linked to the Data Collection Center (DCC) of the HUPO Brain Proteome Project (BPP).

After reanalyzing the data using customized software tools the identified proteins will be validated by partners within the network.

The coordination structure already established in the first funding period deals with administrative issues on different levels: Firstly, the activities within the HBPP are managed by the coordination team. Furthermore, the crossbridging to the

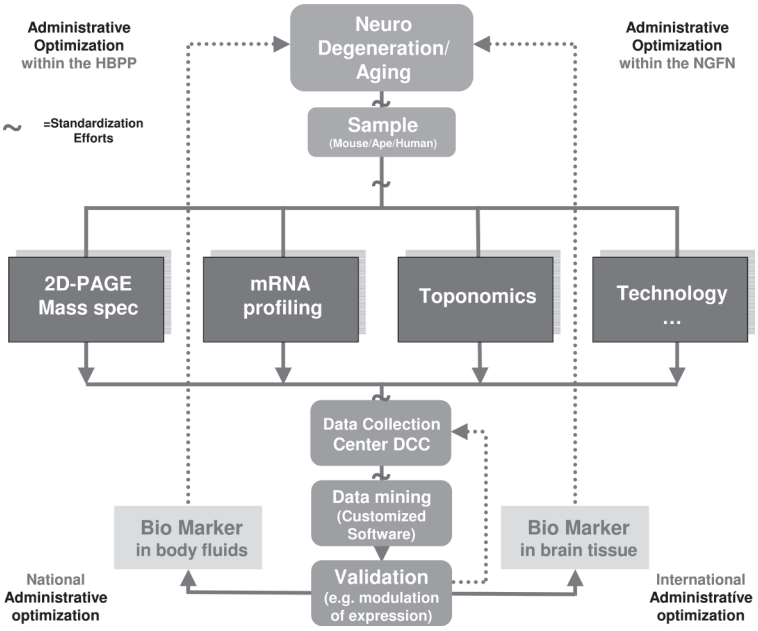

Figure 1.2 Adapted workflow of the Human Brain Proteome Project (HBPP) consortium within the second funding period (compare with Figure 1.1).

NGFN is also part of the administrative task force. In addition, most of the connections to national and international partners are coordinated centrally.

In the next step of the Human Brain Proteome Project it is planned that identified disease-associated proteins will be validated using several different techniques. The pathways they play a role in will be analyzed. This will lead to help in understanding the analyzed diseases and to develop diagnostic and/or therapeutic approaches.

1.6.2 Human Proteome Organisation: the Brain Proteome Project

At about the same time as the German HBPP was founded in 2001, the international HUPO was established as a part of the Human Genome Organisation (HUGO) (Hanash, 2004a; www.hupo.org). HUPO is a nonprofit organization promoting proteomics research and proteome analysis of human tissues. Several initiatives have been established under the roof of HUPO that analyze the proteome of a distinct human organ, e.g., the Plasma Proteome Project (PPP), the Liver Proteome Project (LPP) (Hanash, 2004b), and the Brain Proteome Project (BPP) (Meyer et al., 2003; www.hbpp.org). The HUPO Proteomics Standards Initiative (HUPO PSI) aims to establish definitive bioinformatics standards (Hermjakob et al., 2004a) and is therefore an overlapping project chaired by Rolf Apweiler from the European Bioinformatics Institute (EBI, Hinxton, UK). Standards include mass spectrometry (mzData, mzIdent), protein–protein interaction (IntAct) (Hermjakob et al., 2004b) and General Proteomics Standards (GPS), e.g., minimum information about a proteomics experiment (MIAPE) (Orchard et al., 2004). More information about this modular system is available at http://psidev.sourceforge.net.

The HUPO initiative concentrating on the brain is the HUPO BPP. After the 1st HUPO World Congress in Versailles, it was started by Helmut E. Meyer, Bochum, and Joachim Klose, Berlin, both in Germany in 2003. At a kick-off meeting in Frankfurt, Germany, in late April 2003, the first interested colleagues from around the world met to discuss the shape of the project. Since then, numerous meetings and discussions have taken place, often in close collaboration with the HUPO PSI and the EBI (e.g., Stephan et al., 2005; Hamacher et al., 2004). At the 2nd HUPO BPP workshop at the ESPCI in Paris, April 2004, attendees expressed the HUPO BPP vision as "Towards an understanding of the pathological processes of the brain proteome in neurodegenerative diseases and aging". The postulated vision of the HUPO BPP is the understanding of the pathological processes of the brain proteome in neurodegenerative diseases and aging. This will be achieved by deciphering the normal brain proteome, by correlating the expression pattern of brain proteins and mRNA and by the identification of disease-related proteins involved in neurodegenerative diseases. A pilot phase began in 2004 that addresses a quantitative proteome analysis of mouse brain of three different ages (all samples obtained and prepared by one source) and a differential quantitative proteome analysis of biopsy and autopsy human brain samples (Hamacher et al., 2004).

1.6.2.1 The Pilot Phase

Several conditions have to be met before the main project can be commenced, e.g., a broad community and reliable infrastructure. Without question, a detailed phenotyping of mouse models/patients, a complete characterization of tissue samples before proteome analysis and a high degree of standardization are extremely important in obtaining reliable results. Thus, in the HUPO BPP two pilot studies were initiated, limited to December 2004 (practical work) and March 2005 (data submission), respectively (Stephan et al., 2005; Hamacher and Meyer, 2005b): In order to collect the heterogeneous data of the HUPO BPP pilot studies in one database, the right database concept had to be chosen. The software ProteinScape (Bruker Daltonics Bremen & Protagen AG Dortmund, both Germany; free licenses by Bruker Daltonics) has been chosen for handling the heterogeneous data, as it is a feasible system for importing all the different data of a proteomics study, e.g., 1D gel electrophoresis, 2D liquid chromatography, 2-D difference gel electrophoresis (DIGE), etc. To learn the handling of this software, several ProteinScape training courses took place at periodic intervals and more will be held in the future.

The DCC is installed and a modified version of ProteinScape is running at 12 laboratories taking part in the pilot phase of the HUPO BPP. At the DCC all data has been imported with user-specific IDs. Dozens of gels and more than one million MS spectra were generated and transferred into the DCC. Data are being reprocessed according to a stringency set (Reprocessing Guideline, http://www.hbpp.org) and will be interpreted by invited analysts. Subsequent to the analysis phase all collected data will be exported by a newly designed exchange tool based on a mzData format into the database PRIDE hosted by the European Bioinformatics Institute (EBI) for worldwide access.

After the reprocessing phase the analysis phase will start with different task forces and different goals. The major analysis aspects are, among others, to match mRNA array data and protein data as well as peptidomics data, to analyze identified regulated proteins by interpretation of submitted protein lists (by participating groups) and gel images, to perform data mining and an overall analysis (summary, comparison pilot studies HUPO BPP and HUPO PPM, matching the results with literature).

The next steps prior to the master phase will be the completion of the pilot studies by presenting the results at the 4th HUPO World Congress in Munich (27 August–1 September 2005), by finalizing the interpretation at a bioinformatics jamboree and by preparing a joint publication similar to the HUPO PPP (Omenn, 2004a,b).

The activities of the HUPO BPP have been reported in several publications (Stephan et al., 2005; Hamacher and Meyer, 2005b; Marcus et al., 2004; Bluggel et al., 2004; Habeck, 2003; Service; 2003), newspapers, and other media. One of the most important interfaces with the scientific community is the HUPO BPP homepage, htpp://www.hbpp.org, as well as the discussion forum http://forum.hbpp.org, that offer an overview, news and the contact address of the project.

The results and considerations of the pilot phase will be used as the basis for the activities in the main phase. The network and the bioinformatics infrastructure will allow the performance of standardized differential analysis of neurodegenerative diseases.

In order to choose suitable and freely available (mouse) models for this next phase, the "HUPO BPP Symposium on Mouse Models" took place during the 4th Dutch Endo-Neuro-Psycho Meeting in Doorwerth/Arnhem, The Netherlands on 1 June 2005. Here, the most promising mouse models for Alzheimer's and Parkinson's diseases were presented and discussed, revealing the advantages as well as pitfalls of the different strains. Currently (at the time this review was written) the selection of the models to be analyzed is in progress, but will be finished by the 5th HUPO BPP Workshop that is planed for Dublin in February 2006.

The DCC and the bioinformatics tools, the network of the consortium and the developing structure of HUPO itself will definitely facilitate the reliable and reproducible analysis of neurodegenerative diseases by proteomics means. Nevertheless, HUPO BPP has several inherent peculiarities that are typical for large consortia projects, especially in regard to how willing the participating groups are to volunteer. First of all, active key players had and have to be identified throughout the scientific world by prominent intercessors, using existing email address lists and a publicity domain (announcements, articles, and contact with scientific journalists). Addressed researchers from both academia and industry had to be convinced that is essential to work together in this brain project though direct funding is not available. The motivations of the partners can be classified as follows:

- the conviction that these tasks can not be managed by single groups;
- the need for standardization and comparable results;
- contact with colleagues and the possibility for collaborations and discussions;
- increased publicity, less lobby work and national/EU funding applications.

Major problems mostly result from missing funding, e.g., most participating groups have to finance their HUPO BPP efforts from other sources, while other laboratories could not take part for this reason. As a consequence sometimes suboptimal analysis and unclear responsibilities are still prominent. This can only be overcome by the constant help and requests of the administrative partners and/or by long-term funding, e.g., by consolidation of HUPO, governmental support or industrial sponsoring.

Acknowledgements

The HUPO BPP is supported by the German Federal Ministry of Education and Research (BMBF) with grant 0313318B, the German HBPP is founded by the BMBF with grant 01GR0440.

References

Bluggel, M., Bailey, S., Korting, G., Stephan, C., Reidegeld, K. A., Thiele, H., Apweiler, R., Hamacher, M., Meyer, H. E. (**2004**). Towards data management of the HUPO Human Brain Proteome Project pilot phase. *Proteomics* **4**, 2361–2362.

Burt, R. S. (**1980**). Models of network structure. *Annu. Rev. Sociology* **6**, 79–141.

Cooper, R. G. (**2001**). *Winning at New Products: Accelerating the Process from Idea to Launch*, 3rd ed. Perseus Publishing, Cambridge.

Habeck, M. (**2003**). Brain proteome project launched. *Nature Medicine* **9**, 631.

Hamacher, M., Klose, J., Rossier, J., Marcus, K., Meyer, H. E. (**2004**). Does understanding the brain need proteomics and does understanding proteomics need brains? Second HUPO HBPP Workshop hosted in Paris. *Proteomics* **4**, 1932–1934.

Hamacher, M., Meyer, H. E. (**2005a**). HUPO Brain Proteome Project: aims and needs in proteomics. *Exp. Rev. Proteomics* **1**, 1–3.

Hamacher, M., Meyer, H. E. (**2005b**). Great mood in proteomics: Beijing and the HUPO Human Brain Proteome Project. *Proteomics* **5**, 334–336.

Hanash, S. (**2004a**). Building a foundation for the human proteome: the role of the Human Proteome Organisation. *J. Proteome Res.* **3**, 197–199.

Hanash, S. (**2004b**). HUPO initiatives relevant to clinical proteomics. *Mol. Cell Proteomics* **3**, 298–301.

Hermjakob, H., Montecchi-Palazzi, L., Bader, G., Wojcik, J., Salwinski, L., Ceol, A., Moore, S., Orchard, S., Sarkans, U., von Mehring, C., Roechert, B., Poux, S., Jung, E., Mersch, H., Kersey, P., Lappe, M., Li, Y., Zeng, R., Rana, D., Nikolski, M., Husi, H., Brun, C., Shanker, K., Grant, S. G., Sander, C., Bork, P., Zhu, W., Pandey, A., Brazma, A., Jacq, B., Vidal, M., Sherman, D., Legrain, P., Cesareni, G., Xenarios, I., Eisenberg, D., Steipe, B., Hogue, C., Apweiler, R. (**2004a**). The HUPO PSI's molecular interaction format – a community standard for the representation of protein interaction data. *Nat. Biotechnol.* **22**, 177–183.

Hermjakob, H., Montecchi-Palazzi, L., Lewington, C., Mudali, S., Kerrien, S., Orchard, S., Vingron, M., Roechert, B., Roepstorff, P., Valencia, A., Margalit, H., Armstrong, J., Bairoch, A., Cesareni, G., Sherman, D., Apweiler, R. (**2004b**). IntAct: an open source molecular interaction database. *Nucleic Acids Res.* **32**, D452–D455.

Humphery-Smith, I. (**2004**). A human proteome project with a beginning and an end. *Proteomics* **4**, 2519–2521.

Kami, K., Doi, R., Koizumi, M., Toyoda, E., Mori, T., Ito, D., Fujimoto, K., Wada, M., Miyatake, S., Imamura, M. (**2004**). Survivin expression is a prognostic marker in pancreatic cancer patients. *Surgery* **136**, 443–448.

Klose, J., Meyer, H. E., Hamacher, M., van Hall, A., Marcus, K. (**2004**). Human Brain Proteome Project – towards an inventory of the brain proteins. *BioForum Eur.* **8**, 28–29.

Marcus, K., Hultschig, C., Frank, R., Herberg, F. W., Schuchhardt, J., Seitz, H. (**2003/2004**). Innovative Forschungsansätze im NGFN Verbund 'The Human Brain Proteome Project HBPP'. *GenomXPpress.* **Dec.**, 5–8.

Marcus, K., Schmidt, O., Schaefer, H., Hamacher, M., van Hall., A., Meyer, H. E. (**2004**). Proteomics – application to the brain. *Int. Rev. Neurobiol.* **61**, 285–311.

Meyer, H. E., Klose, J., Hamacher, M. (**2003**). HBPP and the pursuit of standardisation. *Lancet Neurol.* **2**, 657–658.

Nakata, B., Takashima, T., Ogawa, Y., Ishikawa, T., Hirakawa, K. (**2004**). Serum CYFRA 21-1 (cytokeratin-19 fragments) is a useful tumour marker for detecting disease relapse and assessing treatment efficacy in breast cancer. *Br. J. Cancer.* **91**, 873–878.

Omenn, G. S. (**2004a**). Advancement of biomarker discovery and validation through the HUPO plasma proteome project. *Dis. Markers* **20**, 131–134.

Omenn, G. S. (**2004b**). International collaboration in clinical chemistry and laboratory medicine: the Human Proteome Organisation (HUPO) Plasma Proteome Project. *Clin. Chem. Lab. Med.* **42**, 1–2.

Orchard, S., Hermjakob, H., Julian, R. K. Jr., Runte, K., Sherman, D., Wojcik, J., Zhu, W., Apweiler, R. (**2004**). Common interchange standards for proteomics data: Public availability of tools and schema. *Proteomics* **4**, 490–491.

Palomaki, G. E., Neveux, L. M., Knight, G. J., Haddow, J. E., Pandian, R. (**2004**). Maternal serum invasive trophoblast antigen (hyper-glycosylated hCG) as a screening marker for Down syndrome during the second trimester. *Clin. Chem.* **50**, 1804–1808.

Sarbia, M., Bitzer, M., Siegel, D., Ross, D., Schulz, W. A., Zotz, R. B., Kiel, S., Geddert, H., Kandemir, Y., Walter, A., Willers, R., Gabbert, H. E. (**2003**). Association between NAD(P)H: quinone oxidoreductase 1 (NQ01) inactivating C609T polymorphism and adeno-carcinoma of the upper gastrointestinal tract. *Int. J. Cancer* **107**, 381–386.

Service, R. F. (**2003**). Proteomics. Public projects gear up to chart the protein landscape. *Science* **302**, 1316–1318.

Stephan, C., Hamacher, M., Meyer, H. E. (**2005**). 3rd HUPO Brain Proteome Project Workshop promises successful pilot studies. *Proteomics* **5**, 615–616.

Veraguth, T. (**2004**). Zukunft Biotechnol. *BIOforum* **7–8**, 20–22.

Zolg, J. W., Langen, H. (**2004**). How industry is approaching the search for new diagnostic markers and biomarkers. *Mol. Cell Proteomics* **3**, 345–354.

Useful World Wide Web links

http://www.biomedcentral.com:
The Scientist online journal.

http://www.scottish-enterprise.com/proofofconceptfund:
Scottish Proof of Concept Fund.

http://www.ruhr-uni-bochum.de/rubitec/start.htm:
RUBITEC – Society for Innovation and Technology of the Ruhr-University, Bochum, Germany.

http://ott.od.nih.gov:
Office of Technology Transfer of the National Institutes of Health.

http://www.zwm-speyer.de/:
Center for Science and Research Management Speyer.

http://prod-dev.com:
The Product Development Institute.

http://www.smp-proteomics.de:
The Human Brain Proteome Project as SMP within the NGFN.

http://www.ngfn.de:
Nationales Genomforschungsnetz, Germany.

http://www.hupo.org:
Human Proteome Organisation.

http://www.hbpp.org & http://forum.hbpp.org:
HUPO Brain Proteome Project.

http://psidev.sourceforge.net:
HUPO Proteome Standards Initiative.