Protein Binders for Characterisation of Human Proteome Function: Generation, Validation, Application

Reporting

Project Information

**AFFINOMICS**
Grant agreement ID: 241481

Funded under
Specific Programme "Cooperation": Health

Total cost
€ 14 620 568,35

EU contribution
€ 11 000 000,00

Closed project

Start date
1 April 2010

End date
31 March 2015

Coordinated by
THE BABRAHAM INSTITUTE
United Kingdom

Final Report Summary - AFFINOMICS (Protein Binders for Characterisation of Human Proteome Function: Generation, Validation, Application)

Executive Summary:
The AFFINOMICS project addresses a widely acknowledged demand for resources of well characterised and highly quality controlled protein-binding molecules (‘binders’) as affinity reagents - the most familiar of which are antibodies - for application in analysis of human proteins. The aim is to initiate the generation of a proteome-wide binder collection and to this end the consortium integrates the expertise and technologies available in 19 leading European centres and one SME in order to create an efficient pipeline for target and binder production, validation and quality control. As well as including the classically raised polyclonal and monoclonal antibodies, AFFINOMICS has made a strong effort towards technological improvements in
molecular selection systems, such as phage display and ribosome display for recombinant binder formats, to increase throughput and reduce cost. It has also put into practice novel technologies applying binders in high throughput, sensitive, specific protein detection and structure/function determination of protein complexes. The biomedical focus is on characterisation of 5 classes of proteins involved in signal transduction pathways in normal and cancer cells (kinases, phosphatases, SH2 domains, proteins mutated in cancer and plasma biomarkers) where the current need for specific, quality binders and novel approaches to detection, structure and function is especially great. Over 1400 protein targets were selected for inclusion. Particular attention has been paid to thorough characterisation and validation of the binders, an issue of particular concern to the antibody user community. A proteome-scale, binder-based approach to signal transduction has enabled analyses of how signalling complexes are assembled and connected to form larger networks and how they become dysfunctional. Information on all the protein and binders produced can be found in the publicly accessible project database portal, ProteinBinders (www.proteinbinders.org) which contains details of 1429 target proteins, 2900 antigens and 5595 binders.

Project Context and Objectives:
Investigation of the human proteome, including the composition, expression, modifications and interactions of all the proteins, is a major challenge with profound significance in both basic and medical research, as well as for the biotechnology and pharmaceutical industries. The ‘affinity proteomics’ approach to proteome analysis calls for systematic generation of specific binding molecules (binders) against all genome-encoded proteins and their variant forms. The development and impact of proteomic research in the coming years will be critically dependent on the availability of such reagents, as is now widely recognised. The research applications of affinity binders encompass many molecular protein analysis and separation technologies, while the medical applications in diagnostics, and increasingly therapeutics, play a major part in healthcare.

The most familiar affinity reagents are antibodies, yet despite the commercial availability of a very large number of antibodies, currently only a fraction of the human proteome is served by renewable binders. Moreover, the variable quality of commercially available antibodies has frequently given rise to concern among researchers, as it often leads to unreliable or artefactual results across different analysis platforms and a potentially untold waste of valuable time, effort and funds. Addressing the quality issue has recently become particularly prominent in the literature and at meetings. To solve or at least improve it requires increased binder characterisation and elaboration of their suitability for defined purposes; quality of binders is very much an application and context dependent matter. One solution proposed is to increase the deployment of recombinant binders so that recombinant antibody fragments and engineered protein scaffolds become viable alternatives to whole antibody molecules, significantly extending the range of structural possibilities. Another is to ensure that wherever possible binders are sequenced and those sequences are made available to users.

Thus, key objectives of AFFINOMICS are: (1) To enable systematic and rapid production of large sets of well-characterised, validated antibodies and other protein binders to meet the expanding demands to phenotype the human proteome. (2) To develop and implement novel concepts in automated methods for high throughput recombinant binder production. (3) To develop stringent and uniform quality control and characterisation methods assuring high binder quality. (4) To develop application tools for binders and to use them to analyse potential disease biomarkers. (5) To use binders to study the functional cellular
proteomics of signal transduction pathways and cancer-related proteins, particularly protein-protein interactions and complexes involving protein kinases, tyrosine phosphatases and SH2-domain proteins.

The key outputs of the project include: (1) Physical resources of validated binders and their antigens. (2) Standards and protocols for quality control procedures. (3) Recombinant binder production facilities and systems capable of the high throughput required to expand to a future proteomic scale. (4) Novel binder-based tools and technologies for both extracellular and intracellular protein detection and functional analysis. (5) Data on molecular profiling of signalling pathways, cancer related gene products and biomarkers. (6) A publicly accessible database of available binders created during the project, including information on availability.

AFFINOMICS targets are part of key pathways in cell signalling, which comprise networks of interactions, often involving protein phosphorylation and its recognition by modular interaction domains. In human cancers, dysregulated proteins in these pathways behave as disease drivers, making them important targets for therapeutic intervention. The proteome-scale binder-based approach to signal transduction as adopted in AFFINOMICS enables analyses of how signalling complexes are assembled and connected to form larger networks and how they become dysfunctional. Project partners aimed to generate sets of binding reagents for groups of structurally and functionally inter-related proteins, identified from the human genome sequence and literature sources. The first part of the AFFINOMICS project was primarily concerned with establishing the binder resource by means of a pipeline of production, from target to binder to characterisation and quality control, and deposition of the information on each one in a public database. In the second part of AFFINOMICS, the binders have been used to refine and apply novel tools and technologies of affinity proteomics to dissect signal transduction pathways in normal and cancer cells and characterise cancer-associated proteins and biomarkers. Clearly this is something that it is anticipated will be taken up by the community and continued after the end of EU funding.

To achieve the appropriate binder properties for diverse applications has required careful antigen design, from fully folded proteins or individual domains to unique peptide sequences and sequence specific modifications. The proteins have to be produced in an appropriate form for recombinant binder selection procedures, which in practice has meant producing each one with an immobilisation tag, such as biotin, as well as untagged (i.e. two forms for every protein). Clearly, this places special demands on the protein producers. Once produced, the binders themselves may need to be derivatised by addition of detection and amplification modules. An example of the latter is the ultrasensitive proximity ligation method where binders must be conjugated with oligonucleotides for subsequent detection by rolling circle amplification or PCR.

Thus, in terms of overall strategy, one strand of AFFINOMICS comprised a binder production and QC pipeline, with tasks distributed to partners according to their specialisations, and including development of novel recombinant selection systems for proteome scale delivery. A second strand has been to develop high throughput, analytical, binder-based technologies for investigation of signal transduction protein complexes. The main research sections of the project were divided into 7 workpackages (WPs). In WP1, antigens corresponding to about 1400 protein targets have been produced, including folded full length proteins and domains, PrESTs (long peptide fragments chosen for low inter-protein homology) and short peptides, including specific modifications. In WP2, the binders have been generated, including affinity-
purified polyclonal antibodies, monoclonal antibodies, recombinant single-chains, single VH domain camelid Nanobodies and non-immunoglobulin alternatives built on the protein A (Affibody) and ankyrin repeat (DARPin) molecular scaffolds. Recombinant selection systems have included phage and ribosome display. Characterisation, validation and quality control (QC) are the focus of WP3. In-depth analyses included epitope mapping, specificity screening on protein arrays and kinetic analysis, combined with application-specific QC in Western blotting, immunohistochemistry and immunofluorescence. In WP4, ‘next-generation’ selection systems for recombinant protein binders have been developed. The aim here was to achieve high throughput recombinant selection of binders expandable to the proteome scale. Streamlined robotic protocols and novel concepts have been introduced, scalable for 100-500 targets per 6 week period. WP5 has involved the development of new tools and applications for characterisation of proteins and biomarkers, especially for sensitive multiplexed analyses. Methods include protein and binder arrays and in situ detection of complexes with proximity probes and DNA amplification. Affinity-based functional proteomics of signalling protein complexes was implemented in WP6, where binders from WP2 have been combined with mass spectrometry for characterisation of signalling complexes. Finally, WP7 dealt with the bioinformatics and databases necessary for a functioning resource. The ProteinBinders portal contains detailed binder QC data and facilitates user access. Functional proteomics data is centralised in the IntAct database and laboratory protocols are publicly accessible through the MolMeth resource.

Project Results:
The AFFINOMICS workplan was divided into 7 workpackages, the results of which are described below.

WP1: Protein production
The task for the three protein producing partners (UOXF, UCPH, KI) was to deliver purified recombinant functional domains or full length proteins (kinases, phosphatases, SH2-proteins and others) expressed in bacteria, yeast, mammalian or insect cells, as target antigens for binder production. For UOXF and KI, the selected targets included 120 SH2 domains, 38 phosphatases (some with two catalytic domains) as well as 40 selected cancer targets (total of 250 proteins), while the focus target areas for the UCPH were protein kinases, proteins involved in ubiquitin signaling and other cell signaling proteins (total of 240 proteins). Production was staged over the 4 years. Wherever possible, proteins were to be produced in both biotinylated and nonbiotinylated forms for the purposes of selection of recombinant binders and assay of binders respectively.

During the first months a target list of more than 1000 targets was assembled in conjunction with all partners in the project. Targets were selected based on the criteria outlined in the proposal, namely protein kinases, SH2 domain-containing proteins, protein tyrosine phosphatases, proteins somatically mutated in cancers and candidate cancer biomarkers. To establish a functioning pipeline, targets were selected initially that are known to express well and for which purification protocols have been established in the protein producing labs. Proteins were purified to crystallisation grade purity and verified by mass spectroscopy (ESI-MS). For different binder producing labs, the proteins had to be expressed in different formats, e.g. UZH, ULUND and TUBS requested antigen in in vivo biotinylated form as well as non–biotinylated. All targets were recloned into the new vectors for biotinylation and the protocol for in vivo biotinylation was optimised. The output of proteins at the due date of 48 months showed that all targets were met, i.e. all 490 proteins were produced. Lists of the proteins can be found in the periodic reports as well as on the ProteinBinders website (www.proteinbinders.org).
For KTH, the objective in WP1 was the generation of antigens to be used in the polyclonal antibody production pipeline, namely Recombinant Protein Epitope Signature Tags (PrESTs) based on regions with low identity to other human proteins. An in-house developed software package, PRESTIGE, is used to select at least two separate and non-overlapping PrESTs to all 1000 proteins selected as targets. All AFFINOMICS targets were initiated in our production pipeline and production was completed on schedule, with 1680 MS-verified PrESTs to 1010 targets.

The group at DKFZ had the task of producing sets of relevant phospho- and nonphosphopeptide epitopes for immunisation and use in WP6 applications. Reactome and IntAct data were the basis for target selection. A target repertoire of relevant domains was identified and peptides mimicking the phosphosite, as well as the de-phosphorylated state, were generated. The peptides cover well-validated targets of published phosphosites within the MAPK pathway as well as novel ones identified within the consortium. 31 protein epitopes were chosen by HMGU from different signalling pathways. Ultimately, 58 phospho and 36 nonphosphopeptides from project target proteins were synthesised and delivered to partners HMGU and EKUT for antibody production and validation.

WP2: Binder production
A full list of the binders made over the course of the project can be found with validation data at www.proteinbinders.org.

1. Polyclonal antibodies: Partner KTH has been responsible for producing affinity-purified polyclonal rabbit antibodies against 1000 targets. They were generated through immunisation of rabbits with PrESTs from WP1 above. Sera were affinity purified on a matrix using the PrEST as antigen, resulting in monospecific antibodies. The antibodies produced cover genes from all the main categories of defined AFFINOMICS targets, i.e. kinases, phosphatases, SH2 proteins and proteins mutated in or related to cancer. During the project KTH generated 1219 monospecific antibodies for 1009 targets.

2. Monoclonal antibodies: EMBL MACF and PARATOPES together received 207 targets from WP1 partners for production of monoclonal antibodies (mAbs) in mice and rats. Between them, EMBL and PARATOPES produced a total of 127 of 165 targets attempted, a success rate of 77%. HMGU has produced antibodies against 47 out of 51 different target proteins (92%). The primary clones were further validated by immunoblotting (HMGU) and positive clones subcloned and stabilised. All the mAbs were delivered to BI for analysis on protein arrays.

HMGU has in addition focused on a peptide based approach for antibodies against pst-translational modifications (phosphorylations), in collaboration with DKFZ and EKUT, targeting important regulatory sites of proteins relevant to human disease. Mabs were raised against sequence defined phosphopeptides for further application in analysis of signalling pathways in WP6. Hybridomas were made against 46 phosphopeptides and their nonphosphorylated cognate peptides and 57 hybridomas were established. After an initial ELISA based screen, the supernatants of the positive clones were further characterised for specificity and selectivity by Western blot analysis and immunofluorescence. A series of phospho-specific mAbs was produced which have been tested in HEK293 cells expressing either the respective antigen in its phospho-specific form or in a variant, lacking this phosphorylation site.

3. Recombinant binders: ULUND and TUBS were tasked with producing binders of the scFv (ULUND) or scFv-Fc (TUBS) formats in 600 pannings against 600 targets, using proteins, PrESTs or peptides as antigens, depending on the supply of material from the target producers. During the entire project, TUBS generated 898 unique antibodies (scFv and ScFv-Fc fusions) against 191 different targets, while ULUND produced 1086 unique scFv against 206 different targets. An average of 5 binders was obtained per...
antigen, though in many cases considerably more were identified. The number of targets corresponded to the total number of biotinylated targets obtained from the antigen producers, making antigen availability a limiting factor in the number of successful selections. Although, phage selections were also performed on non-biotinylated targets, the success rate was much higher when using biotinylated targets compared to non-biotinylated targets. The relative shortage of biotinylated targets largely accounts for the fact that somewhat fewer targets were covered than originally in the workplan. After validation, binders made at TUBS were reformatted as scFv-Fc fusions, particularly useful for assays involving secondary anti-Ig reagents, and produced using the optimised mammalian protein production system. Their antigen binding properties were characterised in details by titration ELISA followed by protein arrays (BI).

4. Single camelid VHH domains (Nanobodies): At partner VIB, the aim was to enable rapid production of sets of characterised, validated recombinant camelid single VHH domain antibody fragments known as Nanobodies (Nbs), as a renewable resource of binders against various protein targets. In contrast to classic antibodies, camelid-specific antibodies are devoid of light chains and the first constant domain (CH1) of the heavy chain is absent. Camelid Heavy-chain only antibodies are a convenient source of single domain antigen-binding fragments, obtained by cloning the exon of its variable fragment. After immunising dromedaries with the protein targets following a 6 weeks vaccination protocol, an ‘immune’ library of the Nbs was constructed, cloned from the lymphocytes of the immunised animal. Selection of antigen-specific Nbs from this library was performed by panning on antigens coated on microtiterplates, using standard phage display. During the project, 20 animals (17 dromedaries, 3 alpacas) were immunised with a total of 220 antigens, with a resulting 268 unique Nanobody binders retrieved. Total targets covered include kinases, phosphatases, ubiquitin pathway proteins and cytokines.

5. Scaffold proteins: Affibodies: Small affinity proteins such as Affibody molecules, based on the protein A domain, with constrained binding surfaces on secondary structure motifs, have properties that are attractive in a number of different applications (e.g. in vivo tumour targeting). The aim for KTH was to select Affibody molecules using PrEST antigens using a selection strategy combining phage display with a recently developed Gram-positive bacterial display system (Gram-positive Staphylococcus carnosus). 50 validated binders against three targets (CD14, MAPK9 and ErbB) with affinities in the nM to sub-nM range have been selected. Binders have been extensively characterised and used in recent cancer-related studies. The performance of one ErbB3-specific binder has been evaluated for molecular imaging in a xenografted in vivo murine model of breast and colon cancer. The Affibody was site-specifically labelled with 99m-Tc and the analysis demonstrated specific uptake in ErbB3 positive tumours as well as healthy organs with ErbB3 expression. The work on the antibody library and the selection of ErbB3-binding Affibody molecules demonstrated that the bacterial display technology in combination with FACS is very powerful for isolation of the strongest binders from in vitro libraries.

6. Scaffold proteins: Designed Ankyrin repeats (DARPins) to native proteins were generated by the highly efficient ribosome display system established by UZH (see below), adapted for the selection against 94 targets in parallel. Native targets or domains provided from WP1 in biotinylated form using an Avi Tag, were used for selections in solution followed by capture using magnetic beads carrying Neutravidin or Streptavidin. The expectation was to produce DARPin binders against 250 protein targets, dependent on receipt of the antigens in the format of biotinylated/nonbiotinylated pairs, for selection and assay. In the event, DARPin have been selected against 185 target proteins and validated in different applications. Secondary screening was by competition ELISA using the soluble native protein as a competitor and tertiary characterisation involved semi-quantitative estimation of KD by competition with different concentrations of inhibitors. Specificity of purified DARPin expressed from E. coli was tested with suitable
controls, e.g. including related kinases or different phosphorylation forms. Positive clones were sequenced and local software employed to rapidly cluster selected binders, identify positively selected framework mutations and sequence correlation with ELISA results.

WP3: Binder validation
The focus of WP3 is quality control (qc) and validation to ensure the utility of the binders generated in the programme.
1. Protein Arrays: The aim was to characterise binders though multiplexed screening for specificity and cross-reactivity analysis using protein arrays (Partner BI). This was performed by exposing protein arrays to unlabelled binders, with binding detected by fluorescent secondary reagents. BI used a combined approach of DNA-encoded protein arrays (DAPA), together with conventionally spotted protein arrays, comprising all the project proteins and domains distributed by UCPH, KI and UOXF, and very high density proteome-wide (HuProt) arrays with up to 19,000 human proteins on a single microscope slide, for binder affinity and specificity screening. The entire collection of AFFINOMICS binders have been tested for specificity on medium scale protein arrays over the period of the project (years 2-5 inclusive). 2663 binders from 6 partners covering 317 targets have been screened on medium scale protein arrays of target proteins from WP1 partners, measuring reactivity against ~30 targets in parallel. 1003 binders (37.6%) gave scores of either 1 or 2 in the assay, indicating a high specificity for their own target protein. The most highly specific binders cover 209 separate target proteins.

The DAPA (DNA Array to Protein Array) technique was employed for proof of principle screens of binder specificity and cross-reactivity. A total of 139 DAPA constructs relating to AFFINOMICS targets were made successfully, focused on protein kinases, protein phosphatases, SH2-containing proteins, E2/E3 ubiquitin ligases and de-ubiquitinases (DUBs). The results have validated DAPA as a method of testing specificity of antibodies; however, in view of the success of screening on conventional protein arrays (above) this was used for most of the characterisations.

To supplement the above, we have also used the HuProt™ v1 Human Proteome Microarrays. They provided the largest number of unique, full-length, individually purified human proteins known to be included on a single slide at the time, allowing thousands of interactions to be profiled in a high-throughput manner. The arrays each contain more than 17,000 full-length human recombinant proteins, representing at least 13,200 different full-length genes that cover more than 65% of the annotated human genome. These full-length recombinant proteins are expressed as GST fusions in the yeast S. cerevisiae, purified, and printed on glass slides in duplicate, along with a control set of proteins (GST, BSA, Histones, IgG, etc.). Specific interactions could be observed for 7 of 15 mouse mAbs tested (47%) and for 5 of the 12 rat mAbs (42%).

2. Affinity determinations: Characterisation of binders in terms of specificity and affinity, and analysis of detailed kinetic parameters, was carried out using state of the art real time biosensors based on SPR. Kinetic analysis enables determination both of equilibrium binding constants and separate association and dissociation rate constants, providing a complete kinetic profile. For quality control, SPR measurements allow a ranking of binders according to desired properties (e.g. fast association, slow dissociation) and direct side by side comparison of different binder types, independently of a label or specific readout. Affinity and kinetic parameters and rankings for selected binders have been performed for 14 scFv-Fc fusions (from TUBS), 44 DARPins (from UZH), 20 Nanobodies (from VIB) and 9 scFv fragments (from ULUND). Affinities ranged from 10-6 to 10-10 M.
A direct side-by-side comparison of different binder types against the same target allows one to identify customized binders for specific purposes. Due to the availability of multiple binders to one specific target, Biacore based epitope binning experiments were performed in order to identify sets of binders suitable for sandwich assays, biosensor applications, proteomic approaches and diagnostic tools. A particular focus was put on binding analysis of protein kinases as targets, where multiple types of binders were generated by different binder producers. This included the targets/protein kinases AKT3-1, CDK2, MK01, MK08, STK17A, RPS6KA2, RPS6KA6, MAPK3, MAPK6, MAPK8, MAPK9 and GAK. Since multiple DARPin were available against MK08, RPS6KA2 and RPS6KA6 epitope binning experiments were performed. SPR data were in perfect agreement with UZH ELISA data.

3. Western blot: For the Western blot analysis (for polyclonal antibodies at KTH) a standard set up has been used with two cell lines (RT-4, U251 MG), two tissues (liver, tonsil) and human plasma. The blots are then given validation scores: ‘supportive’ if there is a band or bands, depending on which isoforms of the protein the antibody will recognize, of correct size in one or several lanes, ‘uncertain’ if there are no bands, or a band or bands deviating from the predicted molecular weight, ‘non-supportive’ score if there are one or several bands of the wrong size which cannot be explained by isoforms or other bioinformatics data. 40% of the 1200 antibodies tested produced a band of correct size in our WB analysis. Some proteins might not be expressed in the chosen samples and, since all antibodies, to some extent and under certain conditions, will show some cross-reactivity, both uncertain and non-supportive blots might still indicate more a sub-optimal set-up than unspecific antibodies. This is also indicated by 95% of all the antibodies generating a band of expected size when using lysates overexpressing the target protein.

4. Immunohistochemistry: Using full-scale tissue microarrays containing 46 tissues, KTH has performed analysis of 1291 polyclonal antibodies corresponding to 1067 targets. The staining pattern, staining intensity and fraction of stained cells are validated by a pathologist and a score for each tissue is determined. 33% of the antibodies scored as supportive. The uncertain score can reflect possible cross-reactivity, as well as, limited literature to base a decision on. There are no antibodies scored as non-supportive, since these will be filtered out on a first-step small-scale TMA and never stained using the full-scale TMA. An interesting finding is that most genes seem to be expressed, although at variable levels, in a majority of the tissues, while only a minor fraction is tissue specific. This has also been supported by mRNA analysis using RNA seq. Investigating the number of tissues that each protein was present in showed that half of the proteins were expressed in more than 40 of the 46 tissues, while only 0.5% of the proteins showed a more tissue specific expression and were expressed in fewer than 10 tissues.

5. Immunofluorescence: subcellular localisation of targets recognised by polyclonal antibodies in cell lines has been performed for 1316 antibodies corresponding to 985 targets. Results have been analysed regarding staining pattern and subcellular location across 14 compartments, using IF and confocal microscopy. 44% of the antibodies were supportive, 52% uncertain and just 4% unsupportive. An uncertain score most frequently reflects limited literature and only one antibody to base the decision on, while a staining pattern and localisation contradicting extensive literature will result in a non-supportive score.

6. Antibody microarrays: At DKFZ, an antibody microarray was produced with 458 recombinant binders (TUBS) and incubated with five sets of clinical samples. The analysis yielded 63 potentially informative antibodies. Microarray results were compared to Western blot analyses with respect to their information content about the specificity of the antibodies. The antibody microarray analysis was found to be superior in characterising specificity. One important issue in this respect is the ability to analyse many samples on many binders, thus achieving a larger amount of data, on which the evaluation can be based.
WP4: Optimising high throughput recombinant selection systems.

Reliable methods for the selection and screening of a large number of affinity reagents in parallel are a prerequisite to provide specific binders for various scientific projects.

1. Optimised ribosome display systems: for ribosome selection from DARPin libraries, automation at UZH consists essentially of liquid handling to be carried out robotically. The strategy has been to increase the handling from initially 12-20 antigens to 96 in parallel. Importantly, the project design will make it possible always to characterise, if needed, 100s to 1000s of binders to one target to be sure of also finding rare binders. In order to increase the throughput of the selection of specific DARPin binders, UZH focused on optimisation and validation of a semi-automated high-throughput selection and screening pipeline. An emphasis of the work was the further development of the ribosome display technology and its full implementation and optimisation for high throughput selection of DARPins by now 96 independent affinity selections can be performed simultaneously, UZH has been engaged in further improving previous protocols and implementing as many steps of the multistep protocol using various liquid handling workstations as well as the KingFisher® Flex, leading to successful performance of 48 independent affinity selections at a time. This represents a 12-fold improvement over current manual screening of ribosome display libraries and a 2-fold increase compared to the first trials in automation. The major bottleneck of binder generation is not the ribosome display per se, but the subsequent characterisation of the individual binders. Therefore, we further refined the previously established 384-well ELISA so that it can be performed in different variations with crude extract lysates. As no protein purification is required for these analyses, several hundred or even thousand clones can be easily screened within a reasonable time. After analysing the selected DARPins by dilution and competition ELISA, the coding sequences of the selected binders were determined. 80-90% of these DARPins showed unique sequences, emphasizing that there is almost no "focusing" of the ribosome display libraries. Multiple assays to analyse the stabilisation of the targets by DARPin binders as well as "epitope binning" were established. Having established such a massively parallel selection system, the major bottleneck of binder generation is no longer the ribosome display, but the subsequent characterisation of individual binders. Since the quality of the ribosome display has also increased further, this abundance of binders is a rich source from which to pick both high affinity binders as well as those recognising different epitopes, all with high specificity and monomeric behavior. The key challenge is how to do this most efficiently.

2. Optimised phage display systems: ULUND has developed new and improved automated high throughput selection systems by introducing streamlined robotic protocols, ultimately scalable for 50-100 targets per 6 week period. Technologies encompass phage display of antibody fragments from naïve, synthetic libraries. The initial work included setting up of robotic systems to be used for screening processes, such as a liquid handler, colony picker, microtiter plate wash stations, and a Kingfisher (automated magnetic bead dispensing and washing device), as well as adaption of protocols for the new instrumentation. In addition, specialised production vectors, with optimal tags for purification and applications, were designed, constructed and evaluated. These vectors allow efficient re-cloning of selected genes in pools. This semi-automated set-up allows one person to screen (including re-cloning, initial ELISA analysis, cherry picking, validation ELISA analysis, sequence analysis) approximately 2000 clones within a 3-week period, which was well within the aim of AFFINOMICS. The decision was made to focus on an optimized bead selection platform. This has since turned out to be quite successful in delivering the svFv at a rate set out in the workplan.

3. Optimised bacterial display systems: at KTH, the aims was to further develop the selection system using
Gram-positive bacterial display for Affibody binders to allow high-throughput selection of recombinant protein binders. The focus has been on the Staphylococcus carnosus surface display system, which has the advantages of easy recombinant cloning, large libraries and the use of flow cytometry (FACS) sorting for rapid binder selection. The technology is now validated and both antibody libraries and alternative scaffold libraries have been used in the evaluations. The binders selected are primarily Affibodies, but Nanobodies have also been displayed. The main focus has been on optimisation of conditions for reduction of background to obtain an efficient and rapid flow cytometry (FACS) sorting of binders from de novo selection of the recently made naive Affibody library and to perform de novo selections of chosen targets. A panel of new bacterial display vectors has been constructed and generated new important information regarding essential and non-essential regions of the vector as well as a decreased length, which will potentially increase the transformation efficiency. The selection system works well and a number of de novo selections against different targets have been performed, with low nM Affibody binders against three targets (CD14, MAPK9 and ErbB) having been obtained and characterised.

WP5: New tools and applications of binders

1. Binder arrays for serum profiling. This task focused on the application of recombinant scFv binders on arrays in the evaluation of the five major target categories of Affinomics. ULUND has constructed large scale scFv antibody arrays and used them in biomarker detection and discovery for clinical samples. Over the last 10 years ULUND, has developed an affinity proteomics platform, based on specially designed single-framework antibody fragments, which has been used for multiplexed analysis of tumour tissue in gastric adenocarcinomas, and serum analysis of patients suffering from glioblastoma, breast and pancreatic cancer, SLE and systemic sclerosis, chronic lymphocytic leukaemia, etc. The content of these capture arrays comprises hundreds of engineered scFv fragments directed against proteins of the adaptive and innate immune system. Consequently, a wide experience exists in the design and bioinformatics evaluation of capture arrays focusing on biological relevant diseases, events, pathways etc. The focus during AFFINOMICS has been to perform a set of clinical applications within the area of cancer and autoimmunity. We have had direct access to a number of large cohorts of well-characterised clinical serum and/or plasma samples via several pre-established collaborations with leading clinical researchers. The clinical need has been defined (e.g. lack of specific and sensitive blood-based test for diagnosis of pancreatic cancer) together with the clinicians, and we have then used our recombinant antibody microarray platform to decipher serum/plasma biomarker signatures providing the clinicians with actionable information, responding to the pre-defined clinical need(s). Examples include profiling of markers in pancreatic cancer, for disease diagnosis, in breast cancer, for tumour grading, and in Systemic Lupus Erythematosus (SLE) and B cell lymphoma.

2. Data on protein binders evaluated in proximity ligation assays. At UU, different approaches have been used for conversion of the AFFINOMICS binders into proximity probes through attachment of DNA oligonucleotides, either by gene fusion with e.g. SNAP-tag protein domains or by indirect labelling. Together with UZH, various conjugation strategies were explored to couple DNA oligonucleotides to DARPinS. Expressing the DARPinS with a cyclooctyne, introduced as an unnatural amino acid was a relatively easy protocol for cloning and expressing the binder, which generated DARPin-DNA conjugates with a similar efficiency to the SNAP-BG approach. Partner TUBS have now tested the same approach for expressing scFvs for conjugation to azide-modified oligonucleotides.

3. In situ protein arrays. An improved system for cell-free expression of protein arrays based on DNA arrays has been developed by BI. The technology uses an array of DNA constructs for cell-free
expression, which acts as a template instructing the generation of the corresponding protein array. Proteins are expressed locally from these templates by a cell free transcription and translation system, and are immobilised on a separate capture surface overlayed on the DNA array. The method enables protein-coding DNA to be used to create protein arrays with minimal time and effort, since all steps take place within the slide sandwich, avoiding separate protein expression and purification. By simplifying the setup to allow protein diffusion between the slides across a gap filled with the cell-free system, the evenness of the resulting protein microarrays has been markedly improved. A number of optimisations for the DAPA platform were carried out to define the best surfaces and optimal expression conditions.

In a similar concept at DKFZ, protein microarrays are being produced by in situ synthesis from PCR-products made from full-length cDNA libraries. At least 93% of the PCR-products yield full-length proteins, if expressed in situ on the microarray by an in situ transcription/translation process, as determined by the application of binders that are specific for tag-molecules attached to the N- and C-termini of the proteins. The most complex array produced comprises 15,000 proteins. Arrays from human cDNAs have been produced successfully, indicating the wide applicability of the process. Microarrays were also made by a newly developed process of in situ PCR-amplification and subsequent protein expression (DKFZ). They are produced directly from the RNA/cDNA of an individual patient. In consequence, all personal variations in sequence or RNA processing are represented in the on-chip PCR-products. Subsequently, they are translated into the respective protein isoform(s) that characterise the patient. The basic process is established now and used for the creation of individualised protein representations.

4. Epitope mapping on peptide arrays. In situ parallel synthesis of peptides enables the creation of arrays with up to 175,000 unique peptides. The large number of peptides enables array design with overlapping 12-mer peptides with a single amino acid shift covering more than 1000 PrEST-sequences. The single amino acid shift provides very precise mapping of minimal linear epitopes required for binding. At KTH, the planar peptide array set-up has been evaluated by epitope mapping of 947 polyclonal antibodies on an AFFINOMICS specific array containing peptides from 814 AFFINOMICS targets. The data, and corresponding antibodies, have then been used in a targeted proteomics approach resulting in a method (immuno-SILAC) for absolute quantification of proteins in complex samples. The majority of the polyclonal antibodies recognize two or more linear epitopes.

WP6: Applications in proteomics.

1. Exploring the interactome of protein tyrosine phosphatases. The goal of the subtask was to demonstrate how the binder resource can be used to discover new modifiers of medically important pathways. Joint work between EKUT and UTOV on identifying new substrates and network properties of phosphatases critical for cellular signalling pathways complementing our antibody driven approach has led to a comprehensive description of phosphatase networks regulating phosphorylation in human tissues. Combining pull-down proteomics and in silico modelled network context, new phosphatase substrates in growth pathways have been described. Integrating information from functional siRNA- as well as protein interaction screens, we have elaborated a strategy that aims at inferring physiological substrates of phosphatases. The strategy was validated by identifying the ligands of eight phosphatases whose downregulation affect the activation of the PI3K-mTOR and ERK pathways. Graphical analysis was then used to identify protein scaffolds that may link the catalytic subunits to their substrates. By this approach we have characterised two new protein

2. Affinity-based characterisation of protein complexes in the Ras/Raf/MAPK pathway. The EGFR signalling pathway is involved in the regulation of crucial biological processes such as proliferation,
survival, differentiation and growth. Additionally, alterations in this pathway have been associated with human malignancies such as colorectal, pancreatic, lung and breast cancer, which makes it of great interest for targeted therapy and drug design. In this work the goal was to untangle, using a quantitative proteomic strategy, dynamic components of the EGFR pathway. Analysis by partner EKUT indicates that a single mutation can create a cascade of functional changes on the level of protein binding patterns and change the signalling information flow. Accordingly, they experimentally validated interaction modulations of selected driver nodes using AFFINOMICS antibodies derived from WP2 and investigated the possibility, that specific elements within the EGFR pathway serve as potential molecular signalling modulators (either enhancing or attenuating signal quality and signal strength). They pinpointed dynamic modules within the interactome (most importantly around the B-RAF_MEK_ERK1,2 signalling machine) that provide insights for future quantitative dynamic analysis. This antibody driven approach could confirm critical interactions as well as correlating interaction patterns with MEK and ERK activity (quantified by MEK and ERK phospho-specific antibodies created in WP2 in combination with pull downs and mass-spectrometry).

3. Characterisation of the interactome of 5 RNF11 mutants. In another proteomics application, UTOV has used the RNF11 protein as a marker of vesicular trafficking to investigate the dynamic formation of complexes as the protein moves along the secretory/endocytic pathway. RING finger protein 11 (RNF11) is a RING ubiquitin ligase that is often found overexpressed in tumours, with important roles in growth factor signalling, ubiquitination and transcriptional regulation. UTOV have applied affinity purification coupled to mass spectrometry (in collaboration with EKUT) to reveal the network of interactions centered on RNF11, to identify molecular partners involved in the endosomal maturation process. We validated binding between RNF11 and ANKRD13 family members by co-immunoprecipitation and immunofluorescence assays and found that, intriguingly, RNF11 interacts with the ANKRD13 proteins by recognizing a region containing UIM repeats, thus revealing a novel mode of interaction based on a mechanism that possibly mimics the recognition between Ubiquitin and a specific class of Ubiquitin binding domains. Taking advantage of the purified ubiquitinating and de-ubiquitinating enzymes supplied in AFFINOMICS, we are validating some of the interactors and/or putative RNF11 substrates identified by MassSpec analysis.

4. In situ Proximity Ligation assays (isPLA) developed for important interactions in cancer pathways. Mutations that affect intracellular signalling pathways are likely to result in alterations of protein interactions, and such alterations thus represent promising markers of malignancy. Proximity ligation assays (PLA) permit protein interactions to be investigated directly in normal and pathological tissue samples. Recombinant AFFINOMICS binders were used to investigate interactions and posttranslational modifications resulting from cross-talk between pathways prone to mutations in cancer was demonstrated, including those for TGFβ, WNT, EGF and NGF signalling. The TGFβ pathway is involved in cross-talk with pathways in which multiple cancer genes are mutated. The pathway is itself targeted by mutations/deletions of SMAD4 and the type II receptor (TGFβRII). Partner UU developed the assays based on post-translational modifications on SMADs and in particular have investigated a set of scFv binders provided by Partner TUBS developed towards SMAD proteins, with a focus on a binder directed against the Thr179 phosphorylation site in the linker domain of SMAD3. This reagent was used to investigate protein interactions between post-translationally modified SMAD3, SMAD4 and other proteins in several cancer types. The increased levels of linker domain phosphorylated SMADs has been detected in the invasive front of solid tumours, and linked to progression through the cell cycle. Increased levels of p179SMAD3 were detected in mitotic Hs578T metastatic breast cancer cells. There are no quality
commercial reagents available for this SMAD3 linker phosphorylation. This is an example of the detection of a PTM important in cancer, with a potential as a reagent for tumour biology and as a tool in drug screens. They also observed that EGF treatment increased the number of proximity events detected between GSK-3b, a member of the Wnt pathway and the linker phosphorylated SMAD3. The interaction between GSK3b and SMAD is known from the literature, and here we could show that EGF potentiates this interaction.

In order to evaluate recombinant binders from another important cancer pathway, UU chose to validate the performance of a scFv (partner TUBS) developed towards the pro-peptide form of the human Nerve Growth Factor (proNGF). There is insufficient coverage of commercially available antibodies for the human antigen. The NGF circuitry has been implicated in migration/invasion of several solid tumour forms, including melanoma. We have discovered by TAP-MS a novel proNGF binding protein, referred to as PBP, validated the interaction with Biacore and applied in situ PLA to detect the proximity between PBP and proNGF in malignant melanoma cells.

The toolbox of AFFINOMICS binders offers unprecedented possibilities to scale and standardise commercial assays with complete coverage of the SMAD proteins. We aim to combine recombinant binders towards components of both the neurotrophin (proNGF) and the TGFβ (SMAD) pathways. Such assays will offer the pharmacological industry the possibility to standardise cellular read-outs using recombinant binders, avoiding batch-to-batch variations that seriously complicate assays using existing commercial antibodies. We are furthermore developing assay formats where multiple interactions/modifications can be simultaneously assessed using divers affinity reagents. Such assays will be valuable in basic cell biology, as well as for measuring drug responses to targeted and nontargeted signalling pathways in preclinical drug development, and perhaps also for monitoring patient responses to therapy in order to optimise this therapy.

5. Chromobodies. These are intracellularly expressed Nanobodies fused with GFP or RFP. It has been shown that transfected Chromobodies are expressed into mammalian cells and can trace target location and traffic in living cells when followed by confocal microscopy. It appears that 80-90% of nanobodies are indeed functional as intrabodies, even in the reducing environment of the cytoplasm; thus GFP-fusions will be generated from all the nanobodies obtained from WP2 to assess their antigen-tracing capacity in target cells. Partner VIB has generated a versatile vector with a fixed monomeric RFP and restriction enzyme sites to ligate any AFFINOMICS Nanobody genes under control of CMV promotor. This vector is ready for use by Afinomics partners. A vector to express and purify the Chromobody protein was also prepared.

6. Crystallisation of Nanobody–kinase complexes. Binders can be used as ‘crystallisation chaperones’, in particular for highly dynamic proteins such as kinases as they lock a protein in a specific conformation. Not only does that increase ‘crystallisability’ of these proteins, but it can also provide insight into conformational dynamics of protein molecules and their catalytic domain plasticity. Because of their compact size, rigidity and crystallisability, Nanobodies and DARPins are particularly promising for these studies. At VIB, over 40 crystal structures of Nanobodies in complex with their cognate antigen have been solved while DARPins have been used to cocrystallise kinases in different functional conformations. G-associated kinase (GAK) is essential for numerous cell functions, including clathrin trafficking, mediating binding of clathrin to the plasma membrane and the trans-Golgi network, and regulating receptor signalling by influencing trafficking downstream of clathrin coated vesicles. In addition to its role in the cytoplasm, GAK has important functions in the nucleus. Four Nanobodies (Nbs) against GAK were obtained at VIB after phage display of an immune camelid library as described above. Two X-ray crystal structures of GAK with 2 different Nanobodies were solved by Partner UOXF. The complexes lock the kinase into different
conformations. The complex with Nb4 traps the kinase in an inactive dimeric conformation that has also been seen when the kinase is crystallised on its own. In this dimeric conformation the activation segments bind to the substrate binding site of the interacting protomer allowing autoactivation by trans-phosphorylation. In the GAK-Nb1 complex the kinase assumes a conformation that corresponds to the structure of the monomeric active kinase. The structures provided insight into the conformational changes of GAK kinase. The two co-crystal structures revealed very different binding modes of Nanobodies which will help for future design of selective and potent binders. Further co-crystallisations of other kinases (TOPK, RPS6KA6) with Nanobodies and DARPins are being attempted. The results with GAK were recently published.

1. Public release of PSI-PAR 2.0. The PSI-PAR (Proteomics Standards Initiative – Protein Affinity Reagents) standard was earlier developed as a common data exchange format for binder production and application. In Affinomics, PSI-PAR is used and further developed by Partner EMBL EBI to ensure standardised information exchange between all partners, and in particular between the databases (further) developed within this WP. Consultation with the Affinomics user community has indicated that the format has not been employed for use case 1 (production data) and as this proved the most difficult to model in PSI-PAR 1.0 it was agreed that this need not be supported in future versions of the format. Use cases 2 (characterisation/quality control data) and 3 (summary data) were also difficult to model in the standard exchange format for molecular interactions (PSI-MI XML2.5) because this was often abstracted or merged data whereas PSI-MI XML2.5 has a mandatory requirement for each observation to be linked to the details of an actual experiment. The 3.0 schema is now flexible enough to describe characterisation/quality control data as well as summaries of end products. The schema was finalised and agreed at the 2015 PSI Spring workshop and is currently being written up for publication. The development of common interchange formats enabled the initial development of the Proteomics Standard Initiative Common QUery InterfaCe (PSICQUIC) service. A new PSICQUIC specification was developed, encompassing extensions to the MIQL query language and a completely new implementation of the PSICQUIC reference server. This enables binder producers and providers to make experimental data publicly available using the PSICQUIC webservice.

2. IntAct Affinomics data import and dissemination. Capture of Affinomics data across the three specialised resources (ProteinBinders, IntAct, MolMeth) maximises the re-use of previous (EU) resource investments and minimises the risk associated with the construction of large and complex, project-specific data resources. At the same time, the use of a common exchange format between all partners ensures efficient, automated, and standardised data communication with access to detailed data for complex tools like AffinomicsSelect. Users will have efficient access to Affinomics support through all phases of a research project, from exploration of existing molecular networks and pathways data including relevant binders in IntAct via binder-related experiment planning support through AffinomicsSelect and ProteinBinders to detailed application protocols in MolMeth. EMBL EBI have continued to curate papers into the IntAct molecular interaction database which describe data produced using the reagents and techniques developed under the Affinomics grant, in particular high confidence interactions determined by "Proximity Ligation Assay (MI:0813)", the method pioneered by Professor Ulf Landegren (UU). They are also monitoring the ECAS listing of consortium publications, and will curate any relevant publications into IntAct or MINT.

3. MolMeth database. MolMeth is a database of experimental protocols and SOPs, developed and
maintained by UU, which has been further developed to provide a central repository for all Affinomics characterisation and application protocols (www.molmeth.org). The new version of MolMeth is a web-based platform with semantic integration and large focus on user friendliness and search engine optimization for maximum visibility; currently it contains 137 protocols and SOPs. The rapid pace of development of new research tools and methods makes it increasingly important that technical notes, experiences and protocols can be shared even before scientific results are submitted, reviewed and accepted for publication. This is especially important in the field of protein-binding molecules as the number of potential combinations between methods and affinity reagents is exponentially related to the number of methods and affinity reagents available to researchers. The Molecular Methods database was therefore introduced into the Affinomics project as a means to provide swift and efficient publication of protocols and standard procedures to help researchers conduct their experiments more efficiently. Work on MolMeth has resulted in the fulfilment of three objectives: (i) MolMeth Workshop to define database standard and ontology; (ii) MolMeth prototype database containing new modules for enhanced functionality; and (iii) MolMeth database fully functional and deployed with Affinomics web portal (M60). The average number of users per month remains stable at around 900 according to Google Analytics used to continuously monitor web site usage.

4. Release of a new version of the ProteinBinders portal. Partner KTH is responsible for developing ProteinBinders, the project portal for validated protein binder information, providing detailed characterisation data for antibodies as well as other protein binding reagents (www.proteinbinders.org). Data on all antigens and binders generated within Affinomics together with binder characterisation data are submitted to, and made publicly accessible through ProteinBinders. The portal is arranged with a gene-centric structure allowing grouping of antigens and binders related to one gene product. Searching the portal can be done using a simple and intuitive free text search. A more advanced interface is also available allowing combined searches for binders targeting for example a certain protein class and with characterisation data from a certain assay. Properties of antigens, binders and characterisation data including producing partner can also be searched for. The website has been made publicly available and is continuously updated with antigens, binders and validations generated within the consortium. Integration between Molmeth and ProteinBinders enables linkage of protocols with specific antigens and binders. The Human Protein Atlas project (KTH, Stockholm) has generously committed itself to finance and support the portal for one year after project end, but after this period further outside financing will be required to keep the portal running.

Potential Impact:

POTENTIAL IMPACT

It is well recognised that there is an urgent need for well-validated, protein-specific molecular probes – antibodies being the most familiar - to explore the functions of the human proteome and that this is a major challenge in the life and medical sciences. In view of the importance of proteins and proteomics for understanding the normal function of the human (or animal) cells and their alteration in disease, a coordinated effort to structure the area of binding reagents has been urgently needed; a major impact of AFFINOMICS has been as the first project to establish such an effort in Europe. Proteome wide binder resources will have a profound impact on European basic research as well as the biomedical, pharma and biotech industries. Some of the beneficial impacts and opportunities which will result from access to such a resource include biomedical research, drug discovery, diagnostics, therapeutics and biobanking, as described below.
Given the complexity of the proteome, it is difficult (currently impossible) to put a definitive number on the different protein species which exist in a human cell. This is not simply, or even chiefly, a result of the size of the genome, but more the outcome of complex post-translational alterations to the proteins which profoundly affect their function. Therefore with a broad and ambitious programme such as AFFINOMICS, necessarily limited by the available funds to what it can realistically achieve, it has been important to provide an initial focus for what is essentially a preliminary to a much larger entity, potentially a new European infrastructure of binding molecules. The area chosen has been one of major importance to human health, namely that of proteins involved in signal transduction, cancer aetiology and as disease biomarkers. This has proven to be both a challenging and at the same time an appropriate target choice for such a project. AFFINOMICS has used this target focus to demonstrate that it is indeed possible to put into practice a European collaborative research programme aiming to provide publicly accessible resources of validated affinity reagents. The ultimate extension would be to establish an infrastructure directed towards coverage of all human proteins and to apply the reagents so obtained to generate comprehensive data on the human proteome in health and disease. One conclusion of AFFINOMICS is that this very broad aspiration will have to be tailored and approached in an alternative way in the future. Rather than attempting to cover every possibility, which will be near impossible in both financial and logistical terms, a resource will have to be closely responsive to the needs of the research community, i.e. demand driven irrespective of the functional area, rather than ‘stockpiling’ or building up a resource without reference to the shifting requirements of research.

The proteomics area is a combination of protein identification, led by mass spectrometry, and functional description; affinity binders are involved in both. Through their ability to detect proteins with high specificity and sensitivity in dedicated technologies and applications, affinity reagents are among the most universal and powerful analytical tools for life sciences research, diagnostics and, increasingly, as therapeutics for cancer and autoimmunity. They serve not only as probes for detection and discovery, combining with mass spectrometry in techniques such as SISCAPA, but more and more in multiplexed research applications. The discovery of novel targets reaches a major bottleneck when there are no reagents for further investigations, including hypothesis driven and highly focused applications. Therefore, it is clear that the availability of high quality protein binders will be enabling for all areas of investigation into human proteins, making possible comprehensive biological studies of a scale and type not previously achievable and thus guiding the proteomics area in the near and longer term future. The AFFINOMICS programme has contributed significantly to this requirement.

Moreover, accuracy in post-genomic systems-oriented analysis depends on the reliability of the experimental methods and the coverage of the approaches. An area of considerable impact will therefore be in the quality assurance that antibodies and other binders do what they claim to do. This issue is one which has been frequently addressed in meetings of the AFFINOMICS partners (see below).

Improving binder production

Through the pipeline established in WPs 1-4 for target and binder production, together with validation and quality control, AFFINOMICS has significantly improved the technology foundations of a systematic and complete coverage of the proteome. All stages in the process have been optimised, in particular methods for production of recombinant binders, such as phage and ribosome display, which are central to the
demands of high throughput production, and where considerable progress in throughput and automation have been achieved to a degree that proteome-wide questions can now be addressed. Optimisation and improvements have also extended to protein production, given the needs of recombinant selection procedures for efficient immobilisation procedures. At the other end of the affinity reagent pathway, multiplexed selection and screening methods, together with miniaturisation of evaluation assays, have been implemented, which will reduce the target protein requirements which are a large part of current costs. Bringing these steps together will allow the generation of massive numbers of renewable binders in the future. By exploring these aspects in detail, AFFINOMICS has demonstrated the practical aspects of high level generation and application of renewable affinity reagents. The impact of such systems can be likened to that of next generation sequencing in the genomics field, catalysing innovation in high throughput functional proteomics. A quantum leap in terms of speed and cheapness is, however, still awaited; the aim is to be able to receive a validated binder within a few days of requesting it, and while binder production has been accelerated by some impressive new technologies, the testing and quality control are still both the most time consuming and expensive part of the entire process.

Improving reagent quality
Affinity binding reagents, particularly antibodies, and their application systems are among the most commercial of all biological products and important drivers of progress in the life sciences, and the commercial sector has of course for a long time produced those reagents for which there is an immediate customer demand. However, while the number of commercially available antibodies is in the tens of thousands (www.antibodypedia.com) it is clear that validation and reliability of the affinity reagents is frequently less than expected or hoped for. This is therefore particularly important and of pressing concern, given the probability that as much as half of the commercially available antibodies today have question marks around functionality, in many cases giving unreliable results across several analysis platforms. This in itself results in an enormous waste of resources, due to the purchase and use of these poorly functional affinity reagents and it is not inconceivable that many hundreds of millions of euros of research funding in academia and pharmaceutical companies are spent unnecessarily each year. More importantly, huge amounts of time and effort are wasted due to the time and effort on generated data of dubious quality. It is hard to overstate the extent of this problem, which has been highlighted recently in a Nature article which was fully supported by AFFINOMICS partners (Bradbury A, Pluckthun A, 2015, Antibodies: validate recombinants once, Nature 520;295). AFFINOMICS has addressed a much-required strategy for validation and quality control of antibodies and other affinity reagents, with introduction of standards and methods to assess not only those made expressly within the partners’ own facilities, but from all potential providers, commercial as well as academic. The validation methods employed have ranged from Western blotting, immunohistochemistry and immunofluorescence to protein microarrays, binding kinetics and epitope binning. Consistent validation of the affinity reagents so that they become trusted tools irrespective of provider source is particularly vital to their reliable application and will have a major beneficial impact for users. Access to the validation data made has been made publicly available in the project ProteinBinders database (www.proteinbinders.org) and will encourage the adoption of ‘gold standard’ criteria for testing of the reagents which in time will become an essential part of their description and marketing. Another suggestion of the aforementioned Nature article is for every binder in use as a reagent to be sequenced, something which is only really practicable for recombinants. In the AFFINOMICS project, the majority of recombinant binders have been sequenced which adds further reliability to the validation.
Novel binder-based tools for functional protein analysis

As important as the increased throughput of production is the range of novel analytical tools that will become available with the new generation binders. Within AFFINOMICS, there has been a focus on microarray technologies (arrayed binders, protein arrays, peptides, lysates and cells) and proximity ligation systems. Effective technologies to identify the various binary interactions among key signalling proteins in situ are very much needed, since the pattern and extent of various combinations of protein interactions is expected to be indicative of cellular state and tumorigenic potential. Extending such analysis to human cancer samples would also make it possible to characterise signalling systems in patients before and after treatment with anti-receptor therapeutics, as shown in results using in situ PLA. Another area of functional analysis where AFFINOMICS binders have been proven to be ground-breaking tools is in the action of kinases. Nanobodies recognising different functional states of the kinase GAK have been co-crystallised with the enzyme and the structures have yielded important information about the changes which take place on activation. Offshoot studies are planned with other kinases using Nanobodies and DARPins.

Applications in cell signalling and cancer

Binder-based applications are expected to have a major impact in disease related studies. The repertoire of binders raised within AFFINOMICS has been focused on proteins in key cell signalling pathways, which are closely connected with the molecular basis of cancers and other diseases. There are many aspects of these pathways for investigation which await the availability of suitable affinity reagents, making signal transduction an excellent model proteomic system for evaluation of binders in a functional context. A number of the most successful and widely used drugs exert their effects on these pathways, while kinases themselves represent one of the largest druggable class of targets. At the same time, genome sequencing results have shown great variability in cancer mutation patterns and produced many candidates which have yet to undergo biological validation, for which sets of binders can make a unique contribution. During AFFINOMICS partners have utilised the reagents and application technologies to study cellular protein interactions, complexes and networks in signal transduction (see WP6 deliverables). This is another area of impact, given that more than 1% of all human genes are implicated in cancer via somatic or germline mutation, including many encoding signalling proteins; comprehensive sets of binders targeted to these proteins will be vital in providing insights into the processes ultimately underlying neoplastic change. The development of validated reagents for detection and potentially modulation of cell signalling pathways will thus be of immediate benefit to clinical research in this vital area and provide new leads for the pharmaceutical industry.

Understanding the molecular events involved in these pathways is also a major focus of both basic research and drug development. The toolbox of AFFINOMICS binders offers unprecedented possibilities to the pharmacological industry to standardise cellular read-outs using recombinant binders, avoiding batch-to-batch variations that seriously complicate assays using existing commercial antibodies. Moreover assay formats have been developed within AFFINOMICS, where multiple interactions/modifications can be simultaneously assessed using diverse reagents, such as the versatile proximity ligation assays. The assays will be valuable in basic cell biology, as well as for measuring drug responses to targeted and nontargeted signalling pathways in preclinical drug development, and perhaps also for monitoring patient responses to therapy in order to optimise this therapy.

Biomarker discovery and assay
A similar situation pertains to identification of biomarkers, many of which have been reported in preliminary fashion and require more thorough validation before acceptance. Binder-based methods will define cancer protein signatures and detect them as biomarkers in tools such as high content capture microarrays. They promise to generate a unique compendium of proteins associated with disease progression, stratification and therapy not available by other approaches, with clear applications in diagnostics, monitoring of plasma, urine, tissue sections, etc. The standardised reagents and procedures in the AFFINOMICS resource will be of enormous benefit both for applications in monitoring the proteome, and for applications in plasma diagnostics and classification of complex disease states, particularly cancers. This has been well illustrated by the work of Professor Borrebaeck and colleagues (partner ULUND) who have monitored biomarker signatures in pancreatic and breast cancer and in autoimmune disease such as SLE using arrays of recombinant binders made in the AFFINOMICS project.

Impact on Systems Biology

While great advances have been and continue to be made in intracellular protein analysis, studies of the role of endogenous proteins in cells remains challenging. Different approaches to this have been taken in the AFFINOMICS project, through fluorescence microscopy, and proximity assays. In this, the affinity reagents will directly impinge on systems biology approaches, which are in need of analytical methods with improved precision and throughput. In particular, affinity reagents, after suitable labelling or modification, in techniques such as high resolution imaging and proximity ligation, allow individual cells and individual molecules to be investigated, avoiding the complicating effects of tissue heterogeneity, in order to facilitate the ability to compute biological processes. The intracellular studies, exemplified in WP6, will be of particular importance in systems biology, for in situ detection of target protein level, modification and localisation, all equally important cornerstones of cellular function which should be addressed in a comprehensive proteome analysis.

Binders and biobanking

Closely linked to biomarker discovery is the application of affinity reagents in the analysis of biobanked samples of plasma, cells and tissues. In order to take full advantage of the extensive and highly valued European biobank resources of human samples - a noteworthy strength of European science - comprehensive sets of reagent and molecular assays based on them are essential. Many European biobank facilities are currently being brought together under the umbrella of the new FP7 European Biobanking and Biomolecular Resources Infrastructure (BBMRI), which recently completed its preparatory phase. In the subsequent programme, BBMRI-LPC (large prospective cohorts) AFFINOMICS is directly connected through Dr Taussig (BI) and Professor Landegren (UU), who act jointly as coordinators of the Biomolecular Resources section, ensuring a close link between the output of AFFINOMICS and European biobank centres. Integration of AFFINOMICS binders with exploration and utilisation of biobanks of sample collections will be another significant area of impact for this project. Potential drug targets and biomarkers can be validated using the binders and insights relevant for personalised medicine can be gathered, such as larger clinical studies with carefully selected patient cohorts available through biobanks.

Databases and access

In order for binder-based profiling of normal and disease tissue to bring the expected major benefits for biomedical research, and hence have the greatest impact in the future, researchers with an interest in protein targets for therapy, imaging and diagnostics in relation to cell signalling and cancer, will have open
access to the AFFINOMICS data. The availability of individual binders will be a matter for discussion between potential users and the producing partners. The bioinformatics initiatives in WP7 will thus have a particular impact in cataloguing and enabling access to the AFFINOMICS binder information through the ProteinBinders (KTH) and Molecular Methods (UU) databases. The ProteinBinders portal provides annotation of all the AFFINOMICS antibodies and recombinant reagents, with complete evaluation data and systematised validation scoring for different assays, so that users can assess for themselves the quality of reagents. During the period of AFFINOMICS, the Molecular Methods database has grown into a valuable web-based resource of protocols, standard procedures and innovative binder-based methods and applications at a level of detail to enable them to be established in user laboratories. Both databases will have provision for user feedback and comment. By providing a central source of information on available binders and their properties and experimental data, ProteinBinders will make a major contribution to the effective use of validated reagents in biological research and we look forward to publications further validating the AFFINOMICS binders in the research community. Through high-level tools identifying the right binders for a given task, the bioinformatics resources will provide both an efficient planning route for researchers and a dissemination platform for new applications and technologies.

International relations
Given the international nature of proteomics research and the wide demand for binding reagents, it is important that Europe should aspire to be at the forefront of production, technical innovation and application of binders. The AFFINOMICS consortium has been the European representative for such efforts and has linked to parallel programmes in the USA. These are the National Cancer Institute (NCI) Clinical Proteomics Reagent Initiative, aiming to produce and distribute sets of monoclonal antibodies as reagents against cancer targets (http://antibodies.cancer.gov/) and the National Institutes of Health (NIH) Common Fund’s Protein Capture Reagents Program, initiated in 2011 (https://commonfund.nih.gov/proteincapture). The latter has aims similar to those of the AFFINOMICS (“generating low cost, high quality, renewable affinity reagents for all human proteins”) and as a test case has focused on production of monocular and recombinant binders against human transcription factors to create a community resource for chromatin immunoprecipitation (ChIP) studies. The choice of targets in both the NIH and NCI initiatives is highly complementary to those in AFFINOMICS and there has been active communication and sharing of expertise and materials across the Atlantic. AFFINOMICS partner KTH has undertaken testing of NIH samples, and NCI antibodies have been characterised at DKFZ. Both the US programmes have entered into Memoranda of Understanding with the AFFINOMICS consortium for the exchange of materials. NIH and NCI personnel have attended AFFINOMICS reporting meetings and made presentations. In addition, Dr Taussig has been a committee member for the NCI programme and on the expert panel of the NIH programme for the last three years and has been able to keep the US side informed about AFFINOMICS objectives and progress. This has given the EU efforts a positive profile in the US.

MAIN DISSEMINATION ACTIVITIES
The external website www.affinomics.org was designed jointly by BBT and UU, and is maintained jointly by BI and UU. The portal is the central entry point to AFFINOMICS, communicating information about the project and progress towards its goals. For external users, it provides access to information on the project, including the WPs, publications and reports on the meetings, including presentations.
The main source of information for the public on the outputs of AFFINOMICS, i.e. the validated protein binder data, is the ProteinBinders portal www.proteinbinders.org. It provides detailed characterisation for antibodies and other protein binding reagents. Data on all antigens and binders generated within AFFINOMICS submitted by the partners and made publicly accessible. The site is continuously updated and maintained by KTH. The portal is arranged in a gene-centric structure allowing grouping of antigens and binders related to one gene product. A more advance interface is also available allowing combined searches for binders targeting for example a certain protein class and with characterisation data from a certain assay. Properties of antigens, binders and characterisation data including producing partner can also be searched for. Version 3 of the portal contains information on 5408 binders, 2900 antigens and 8355 validations. We recently introduced information on the availability of binders and contact details for requests.

The AFFINOMICS project has led to some invigorating meetings where the issues around affinity binders and proteomics are vigorously debated. One series is the biennial Alpbach workshops on Affinity Proteomics, organised and supported by AFFINOMICS, and previously also by the European Science Foundation (ESF). There have been three such workshops during the period of the grant. The programmes comprised two and a half days of talks, with time set aside for outdoor activities. As in previous years, the focus of the event is the current state of affinity methods for proteome analysis. They have covered a wide area, including the Human Protein Atlas and human proteome projects, biomarker discovery, the application of binders in cell signalling research, complementarity of affinity methods to more conventional mass spectrometry proteomics, new technologies, and microfluidics systems. The following is a resume of the last three meetings.

(1) The first workshop (in fact the 5th in the series) was organised during the first reporting period by BBT and was held at the Alpbach Conference Centre, 14-16th March 2011, with co-funding by the ESF programme on functional genomics. Presentations were made by several partners as well as invited speakers. The workshop was attended by 110 participants of whom about 50 were from nonpartner academic organisations and industry. The programme can be viewed on the Meetings and Presentations section of the website at http://www.affinomics.org/.

(2) The second workshop was held from March 11-13th 2013 and was attended by 104 academic and industrial participants from Europe and the USA. In addition to 16 principal invited speakers, there were 18 short talks selected from submitted abstracts and 20 poster presentations. Participation of younger researchers as speakers and poster presenters was encouraged and the meeting provided an excellent opportunity for training in affinity methods. The meeting was again co-funded with the ESF Programme Frontiers of Functional Genomics and details can be found at http://www.affinomics.org/.

(3) The third workshop was held from March 9-11th 2015 and was attended by 120 academic and industrial participants from Europe and the USA. As in previous years, the focus of the event was the current state of affinity methods for proteome analysis, including the complementary approaches of mass spectrometry and affinity methods, the interests of industry focusing on biomarkers, recombinant binder production methods, alternatives to immunoglobulins, proteome targets and diagnostic and clinical applications. There were 22 principal invited speakers, plus 8 short talks selected from submitted abstracts and 20 poster presentations. Papers by speakers at the 2013 and 2015 meetings are published in special issues of the journal New Biotechnology accompanied by an extended meeting report. The programme of the workshop can be found at http://www.affinomics.org/.
A Workshop on laboratory protocol standards for the molecular methods database was held in Uppsala, November 15th, 2011. The remit was to create an ontological data-modeling framework for Laboratory Protocol Standards for the Molecular Methods Database (MolMeth). The workshop provided a set of short- and long-term goals, the most important being the decision to use the established EXACT description of biomedical ontologies as a starting point. It was divided into a series of lectures, and a session to receive input from experts regarding the development of MolMeth and its supporting ontologies. The discussions have been published (Klingstrom T et al., New Biotechnology 2013, 30:109-113)

A training course on Advanced Molecular Technology and Instrumentation for Proteome Analyses was organised jointly between AFFINOMICS, the DiaTools project and the Science for Life Laboratory and held in Uppsala from August 25-September 5, 2014. There were 14 attendees from Sweden, Belgium, Netherlands and Germany. It provided insight into state-of-the-art technologies for proteome analyses in research and medicine via advanced molecular technologies, reagent resources and instrumentation. Lectures and hands-on training conveyed information about established and emerging methods and technologies relevant in basic research and diagnostics. A full report with the programme can be found in deliverable D9.19.

EXPLOITATION OF RESULTS
The major area for exploitation is the availability of the characterised binders, the major output of the AFFINOMICS project, which are listed on the ProteinBinders portal, www.proteinbinders.org as described above. This also provides contact information for the binder producers who are able to decide on an individual basis to make them available and the conditions of access. The portal will continue to be maintained for at least one year after the end of AFFINOMICS at KTH; thereafter other support arrangements may have to be considered.

In regard to commercialisation, 46 monoclonal antibodies against 18 different targets were licensed nonexclusively from EMBL by EMD Millipore, two of which have passed internal quality control and are now being marketed: See: http://www.emdmillipore.com/US/en/product/Anti-3-PGDH-Antibody%2C-clone-TK-08-13A8-,MM_NF-MABS1145 and http://www.emdmillipore.com/US/en/product/Anti-Fgr-Antibody%2C-clone-SuGr-7-14F3-,MM_NF-MABS1140. Further antibodies are expected to be released for marketing in the near future.

There has also been interest from an antibody distributor (Biorbyt) to commercialise the binders; discussions are currently ongoing. There is also interest from Cambridge Protein Arrays Ltd in creating an antibody portfolio based on the AFFINOMICS sets. Here too discussions are in progress.

All monoclonal antibodies have been sent to the Developmental Studies Hybridoma Bank (DSHB) at the University of Iowa, USA, for storage and onward distribution (http://dshb.biology.uiowa.edu/).

List of Websites:

www.affinomics.org (contact: Dr Erik Bongcam Rudloff, erikbong@me.com)
www.proteinbinders.org (contact: Dr Fredric Johansson, fredric.johansson@scilifelab.se)
Related documents

- final1-beneficiaries.pdf
- final1-affinomics-logo.pdf

Last update: 15 January 2016

Permalink: https://cordis.europa.eu/project/id/241481/reporting

European Union, 2023