Validation of affinity reagents using antigen microarrays

Ronald Sjöberg¹, Mårten Sundberg¹, Anna Gundberg, Åsa Sivertsson, Jochen M. Schwenk, Mathias Uhlén and Peter Nilsson

Science for Life Laboratory Stockholm, School of Biotechnology, KTH – Royal Institute of Technology, Box 1031, SE-171 21 Solna, Sweden

Abstract

There is a need for standardised validation of affinity reagents to determine their binding selectivity and specificity. This is of particular importance for systematic efforts that aim to cover the human proteome with different types of binding reagents. One such international program is the SH2-consortium, which was formed to generate a complete set of renewable affinity reagents to the SH2-domain containing human proteins. Here, we describe a microarray strategy to validate various affinity reagents, such as recombinant single-chain antibodies, mouse monoclonal antibodies and antigen-purified polyclonal antibodies using a highly multiplexed approach. An SH2-specific antigen microarray was designed and generated, containing more than 6000 spots displayed by 14 identical subarrays each with 406 antigens, where 105 of them represented SH2-domain containing proteins. Approximately 400 different affinity reagents of various types were analysed on these antigen microarrays carrying antigens of different types. The microarrays revealed not only very detailed specificity profiles for all the binders, but also showed that overlapping target sequences of spotted antigens were detected by off-target interactions. The presented study illustrates the feasibility of using antigen microarrays for integrative, high-throughput validation of various types of binders and antigens.

Introduction

It is of great importance to have access to well-characterised and renewable affinity reagents within the different disciplines of life science. Even though there are several hundred thousand antibodies and other affinity reagents and binding molecules available, many lack a solid validation and this make them less useful for the research community [1–3]. There are two principally different ways to validate affinity reagents, firstly using an application specific assay, such as Western blot, immunofluorescence-based confocal microscopy or immunohistochemistry, aimed to determine the functionality of an affinity reagent in a particular application. Secondly more generic binding assays, such as ELISA or surface plasmon resonance, aimed to determine the binding of the reagent to its antigen, usually performed using the antigen as ligand in the assay. For the latter application, microarrays provide an attractive possibility to perform multiplex assays with many antigens displayed on a single surface making it possible to determine the selectivity and specificity in a background of many related and unrelated antigens.

Recently, an international effort was initiated to generate renewable affinity reagents by various methods to SH2-containing human proteins [4,5]. There are 110 such proteins encoded in the human genome and there is a lack of reagents to many of these proteins. The effort took advantage of several different research groups developing different types of affinity reagents. Antigens corresponding to these proteins were produced either by the Structural Genomics Consortium (SGC) [6] or by the Human Protein Atlas (HPA) project [7,8]. As affinity reagents, antigen-purified polyclonal HPA antibodies were produced in rabbits,
mouse monoclonal antibodies and recombinant single-chain variable fragment (scFv) antibodies were produced in five different academic laboratories.

We have previously used a microarray format with 14 identical subarrays with 384 antigens in each subarray, enabling the simultaneous analysis of 14 binders on each slide. This setup has for several years routinely been used to validate antibodies within the Human Protein Atlas project and more than 29,000 antibodies have been validated since 2005 with an approximate throughput of 150 new antibodies to validate every week. The challenge here has been to broaden the application by combining different antigens and different types of binders with different principles for detection and also in many cases with unknown concentrations.

Here, we have developed a microarray assay for validation of affinity reagents of different sources using a standardised assay involving a fluorescent read-out. The antigen microarray format developed here is a high-throughput and versatile validation tool for the analysis of the specificity of different affinity reagents being profiled with different antigens.

**Material and methods**

**Antigen production**

In total, 64 SH2-domain containing proteins were represented on the antigen microarrays. They were distributed among 85 different HPA antigens or as they are denoted Protein Epitope Signature Tags (PrESTs), which have been designed and expressed as protein fragments from the corresponding SH2-domain containing proteins and produced within the Human Protein Atlas project. These 85 PrESTs represent 53 unique proteins and were originally designed to have as low as possible similarity to other proteins. This means that they have varying overlap, or no overlap, with the SH2-domain of the corresponding SH2-domain containing protein and that more than one PrEST can be derived from one unique protein. Together with these PrESTs were also 20 different SH2-domains expressed and purified at Structural Genomics Consortium (SGC), Karolinska Institute. Eleven of these 20 proteins were not present as PrESTs and 9 proteins were represented by both types of antigens.

The PrESTs used here have all been utilised as antigens for the generation, purification and validation of the monospecific polyclonal antibodies produced in the HPA project. These protein fragments are selected and designed based on regions with low similarity to other human proteins and are approximately 80–100 amino acids in length [9]. The PrESTs are expressed together with an N-terminal hexa-histidine and an albumin binding protein tag (His6ABP), for purification and solubilisation purposes, respectively. The recombinant PrEST proteins are produced in *Escherichia coli* Rosetta DE3 strain [10] and purified with immobilised metal affinity chromatography and subsequently verified with mass spectrometry [11].

The SGC proteins were produced as described previously [12,13]. In short the 20 SH2-domains were expressed in the *E. coli* strain Rosetta 2(DE3)R3 and the purification was done using a two-step protocol including IMAC and gel filtration on an ÄKTA system (GE Healthcare). The proteins were eluted in PBS buffer and the concentration was measured with a NanoDrop spectrophotometer (Thermo Scientific). The batches were quickly frozen in liquid nitrogen and then stored at −80 °C. Mass spectrometry was used to confirm the identities of the proteins and the purity was tested with SDS-PAGE.

**Protein microarray fabrication**

The antigen microarrays were designed to host 432 antigen spots replicated in 14 identical subarrays (Fig. 2a). As solid support slides with epoxy coated glass surface (CapitalBio) were utilised and for the dispensing of the protein solution, that is the actual spotting, a non-contact printer (GeSIM Nanoplotter 2.0E and subsequently also an Arrayjet Marathon) was used.

The SH2 related antigens were complemented with 301 non-SH2 related PrEST representing 172 different proteins encoded on human chromosome 21. In the HPA project there were, at the time of the initiation of the study, 85 PrESTs corresponding to 53 unique SH2-domain containing proteins available and those were all included on the microarray. All of these PrESTs were randomly distributed on the microarray at a concentration of 40 g/ml in 0.1 M urea and 50 mM sodium carbonate-bicarbonate, pH 9.6, complemented with 100 g/ml BSA. The 20 SH2-domains produced by the SGC were positioned in the last two rows of each subarray (Fig. 2a). They were printed at concentrations ranging from 0.4 to 4.0 mg/ml in PBS and also as 1:1 dilutions in the above mentioned sodium carbonate-bicarbonate, buffer. Six of them were also printed in triplicates to utilise the full 432 possible spot positions in the used array design. This resulted in a total of 46 spots from the 20 SH2-domains produced by the SGC.

After printing, the slides were incubated at 37 °C overnight and then blocked for 1 hour in PBST (0.1% Tween 20) supplemented with 3% BSA. They were then washed 2 min × 5 min in PBST followed by 1 min × 5 min in PBS. Finally, slides were washed with deionised water and spun dry and stored dark and dry in 4 °C until use.

**Affinity reagents**

In total 398 different affinity reagents were analysed on the protein array. These reagents were produced at five different laboratories: Human Protein Atlas, Stockholm (HPA), PI: Professor Mathias Uhln, 49 mono-specific polyclonal antibodies [14]; Monash University, Melbourne (MON), PI: Dr. Alan Sawyer, 97 mouse monoclonal antibodies [15]; Beijing Institute of Genomics (BEI), PI: Professor Siqi Liu, 64 mouse monoclonal antibodies; Technische Universität Braunschweig (BRA), PI: Professor Stefan Düb, 30 recombinant scFv [13], University of Cambridge (CAM), PI: Dr. John McCafferty, 90 recombinant scFv [12].

The study of the generated affinity reagents was complemented with the analysis of 68 different external commercially available antibodies targeting SH2-domain containing proteins and that had been validated within the Human Protein Atlas. These antibodies were both monoclonals and polyclonals and of both mouse and rabbit origin. The origin of the commercially available binders included several companies, these being Santa Cruz Biotechnology, Invitrogen, Lab Vision, AbFrontier, Zymed, Epitomics, Cell Signalling Technology, Abcam, Upstate, Strategic Diagnostics and Sigma.

**Experimental microarray procedures**

For the incubation, a mask with 16 incubation chamber wells (GE Healthcare) was used on which up to 16 different reagents can...
be analysed simultaneously on one slide. After incubation with the primary binder the slides were washed and the secondary, and in some cases the tertiary binder, was added in 10 ml solution.

The HPA antibodies were incubated according to the normal procedure in the HPA workflow where approximately 150 new antibodies are analysed every week [14] at a dilution from 1:500 to 1:3000 in PBST for 1 hour. The secondary antibody (anti-rabbit-Alexa647, Invitrogen) was incubated at 1:60,000 dilutions in PBST for 1 hour.

The mouse monoclonal antibodies produced at MON [15] and BEI were incubated at dilutions in PBST ranging from 1:10 to 1:10,000 for 1 hour. The secondary antibody (anti-mouse-Alexa647, Invitrogen) was incubated at a dilution of 1:60,000 in PBST for 1 hour.

The scFvs produced at CAM [12] were incubated at a dilution ranging from 1:1 to 1:3000 in PBST for 1 hour. The secondary antibody (anti-FLAG-biotin) was incubated at a 1:1000 dilution in PBSTM (1% milk and 0.1% Tween 20) for 1 hour and the tertiary detection reagent (Streptavidin-Alexa647) was used at 1:1000 dilutions in PBST.

The scFvs produced at BRA [13] were incubated in a dilution range from 1:1 to 1:100 dilution in PBST for 1 hour. The secondary detection reagent (mouse hybridoma supernatant specific to the myc-tag (9E10) was incubated at 1:1000 in PBSTM (1% milk and 0.1% Tween 20) for 1 hour and the tertiary antibody (anti-mouse-Alexa 647, Invitrogen) was used at 1:1000 dilution in PBST.

The external commercially available antibodies were of both mouse and rabbit origin and they were incubated in a similar manner as the HPA antibodies, and for the mouse antibodies, an anti-mouse-Alexa647 (Invitrogen) was used as secondary antibody.

All the slides were analysed in a dual-colour format because they were co-incubated with a chicken anti-His6ABP antibody together with the primary antibody at a dilution of 1:80,000. The anti-His6ABP antibody was detected with an anti-chicken-Alexa555 (Invitrogen) at a dilution of 1:60,000. This antibody was used as a marker for the presence of the PrESTs on the microarray surface, because they all have a His6ABP tag, and to enable an optimised image analysis (Fig. 2b). The slides were washed 2 min × 5 min in PBST and 1 min × 5 min in PBS between primary and secondary incubations. The slides were scanned with a G2565BA array scanner (Agilent) and the images quantified using the image analysis software GenePix 5.1 (Molecular Devices).

**Results**

**Principle of the assay**

There is a need to establish a platform for validation of affinity reagents from different sources and protein microarrays offer one
such possibility. In Fig. 1, the main types of affinity reagents and antigens are schematically shown centred around the protein microarray platform. By immobilising relatively large numbers of different types of related and non-related antigens on the same array surface, a specificity profiling of antibodies and other binders is enabled. An antigen microarray as outlined in Fig. 2 was designed here to host 432 protein spots replicated in 14 identical subarrays. One microarray slide can thereby be used for the simultaneous incubation and analysis of 14 different binders or conditions. Different primary reagents are combined on each slide such that the same secondary and any potential tertiary reagents can be used throughout the slide.

The human SH2-domain microarray

In this study we designed, printed and utilised microarrays with a focus on SH2-related antigens. In total, 64 unique SH2-domain containing human proteins were represented on the arrays. They were distributed among 85 different antigens representing 53 unique SH2-domain containing proteins that were designed and expressed as protein fragments within the Human Protein Atlas (HPA) project. Included in the last two rows of each subarray were also 20 different SH2-domains expressed and purified at the Structural Genomics Consortium (SGC). To obtain a broader picture of the selectivity and potential off-target interactions from the binders, the SH2 related antigens were complemented with 301 non-SH2 related HPA antigens representing 172 different proteins encoded on human chromosome 21. In total 406 antigens representing 64 unique SH2-domain containing proteins and 172 proteins from the human chromosome 21 were printed in a total of 432 spots. See Fig. 2a for a schematic visualisation of the array layout and Fig. 2b for a picture of the actual spots taken during a quality assurance step to verify the presence of all spotted HPA antigens with the detection of a co-expressed affinity tag.

The protein binders

The various binders analysed here are produced at different academic research centres and constitute in total 398 different SH2-domain containing protein binders. The binders cover mono-specific polyclonal antibodies, mouse monoclonal antibodies and recombinant single chain variable fragments. All these binders have been profiled on 406 antigens and including replicated
experiments were 260,000 data points of antigen-antibody interactions generated and analysed.

Each set of binders required their own optimisations to establish suitable dilutions and concentrations. Briefly, the working dilution range of a set of binders was initially tested on a subset of those binders on one slide based on information provided by the reagent producer, such as dilutions used in other assays, the concentration of the binders, or based on the HPA validation protocol if no prior knowledge of the binders existed. The final working dilutions for the whole set of binders were then decided depending on whether the signals from the test-set were useful or too high or low. No fine tuning of individual binder concentrations were done. The levels of the resulting fluorescent signals varying from binder to binder but the absolute intensities are less important because it is only the relation between the on-target and off-target signals for each binder that is relevant. That is also why it is most appropriate to visualise the resulting array signals in a relative scale, which is done for each affinity reagent separately where the fluorescent intensities are plotted in a histogram to visualise the specificity profile. A dual-colour format was used here, where one colour channel indicates the presence of all immobilised antigens by using a common tag. The main advantage with the dual-colour assay format is that it enables all spots to be visualised and thereby improves the reliability in the image analysis. It also works as a positive control that the antigen is actually present on the surface, which is useful when there is no specific signal from individual binders.

The binders analysed were of diverse origin as some binders, like the HPA polyclonal antibodies, were already purified and well characterised and other binders were supplied with unknown properties and composition. This coupled with the fact that the antigens consisted of two different sets of antigens made it difficult to apply a consistent set of criteria for classification of a binder as validated. In short a binder would be considered to be a validated binder if the majority of its binding is towards an antigen from the correct target.

Despite these complexities, highly specific affinity reagents of all analysed types and from all providers were obtained and all antigens utilised as targets generated validated binders [4]. Approximately ninety percent of all tested binders had specificity profiles where the highest signal was from their intended target while around ten percent showed specificity profiles with little to no binding to its intended target but high signals from other antigens. The microarrays revealed not only very detailed and informative specificity profiles for all the binders, but also illustrated how sequence similarity and overlap in amino acid sequence influence the off-target interactions. In the following, we describe a few illustrative examples of the results from the microarray assay. In Fig. 3 we illustrate how two polyclonal binders against different parts of the same target protein can generate different profiles of specificity and how a binder towards a part of the target protein that does not contain a SH2-domain still can give higher unspecific signals from other SH2-domains, although in this case the higher unspecific binding was also accompanied by a stronger specific signal. In Fig. 4 we show two examples of how a polyclonal binder against a longer peptide also shows specific recognition of shorter peptides from the same target protein. They also show how higher antigen concentration can give rise to higher signal intensity from shorter antigens. Figure 5 shows how off target binding can occur between peptides with sequence similarities and how the off target signals often are of lower intensity compared to on target signal. Figure 6 illustrates how monoclonals with different amounts of overlap towards other peptides can generate

**Validation of polyclonal antibodies towards human protein PLCG2**

In the upper histogram in Fig. 3, a polyclonal antibody produced towards one PLCG2-PrEST displayed strong relative binding to its target antigen, in this case PrEST number 206 on the microarray. This PrEST consists of an 89 amino acid long part of the target protein between amino acids 912–1000, and it does not overlap any of the two SH2-domains present in the full-length protein. In the lower histogram, a second antibody produced against another PrEST from the same full-length protein show equally strong relative binding to its PrEST, number 208, on the array. This PrEST is 112 amino acid long and located between amino acid 425–536 on the full-length protein and overlaps one of the SH2-domains by four amino acids. Both antibodies were analysed at the same dilution and had the same original concentration, but show different amounts of background binding to the 20 SGC produced SH2-domains. The antibody showing a higher unspecific binding also showed a specific signal, almost twice as strong as the antibody with lower unspecific signals.

**Validation of polyclonal antibodies towards human proteins SH2D1A and GRB2**

The two histograms in Fig. 4 show two different antibodies raised against two different PrESTs that fully overlap the SH2-domains on their respective full-length protein. Both antibodies show binding to their specific PrEST as well as to the corresponding SGC produced SH2-domain and limited binding to other PrESTs and SGC SH2-domains in the array. This shows that the antibodies can recognise both the actual antigen, that is the PrEST generated as a fusion protein as well as the SH2-domain if there is sufficient overlap.

**Validation of polyclonal antibodies towards human proteins VAV1 and BLNK**

As can be seen in the upper histogram in Fig. 5, an antibody raised against a PrEST with no overlap with the corresponding SGC SH2-domain nevertheless recognised a SH2-domain corresponding to a different SH2-protein than that which the target PrEST was derived from, in this case the SGC RASA1 SH2-domain. This could be a result of an antibody that requires an epitope of very short length or even just a few individual amino acids and in this case the PrEST and the SGC RASA1 SH2-domain show a 66% sequence similarity match over a 15 amino acid long stretch. The lower histogram in Fig. 5 shows an antibody raised against a PrEST that overlaps fully with the SH2-domain of the protein that the PrEST is derived from, in this case BLNK. The antibody recognises the PrEST and the SGC SH2-domain derived from another protein, the SYK-protein. The BLNK-PrEST and the SGC SYK SH2-domain show a sequence similarity of 58% over a 34 amino acid long stretch. This shows once more that off-target binding can occur towards proteins with similar sequence but that there is a difference in the strength of the resulting signals.
Validation of monoclonal antibodies towards human proteins SLA, ZAP70 and GRB2

In histogram 6, the first diagram shows a monoclonal antibody raised against a PrEST derived from the SLA-protein, but without overlap of the SH2-domain, binding specifically to its target. The second histogram in Fig. 6 shows a monoclonal antibody raised against the PrEST derived from the ZAP70-protein binding to the corresponding SGC ZAP70 SH2-domain spot in the array. The PrEST and the SGC ZAP70 SH2-domain overlap with 22 amino acids which is enough in this case to give a clear signal from the SGC SH2-domain. The ZAP70-PrEST was absent from the array. In the third histogram of Fig. 6 a monoclonal antibody raised against the SGC GRB2 SH2-domain binds to both the SGC SH2-domain and to two PrESTs derived from the GRB2-protein that overlap the SGC SH2-domain.

Validation of monoclonal antibodies and recombinant antibody fragments towards human proteins ABL1, GRB2 and VAV1

The first histogram in Fig. 7 illustrates one monoclonal antibody raised against the SGC ABL1-SH2-domain that manages to discriminate between ABL1 and ABL2 despite the 89% sequence overlap between the two SH2-domains. In the second diagram in Fig. 7, an

![Validation of monoclonal antibodies towards human proteins SH2D1A and GRB2. See legend to Fig. 3.](image-url)
scFv raised against the SGC GRB2 SH2-domain recognises not only its target SH2-domain but also the two GRB2-PrESTs which both completely overlap the SGC SH2-domain; this is similar to what can be seen in the third histogram of Fig. 6. Finally the third histogram in Fig. 7 shows a binder towards the SGC VAV1 SH2-domain and which also recognises one of the three PrESTs generated from VAV3. This particular PrEST and the SGC VAV1 SH2-domain show a 90% sequence identity match over an 11 amino acid long stretch.

**Discussion**

We have shown here the potential of antigen microarrays as an informative and important validation tool to analyse a wide variety of affinity reagents. To illustrate the feasibility and potential of the antigen microarray based platform, we have utilised binders from the international SH2-consortium formed to demonstrate the potential in systematic and high-throughput generation of antigens and corresponding affinity reagents [4,5]. This geographically well dispersed set of academic research groups formed a constellation which has been denoted the Renewable Protein Binder Working Group [4,12,13]. Several types and numbers of different protein binders were generated; antigen-purified polyclonal antibodies were produced in rabbits, mouse monoclonal antibodies and recombinant single chain variable fragments. To be able to thoroughly evaluate and validate the specificity and selectivity of a larger number of reagents, a high-throughput platform suitable for the simultaneous analysis of different types of protein binders is not only attractive, but essentially required. The total number of affinity

**FIGURE 5**

Validation of polyclonal antibodies towards human proteins VAV1 and BLNK. See legend to Fig. 3.
reagents incubated on these arrays was over 600, which includes replicated measurements, multiple assay conditions as well as concentration ranges evaluated for the 398 binding molecules of different classes.

Here, microarray printing equipment allowing 432 spots per subarray was used. It is possible to use higher density spotting to allow one thousand or more spotted antigens per subarray. Furthermore, by increasing the number of subarrays from 16 to 24, together with the use of a slide holder and a gasket with four slide positions, a 96 microplate formatted microarray setup is possible. That allows for the simultaneous analysis of 96 affinity reagents towards a thousand antigens in an automated manner. Therefore, a protocol for automated and normalised dilutions and transfer of antibodies to array slides using liquid handling devices has been set up. This improves our ability to perform large numbers of validation assays for example 192 affinity reagents per day per person including scanning, but excluding image analysis and data processing.

The two proteins ABL1 and ABL2 are very similar in their protein sequences. They share 89% identity [12] and therefore it was interesting that one of the tested monoclonal antibodies discriminates between these two proteins even though they were similar in sequence. Although only one of the binders could distinguish between the two proteins it still showed how informative a tool protein microarrays can be to establish well validated affinity reagents.

There are no commonly agreed standards for antibody validation and there are no proposed guidelines on what can be defined as a specific binding reagent to be sold on the antibody assay market today [3]. Therefore, it is in generally up to the individual researcher to validate any commercial antibody for its planned use [16]. The aim should be to have common and generally accepted specificity assessment strategy for antibody validation. The antibodypedia.org [17] site is one example of where all available validation data can be collected and is publicly available to all potential users. This is one important step in the direction of better validation of affinity reagents.

There are many different procedures or technologies that can be used to validate an antibody or other affinity reagents. The planar antigen microarrays used here as a validation tool to investigate the specificity of affinity reagents are only one of many specific application driven validation technologies to use. Another highly multiplex format for immobilising antigens on a solid support to

---

**FIGURE 6**
Validation of monoclonal antibodies towards human proteins SLA, ZAP70 and GRB2. See legend to Fig. 3.

**FIGURE 7**
Validation of monoclonal antibodies and recombinant antibody fragments towards human proteins ABL1, GRB2 and VAV1. See legend to Fig. 3.
measure specificity is for example a comparable system, which is on colour coded particles and flow cytometry [18].

In conclusion, the presented antigen microarray demonstrates an integrated analysis of binding reagents of different types using a single layout. The study shows the potential of the method to become an important – if not essential – component in future studies aiming at covering the human proteome with antibodies. With future technological advances and a growing number of several different proteins and binding reagents becoming available, arrays of even greater density and diversity can be created to expand selectivity analysis into arrays hosting peptides, protein fragments, recombinant or purified full-length proteins and proteins with post translational modifications and splice variation.

**Acknowledgements**

We are thankful for the collaborations with all participants within the SH2-consortium and especially here the providers of antigens and binding reagents: Dr. Susanne Gräslund et al., Structural Genomics Consortium; Dr. Alan Sawyer et al., Monash University; Professor Siqi Liu et al., Beijing Institute of Genomics; Professor Stefan Dübel et al., Technische Universität Braunschweig; Dr. John McCafferty et al., University of Cambridge.

This study was supported by AFFINOMICS, an EC FP7 Collaborative Project, by the ProNova VINN Excellence Centre for Protein Technology (VINNOVA, Swedish Governmental Agency for Innovation Systems) and by the Knut and Alice Wallenberg Foundation. Marie Berg and Ida Hossar are acknowledged for fruitful discussions.

**References**

10 Tegel, H. et al. (2010) Increased levels of recombinant human proteins with the Escherichia coli strain Rosetta(DE3). Protein Exp. Purif. 69, 159–167