A roadmap to generate renewable protein binders to the human proteome

Karen Colwill[1], Renewable Protein Binder Working Group[2] & Susanne Gräslund[3]

Despite the wealth of commercially available antibodies to human proteins, research is often hindered by their inconsistent validation, their poor performance and the inadequate coverage of the proteome. These issues could be addressed by systematic, genome-wide efforts to generate and validate renewable protein binders. We report a multicenter study to assess the potential of hybridoma and phage-display technologies in a coordinated large-scale antibody generation and validation effort. We produced over 1,000 antibodies targeting 20 SH2 domain proteins and evaluated them for potency and specificity by enzyme-linked immunosorbent assay (ELISA), protein microarray and surface plasmon resonance (SPR). We also tested selected antibodies in immunoprecipitation, immunoblotting and immunofluorescence assays. Our results show that high-affinity, high-specificity renewable antibodies generated by different technologies can be produced quickly and efficiently. We believe that this work serves as a foundation and template for future larger-scale studies to create renewable protein binders.

Antibodies are key research tools in biomedical research. Although there are more than 500,000 commercially available antibodies, scientists widely acknowledge difficulties in accessing high-quality antibodies for their research[1,2]. This has spurred efforts to systematically generate and characterize high-quality antibodies, typically polyclonals[3–5]. But polyclonal antibody supplies are finite, and each new batch is unique, putting extra burdens on quality-assurance systems. Efforts to produce polyclonal antibodies must be complemented by efforts to produce high-affinity, selective and renewable antibodies or antibody-like reagents.

Two general methods to develop renewable antibodies, hybridoma and recombinant display technologies, have been previously applied to large protein sets[6–8]. These studies revealed that reagent generation on a proteome scale requires a streamlined selection process[9–11] and highly purified, correctly folded antigens[12]. Large-scale structural genomics efforts now enable the production of tens of thousands of proteins or protein domains in milligram quantities. The availability of these potential antigens, together with the increasing efficiency of recombinant antibody engineering methodologies and monoclonal antibody production, prompted us to again explore the prospects of systematically developing high-quality antibodies to human proteins[13]. In this pilot study, we selected SH2 domains as prototype antigens because they are of broad interest to the scientific community[14,15]. Additionally, these domains are autonomously folded, are stable and can be produced in milligram quantities. They are also a challenging test set, owing to their high degree of sequence and structural similarities.

We produced 20 SH2 domain proteins and distributed the antigens to researchers in five laboratories for antibody generation. One group generated monoclonal antibodies by hybridoma technology using a high-throughput robotic approach. The other groups generated recombinant Fab or single-chain Fv (scFv) reagents using phage display[16,17]. Here we summarize the results of this project, compare and contrast the different experimental approaches, and demonstrate the application of these reagents in relevant immunological assays. Our results show that specific, high-affinity renewable antibodies can be produced quickly and efficiently. We believe our results demonstrate the potential for renewable antibody generation within a larger-scale project.

RESULTS

Generation of renewable protein binders

We selected 20 SH2 domains from a library of 130 SH2 domain expression constructs, then purified these antigens to >95% homogeneity[17]. We selected these SH2 domains for their biological relevance and because they represent diverse sequences in the SH2 family and include close family members (for example, Abl1 and Abl2). We distributed these antigens to five groups to produce either monoclonal antibodies via immunization or recombinant antibody fragments by phage display (scFvs or Fabs) to the folded domains.

The study workflow is summarized in **Figure 1**. To achieve the required throughput with hybridoma technology, we used automated robotic systems to facilitate the tissue-culture workload and an antigen microarray assay (AMA) to screen hybridoma supernatants for antibodies that bind to the purified antigen ('positive'

[1]Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada. [2]A full list of authors appears at the end of this paper. [3]Structural Genomics Consortium, Department of Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden. Correspondence should be addressed to S.G. (susanne.graslund@ki.se) or K.C. (colwill@lunenfeld.ca).

RECEIVED 21 MAY 2010; ACCEPTED 11 APRIL 2011; PUBLISHED ONLINE 15 MAY 2011; DOI:10.1038/NMETH.1607

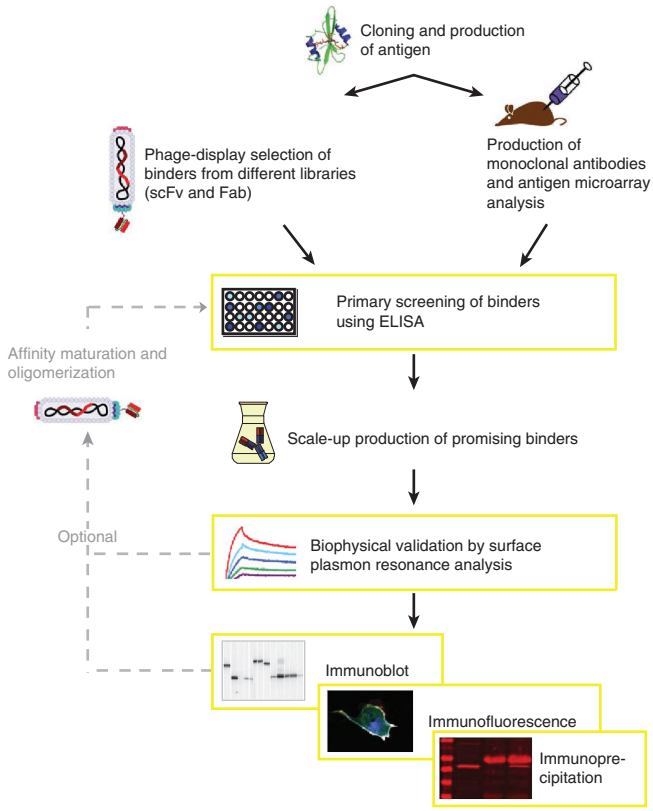

Figure 1 | Flowchart showing methodologies used to systematically produce and validate renewable antibodies. Validation steps are outlined in yellow.

antibodies)[8]. AMA allows IgG-type antibodies to be specifically selected and antibody specificity determined on multiple targets. Using AMA, we screened 34,800 culture supernatants, containing ~11,500 hybridomas. For the recombinant binders, we selected both scFvs and Fabs from libraries with repertoires containing $>10^9$ members, using standard methods. For scFv screening, we used two human natural libraries: HAL4/7 (refs. 16,18) and the 'McCafferty library'[7]. For Fab screening, we used synthetic libraries E and F, which are the fifth and sixth generation of our synthetic library (unpublished data; H.P. and S.S.).

Primary characterization of affinity and selectivity

Both hybridoma and recombinant technologies generated more clones than were feasible for downstream processing. Therefore, we established criteria to prioritize the most promising clones from each step (**Fig. 1**). For hybridomas, we validated by enzyme-linked immunosorbent assay (ELISA) up to 30 AMA hits per antigen. Of these, 165 clones passed ELISA validation (**Table 1** and **Supplementary Table 1**). For the phage display technologies, of 6,972 clones screened from the initial panning, 1,788 (26%) were positive by ELISA (**Table 1** and **Supplementary Table 1**). Of the 1,788 clones, we identified at least 340 unique clones as determined by restriction enzyme digestion or sequencing. For recombinant binders, the average hit rate among the different groups was 14–34%, and across all antigens the average hit rate was 11–46%. Certain antigens worked better in some laboratories than in others, possibly reflecting differences in the available library diversities. Together, the hybridoma and recombinant approaches yielded at least ten unique binders for each target.

We analyzed the selectivity of binding for the cognate SH2 domain over other SH2 domains in two different screens. First, we screened a subset of primary positive scFv clones against all other SH2 domains using ELISA. Of 695 clones, 379 (55%) were specific for the target and did not bind any of the other 19 SH2 domains[17] (**Table 1**). Second, we screened a subset of scFvs and hybridoma supernatants on protein microarrays that contained duplicates of the 20 folded SH2 domains and 386 different protein fragments (protein epitope signature tags)[16,17] (**Supplementary Fig. 1**). In this assay format, we scored binders as 'supportive' if the signal for their cognate targets was ten times above the off-target background (**Supplementary Table 2**). Initially, the hybridoma-generated binders appeared to be more selective in this assay format (unpublished data; P.N. and M.U.) but optimizing conditions for the recombinant binders, mainly in terms of concentration, dramatically reduced or eliminated this difference (data not shown).

Large-scale purification of renewable protein binders

We purified monoclonal antibodies using protein G and obtained an average yield of 23 mg l^{-1} of hybridoma supernatant (**Supplementary Table 1**). As we built synthetic Fab libraries on a single variable heavy (V_H) chain framework that binds protein A, we purified Fabs on protein A–sepharose beads, giving an average

Table 1 | Renewable antibodies targeting SH2 domains

Antigen	Dübel scFv[a] (unique)[b]	McCafferty scFv[c]	McCafferty scFv: specificity[d] (unique)[b]	Hybridoma[e]	Fab[f] (unique)[b]	Total (unique)[b]
ABL1	5 (4)	90	14 (14)	5	5 (4)	29 (22)
ABL2	22 (4)	9	2 (1)	1	8 (7)	33 (12)
BCAR3	12 (3)	55	18 (8)	3	8 (8)	41 (19)
BTK	70 (4)	11	12 (6)	3	5 (5)	90 (15)
CRK	40 (2)	46	28 (2)	6	11 (8)	85 (12)
FYN	8 (3)	61	25 (12)	0	6 (3)	39 (18)
GRAP2	44 (2)	21	1 (1)	14	12 (11)	71 (14)
GRB2	7 (5)	82	4 (2)	20	12 (6)	43 (13)
LCK	20 (3)	99	33 (12)	16	5 (5)	74 (20)
LYN	20 (4)	106	43 (14)	3	10 (6)	76 (24)
NCK1	9 (3)	97	40 (6)	8	3 (3)	60 (12)
PIK3R1 C[g]	14 (5)	60	1 (1)	3	1 (1)	19 (7)
PLCG1 C	10 (7)	105	1 (1)	0	3 (3)	14 (11)
PTPN11 C	6 (6)	120	40 (17)	7	5 (5)	58 (28)
RASA1 C	4 (3)	23	11 (6)	6	0 (0)	21 (9)
SH2D1A	10 (6)	85	14 (5)	13	8 (8)	45 (19)
SHC1	12 (10)	62	42 (13)	7	8 (4)	69 (27)
SYK N[h]	15 (4)	42	1 (1)	6	3 (3)	25 (8)
VAV1	29 (10)	77	20 (7)	15	5 (5)	69 (22)
ZAP70 tandem	5 (3)	41	29 (19)	29	16 (6)	79 (28)
Total	362 (91)	1,292	379 (148)	165	134 (101)	1,040 (340)

The number of ELISA-validated affinity reagents that were produced for the SH2 domains from each laboratory is shown.
[a]We screened 2,668 clones (92 for 14 targets, 276 for FYN, BTK, GRAP2, 184 for GRB2, VAV1, ZAP70) by ELISA during primary validation[16]. [b]Unique clones, where identified. [c]We screened 190 clones from each target (3,800 in total) by ELISA during primary validation and 1,292 positive clones were identified[17]. [d]We screened 695 of the 1,992 positive clones from the McCafferty lab for cross-reactivity to the other SH2 domains and the numbers of monospecific hits (signal for respective targets >tenfold over other SH2 domains) are listed. Sequence diversity for the monospecific hits is indicated in parentheses. [e]For these 20 antigens, 38,400 supernatants in total were screened by AMA. Not all of the wells contained viable hybridomas. Selected AMA positives were subsequently screened by ELISA during primary validation. [f]We screened 504 clones (24 clones for 18 targets, 36 clones for GRB2 and GRAP2) by ELISA during primary validation in the Sidhu and Koide laboratories. [g]C, C-terminal SH2 domain. [h]N, N-terminal SH2 domain.

Table 2 | SPR analysis and biological validation of 73 renewable antibodies

Target[a]	Binder name[b]	Alternate name	Affinity maturation	k_a (M^{-1} s^{-1})	k_d (s^{-1})	K_D (M)	Experimental validation
ABL1	ABL1_JM_scFv_1	JM_043_1H10	Naive	4.84×10^{4}	3.63×10^{-4}	7.51×10^{-9}	
	ABL1_JM_scFv_2	JM_043_1G05	Naive	2.15×10^{4}	1.48×10^{-3}	6.86×10^{-8}	
ABL2	ABL2_JM_scFv_1	JM_061_2E04	Naive	2.10×10^{4}	5.44×10^{-5}	2.59×10^{-9}	
BCAR3	BCAR3_SS_Fab_9	BCAR3-9	Matured	3.31×10^{4}	1.96×10^{-3}	5.94×10^{-8}	
BTK	BTK_JM_scFv_1	JM_060_2D01	Naive	7.99×10^{5}	1.65×10^{-3}	2.07×10^{-9}	
CRK	CRK_JM_scFv_1	JM_070_1E02	Matured	1.78×10^{4}	4.65×10^{-4}	2.61×10^{-8}	
	CRK_JM_scFv_2	JM_070_1H05	Matured	1.32×10^{4}	7.12×10^{-4}	5.41×10^{-8}	
	CRK_JM_scFv_3	JM_070_1E08	Matured	2.15×10^{4}	5.20×10^{-4}	2.42×10^{-8}	
	CRK_SD_scFv_1	dm140-K-H2	Naive	9.11×10^{4}	3.00×10^{-3}	3.30×10^{-8}	Failed ectopic protein IP
	CRK_SD_scFv_2	dm140-K-H10	Naive	6.72×10^{5}	3.00×10^{-2}	4.84×10^{-8}	Ectopic protein IP only
	CRK_SS_Fab_9[c]	CRK-9	Matured	4.21×10^{4}	1.04×10^{-4}	2.47×10^{-9}	Endogenous protein IP
	CRK_SS_Fab_10	CRK-10	Matured	7.13×10^{3}	4.78×10^{-4}	6.71×10^{-8}	
	CRK_SS_Fab_11	CRK-11	Matured	2.80×10^{4}	2.59×10^{-4}	9.25×10^{-9}	Failed ectopic protein IP
FYN	FYN_SS_Fab_2	FYN-2	Naive	4.50×10^{5}	9.13×10^{-4}	2.03×10^{-9}	
GRAP2	GRAP2_SK_Fab_1		Naive	1.36×10^{5}	2.28×10^{-3}	1.68×10^{-8}	
	GRAP2_SK_Fab_2		Naive	8.28×10^{4}	6.06×10^{-3}	7.32×10^{-8}	
	GRAP2_SK_Fab_3		Naive	9.31×10^{4}	1.07×10^{-3}	1.15×10^{-8}	
GRB2	GRB2_AS_mAB_3	GRB2_11H3	Naive	6.72×10^{5}	3.30×10^{-4}	4.92×10^{-10}	Failed ectopic protein IP
	GRB2_JM_scFv_1	JM_047_2A08	Naive	1.21×10^{4}	2.68×10^{-3}	2.22×10^{-7}	
	GRB2_JM_scFv_2	JM_047_2A02	Naive	1.67×10^{4}	3.86×10^{-3}	2.30×10^{-7}	
	GRB2_JM_scFv_3	JM_047_2A10	Naive	9.72×10^{4}	7.91×10^{-4}	8.13×10^{-9}	Failed ectopic protein IP
	GRB2_JM_scFv_4	JM_047_2B02	Naive	9.78×10^{3}	2.27×10^{-3}	2.32×10^{-7}	
	GRB2_SK_Fab_1		Naive	1.17×10^{5}	3.00×10^{-2}	2.96×10^{-7}	
	GRB2_SK_Fab_2		Naive	2.42×10^{5}	1.45×10^{-3}	5.99×10^{-9}	Ectopic protein IP only
	GRB2_SK_Fab_3[c]		Naive	1.88×10^{5}	2.14×10^{-3}	1.14×10^{-8}	Endogenous protein IP
	GRB2_SS_Fab_2	GRB2-2	Naive	5.65×10^{4}	4.08×10^{-3}	7.22×10^{-8}	
	GRB2_SS_Fab_4	GRB2-4	Matured	7.14×10^{4}	2.25×10^{-3}	3.16×10^{-8}	
	GRB2_SS_Fab_5	GRB2-5	Matured	9.84×10^{4}	2.78×10^{-3}	2.83×10^{-8}	
	GRB2_SS_Fab_6	GRB2-6	Matured	1.96×10^{5}	8.43×10^{-3}	4.29×10^{-8}	
LCK	LCK_SS_Fab_6	LCK-6	Matured	1.16×10^{4}	1.82×10^{-3}	1.57×10^{-7}	
LYN	LYN_JM_scFv_1	JM_065_1E12	Matured	2.53×10^{5}	2.27×10^{-3}	8.98×10^{-9}	Failed ectopic protein IP
	LYN_JM_scFv_2	JM_065_1E07	Matured	2.62×10^{5}	2.66×10^{-3}	1.01×10^{-8}	Failed ectopic protein IP
	LYN_JM_scFv_3	JM_065_1E03	Matured	3.12×10^{5}	2.19×10^{-3}	7.02×10^{-9}	Failed ectopic protein IP
	LYN_SD_scFv_1	dm130-F-A9	Naive	9.30×10^{4}	5.72×10^{-4}	6.15×10^{-9}	Ectopic protein western
	LYN_SS_Fab_1	AP03-43j	Naive	1.91×10^{5}	1.46×10^{-3}	7.61×10^{-9}	
	LYN_SS_Fab_2	LYN-4	Naive	2.54×10^{5}	2.27×10^{-3}	8.95×10^{-9}	Ectopic protein IP and western
	LYN_SS_Fab_3	LYN-7	Matured	1.23×10^{5}	3.96×10^{-3}	3.22×10^{-8}	
NCK1	NCK1_JM_scFv_1	JM_067_1E03	Matured	3.52×10^{5}	8.97×10^{-3}	2.55×10^{-8}	Failed endogenous protein IP
	NCK1_JM_scFv_2	JM_075_1B09	Matured	1.74×10^{5}	3.00×10^{-2}	1.47×10^{-7}	Failed endogenous protein IP
	NCK1_JM_scFv_3	JM_067_1F09	Matured	1.25×10^{5}	4.04×10^{-3}	3.22×10^{-8}	Failed endogenous protein IP
PIK3R1 C	PIK3R1C_SS_Fab_2	PIK3R1C-2	Matured	9.76×10^{4}	4.46×10^{-3}	4.56×10^{-8}	
	PIK3R1C_SS_Fab_3	PIK3R1C-3	Matured	2.91×10^{4}	6.84×10^{-3}	2.35×10^{-7}	
PLCG1 C	PLCG1C_SS_Fab_4	PLCG1C-4	Matured	9.22×10^{4}	2.82×10^{-3}	3.06×10^{-8}	
	PLCG1C_SS_Fab_5	PLCG1C-5	Matured	5.71×10^{4}	2.38×10^{-3}	4.17×10^{-8}	
	PLCG1C_SS_Fab_6	PLCG1C-6	Matured	2.14×10^{5}	1.90×10^{-3}	8.84×10^{-9}	
PTPN11 C	PTPN11C_JM_scFv_1	JM_069_1C02	Matured	8.83×10^{5}	9.14×10^{-3}	1.03×10^{-8}	
	PTPN11C_JM_scFv_2	JM_069_1A11	Matured	1.46×10^{5}	1.21×10^{-3}	8.33×10^{-9}	
	PTPN11C_SS_Fab_5	PTPN11C-5	Naive	1.80×10^{4}	8.91×10^{-3}	4.96×10^{-7}	
RASA1 C	RASA1C_JM_scFv_1	JM_056_2A03	Naive	4.06×10^{5}	2.23×10^{-3}	5.48×10^{-9}	
	RASA1C_JM_scFv_2	JM_056_2A10	Naive	3.70×10^{4}	4.20×10^{-3}	1.14×10^{-7}	
	RASA1C_SD_scFv_1	dm124-Q-H1	Naive	1.60×10^{6}	4.56×10^{-4}	2.85×10^{-10}	Failed endogenous protein IP
	RASA1C_SD_scFv_2	dm124-Q-H8	Naive	5.08×10^{4}	0.13	2.61×10^{-6}	
SH2D1A	SH2D1A_JM_scFv_1	JM_051_2B04	Naive	2.72×10^{5}	3.20×10^{-3}	1.18×10^{-8}	
	SH2D1A_JM_scFv_2	JM_051_2B09	Naive	6.21×10^{3}	4.40×10^{-3}	7.08×10^{-7}	
	SH2D1A_SS_Fab_4	SH2D1A-4	Naive	2.79×10^{5}	4.00×10^{-2}	1.56×10^{-7}	
	SH2D1A_SS_Fab_9	SH2D1A-9	Matured	2.58×10^{5}	2.00×10^{-2}	8.48×10^{-8}	
SHC1	SHC1_AS_mAb_1	SHC1_9E11	Naive	2.32×10^{5}	1.74×10^{-4}	7.49×10^{-10}	Endogenous protein IP
	SHC1_AS_mAb_2	SHC1_2H6	Naive	4.16×10^{5}	1.35×10^{-3}	3.25×10^{-9}	Endogenous protein IP
	SHC1_JM_scFv_1[c]	JM_072_1A10	Matured	1.21×10^{5}	5.82×10^{-4}	4.80×10^{-9}	Endogenous Protein IP
	SHC1_JM_scFv_2	JM_072_1A01	Matured	3.53×10^{4}	3.04×10^{-3}	8.61×10^{-8}	
	SHC1_JM_scFv_3	JM_072_1C04	Matured	2.86×10^{4}	9.28×10^{-4}	3.24×10^{-8}	
	SHC1_SD_scFv_2	dm122-G-B6	Naive	7.49×10^{4}	3.82×10^{-4}	5.10×10^{-9}	Failed endogenous protein IP
	SHC1_SD_scFv_3	dm122-G-B9	Naive	9.60×10^{4}	1.79×10^{-3}	1.86×10^{-8}	
	SHC1_SD_scFv_4	dm122-G-C9	Naive	9.64×10^{4}	4.39×10^{-4}	4.56×10^{-9}	Failed endogenous protein IP
SYK	SYK_JM_scFv_1	JM_054_2A04	Naive	4.86×10^{3}	3.00×10^{-2}	5.58×10^{-6}	
VAV1	VAV1_JM_scFv_1	JM_066_1E03	Matured	6.59×10^{4}	1.59×10^{-3}	2.42×10^{-8}	
	VAV1_JM_scFv_2	JM_066_1E09	Matured	8.03×10^{4}	1.76×10^{-3}	2.19×10^{-8}	
	VAV1_JM_scFv_3	JM_066_1H05	Matured	5.39×10^{4}	2.26×10^{-3}	4.20×10^{-8}	
ZAP70	ZAP70_SS_Fab_1	AP03-17t	Naive	1.48×10^{4}	2.33×10^{-3}	1.58×10^{-7}	
	ZAP70_SS_Fab_2	AP03-17ab	Matured	4.29×10^{4}	1.90×10^{-3}	4.42×10^{-8}	
	ZAP70_SS_Fab_7	AP03-17cd	Matured	6.27×10^{4}	2.79×10^{-3}	4.44×10^{-8}	
	ZAP70_SS_Fab_8	AP03-17ef	Matured	1.00×10^{4}	2.72×10^{-3}	2.70×10^{-7}	

k_a, on rate; k_d, off rate; and IP, immunoprecipitation.

[a]For a summary of tested clones, see **Supplementary Table 1**. [b]Standardized binder name: SH2 domain target_principal investigator initials_Binder type_number. [c]Sequences in FASTA format are available in **Supplementary Figure 4**.

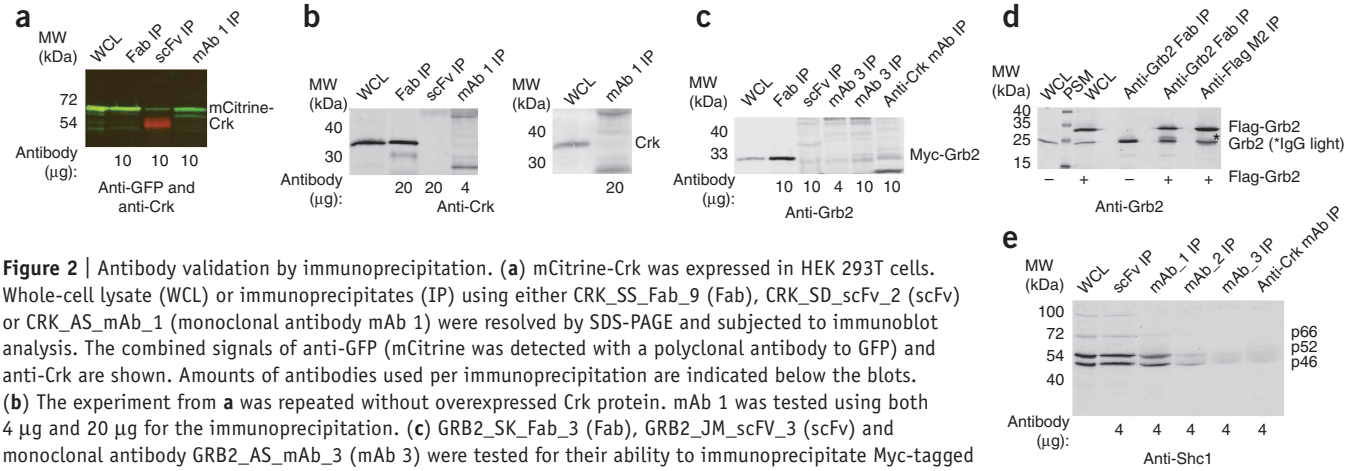

Figure 2 | Antibody validation by immunoprecipitation. (**a**) mCitrine-Crk was expressed in HEK 293T cells. Whole-cell lysate (WCL) or immunoprecipitates (IP) using either CRK_SS_Fab_9 (Fab), CRK_SD_scFv_2 (scFv) or CRK_AS_mAb_1 (monoclonal antibody mAb 1) were resolved by SDS-PAGE and subjected to immunoblot analysis. The combined signals of anti-GFP (mCitrine was detected with a polyclonal antibody to GFP) and anti-Crk are shown. Amounts of antibodies used per immunoprecipitation are indicated below the blots. (**b**) The experiment from **a** was repeated without overexpressed Crk protein. mAb 1 was tested using both 4 μg and 20 μg for the immunoprecipitation. (**c**) GRB2_SK_Fab_3 (Fab), GRB2_JM_scFV_3 (scFv) and monoclonal antibody GRB2_AS_mAb_3 (mAb 3) were tested for their ability to immunoprecipitate Myc-tagged Grb2 expressed in HEK 293T cells. (**d**) HEK 293T WCL or IP using Fab GRB2_SK_Fab_3 resolved by SDS-PAGE and subjected to immunoblot analysis. The presence of ectopic Flag-tagged Grb2 (Flag-Grb2) is indicated below the blot. *, light chain from the anti-Flag M2. (**e**) scFv SHC1_JM_scFv_1 and monoclonal antibodies (mAb 1–3) SHC1_AS_mAb_1, SHC1_AS_mAb_2 and SHC1_AS_mAb_3 were tested for their ability to recognize endogenous Shc1 in HEK 293T cells. The amount of antibody used in each immunoprecipitate is indicated below the gel. Three isoforms of Shc1, p46, p52 and p66, are visible at 46, 52 and 66 kDa, respectively. A monoclonal antibody to Crk (CRK_AS_mAb_1) was used as a control to identify background signal in **c** and **e**. PSM, prestained marker.

yield of 0.6 mg l^{-1} culture (>75% purity) (**Supplementary Table 1**). scFvs were produced with a hexahistidine tag and secreted into the medium. Average yields were 1 mg l^{-1} culture for binders in pSANG10-based vectors and 4.4 mg l^{-1} culture for binders in pOPE101-based vectors, both with >90% purity. Overall, the success rate for protein production from recombinant antibodies was 80%. With additional optimization, we increased yields for purified Fabs to 4–10 mg l^{-1} culture (data not shown).

Measurement of affinity by surface plasmon resonance

Many antibodies have binding affinities in the low nanomolar to picomolar range. To determine whether our systematic approach generated binders in this range, we used SPR to determine on rate, off rate and dissociation constant (K_D) for a selection of renewable antibodies that passed primary validation. We used a parallel format for SPR to simultaneously measure the kinetics of each purified binder to its cognate target and five other SH2 domains. No binders tested showed any binding to noncognate SH2 domains. Of 116 binders tested, 73 (63%) had detectable binding (**Table 2**, **Supplementary Table 1** and **Supplementary Fig. 2**). In this group, 29 binders had K_D values below 20 nM (25%). The success rate is likely an underestimate of antibody binding because we used immobilized antigen in the SPR assay, and the coupling of amines during immobilization may have masked epitopes[19]. For instance, of nine monoclonal antibodies positive by ELISA (**Supplementary Fig. 3**), only three could be validated by SPR. All three of these monoclonal antibodies had high affinity (<4 nM), possibly because of their bivalent nature. Valency of recombinant antibodies can be increased by fusing with bivalent[7] or oligovalent fusion partners[20] resulting in 10–40 fold increase in avidity[20] with concomitant improved performance[21].

The affinity of recombinant antibodies can be increased through additional cycles of mutagenesis and selection[22]. To explore the feasibility of incorporating affinity maturation into a high-throughput pipeline, 93 binders for 19 SH2 domain targets from Fab library F and populations of scFv binders targeting eight SH2 domain targets (unpublished data; M.R.D., K.C., K.P., B.K.K., T.P. and J.Mc.) were subjected to affinity maturation. Many of the affinity-matured binders had affinities in the low nanomolar range (**Table 2** and **Supplementary Table 1**). Such strategies could be used if binders have low affinity or fail in subsequent experimental validation (**Fig. 1**).

Immunoprecipitation of ectopic and endogenous targets

Antibodies are commonly used for immunoprecipitation to capture proteins from their cellular environment to identify binding partners or post-translational modifications. We tested antibodies from each binder type for their ability to immunoprecipitate epitope-tagged and endogenous full-length targets expressed in HEK 293T cells. We focused on three targets (Crk, Grb2 and Shc1) and compared monoclonal antibodies with recombinant antibodies with K_D values <100 nM as defined by SPR.

In anticipation of having to generate antibodies to proteins for which there are no well-characterized antibodies to serve as controls, we expressed epitope-tagged antigens as a general approach for a 'first-pass' characterization of the utility of the renewable antibodies in cell-based assays. Antibodies that recognized ectopically expressed antigen were tested for their ability to recognize endogenous antigens. We used Crk and Grb2 as representative proteins. We named the binders as 'SH2 domain target_principal investigator initials_binder type_number'.

For Crk, all binder types (Fab, monoclonal antibody and scFv) could immunoprecipitate ectopic Crk (**Fig. 2a**). But only the Fab to Crk (CRK_SS_Fab_9) precipitated detectable amounts of endogenous Crk (**Fig. 2b**). Neither the scFv nor the monoclonal antibodies tested could recognize overexpressed Grb2, but the Fab to Grb2 (anti-Grb2 Fab; GRB2_SK_Fab_3) recognized both ectopic and endogenous Grb2 (**Fig. 2c**,**d**). The anti-Grb2 Fab immunoprecipitated Flag-tagged Grb2 as effectively as the antibody to Flag (anti-Flag M2). Although we did not detect endogenous Grb2 or Crk protein using monoclonal antibodies or scFvs, these antibody types immunoprecipitated endogenous Shc1 (**Fig. 2e**). The higher success rate of binders targeting Shc1

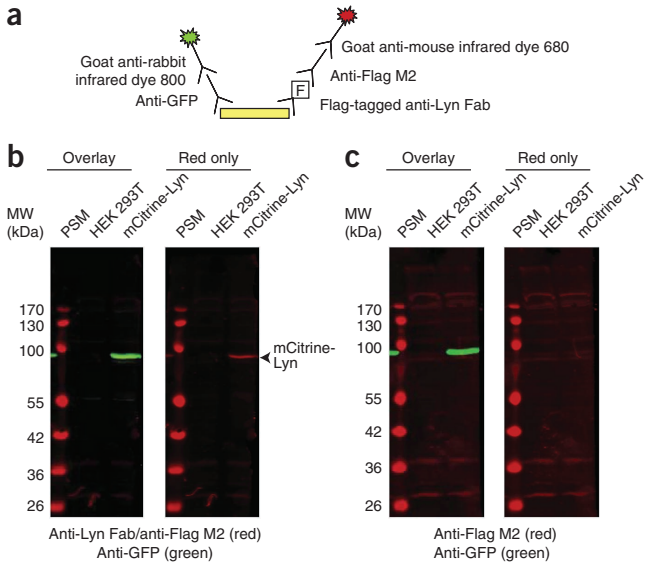

Figure 3 | Anti-Lyn Fab (LYN_SS_Fab_2) recognizes ectopic Lyn expressed in HEK 293T cells. (**a**) Schematic of the antibodies used in the immunoblot analysis. (**b**) Lysates from HEK 293T cells, with (labeled mCitrine-Lyn) or without (labeled HEK 293T) mCitrine-tagged Lyn, were probed with antibodies indicated below the blots. Overlay, signal for anti-GFP and anti-Lyn Fab with the intermediary anti-Flag M2. (**c**) As a control, the same lysates as those described in **b** were probed with anti-GFP and anti-Flag M2. For mCitrine detection, blots were probed with a polyclonal antibody to GFP (anti-GFP).

as compared to those targeting Grb2 or Crk may reflect the accessibility of the Shc1 SH2 domain epitope in the context of the full-length protein.

In total, five of the 18 binders (28%) tested by immunoprecipitation recognized endogenous proteins (**Table 2** and **Supplementary Table 1**), and representatives from each scaffold were successful. We provide protein sequences for recombinant binders that recognized endogenous proteins (**Supplementary Fig. 4**).

Recombinant antibodies can be used for immunoblotting

Recombinant protein binders have been used for immunoblots[23,24] but nowhere near the extent of traditional antibodies. Previously, we showed that 15 of 18 SH2 domain–directed scFvs recognized their cognate purified SH2 domains on immunoblots[16]. To determine whether recombinant binders could recognize their targets in the context of the full-length proteins, we expressed SH2-containing proteins tagged with the yellow fluorescent protein variant monomeric Citrine (mCitrine) in HEK 293T cells and tested the binders using the assay format shown in **Figure 3a**. A Fab to Lyn (anti-Lyn Fab; LYN_SS_Fab_2) recognized a single protein that a polyclonal antibody to GFP (anti-GFP) also recognized (**Fig. 3b**). No signal was generated when we omitted the anti-Lyn Fab from the immunoblot assay (**Fig. 3c**). The mCitrine-tagged Lyn was also recognized by a representative scFv to Lyn (data not shown). Thus, despite being selected against folded proteins, these recombinant binders were effective in immunoblot analysis. It is likely that the small and stable SH2 domains more readily refold after denaturing gel electrophoresis compared with larger and/or unstable proteins, and might be better candidate immunoblot antigens. These results suggest that recombinant antibodies suitable for immunoblotting can be readily generated using our current approaches applied to protein domain antigens.

Immunofluorescence using an scFv to Shc1

To assess whether recombinant antibodies to SH2 domain were functional in immunofluorescence applications, we tested the ability of an scFv to Shc1 (anti-Shc1 scFv; SHC1_JM_scFv_1) to recognize Shc1 in Madin-Darby canine kidney (MDCK) cells. We expressed both YFP-tagged Shc1 (YFP-Shc1) and Erb2 fused to CFP (CFP-ErbB2) together in MDCK cells, and fixed the cells at two different time points after stimulation with epidermal growth factor (EGF): at 5 min when Shc1 is recruited to the plasma membrane by ErbB2 (refs. 15,25) and at 15 min when internalized activated ErbB2 and bound Shc1 accumulate in endocytic vesicles[26].

We tested two different protocols to visualize binding of the anti-Shc1 scFv to Shc1: biotinylated anti-Shc1 scFv binding followed by staining with Cy3-conjugated streptavidin (**Fig. 4a**) or use of the same Flag-tagged anti-Shc1 scFv as the primary antibody, a monoclonal anti-Flag M2 as an intermediary antibody and an Alexa Fluor 588 (Alexa 588) goat anti-mouse IgG as the final antibody (**Fig. 4b**). In both protocols, the anti-Shc1 scFv perfectly localized with ectopically expressed YFP-Shc1 that was bound to activated CFP-ErbB2, and the signal from the anti-Shc1 scFv mirrored the movement of YFP-Shc1 from the membrane to endocytic vesicles. Mock-transfected samples lacking the anti-Shc1 scFv did not show a

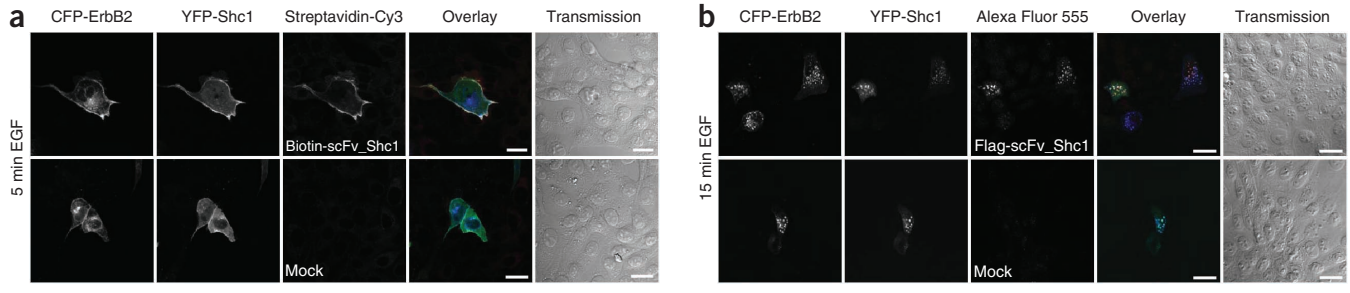

Figure 4 | An scFv recognizes Shc1 in MDCK cells. (**a**,**b**) Confocal images of fixed MDCK cells expressing CFP-ErbB2 and YFP-Shc1 that were stimulated with EGF for the times indicated. In overlay images, shown are signals for CFP-ErbB2 (blue), YFP-Shc1 (green) and anti-Shc1 (red; scFv -SHC1_JM_scFv_1); overlap of all channels appears white. In **a**, cells were stained with biotin-labeled anti-Shc1 scFv (labeled as Biotin-scFv_Shc1) and streptavidin-Cy3 (top) or streptavidin-Cy3 only (bottom). In **b**, cells were incubated with anti-Shc1 scFv (labeled as Flag-scFv-Shc1), anti-Flag M2 that recognizes the Flag-tagged anti-Shc1 scFv and Alexa555 goat anti-mouse IgG (top) or with anti-Flag M2 and Alexa555 goat anti-mouse IgG only (bottom). Scale bars, 20 μm.

specific signal (**Fig. 4**). Thus, anti-Shc1 scFv can specifically recognize ectopic Shc1 protein in MDCK cells.

However, we did not detect endogenous Shc1. This may be due to either low Shc1 expression or high background. Streptavidin-conjugated dyes generate high background signals because streptavidin binds to biotinylated proteins found in the mitochondria[27]. Although we preincubated cells with streptavidin, background staining was not eliminated (**Fig. 4a**). Similarly, the anti-Flag M2 intermediary also produced a background signal (**Fig. 4b**).

DISCUSSION

The success of the integrated pipeline augurs well for a larger-scale initiative. We estimated success rates for each step of the process, and based on this information estimate that 242 initial clones should be screened to have a reasonable chance of obtaining one antibody that recognizes an endogenous target in an immunoprecipitation assay (**Supplementary Table 1**). The accuracy of this attrition rate is likely to increase as more data are collected, but in practice success rates will always be target-dependent.

High-quality antigens were fundamental to our strategy. Assuming that most protein targets will contain at least one domain that can be expressed in reasonable amounts, and that an expression construct has already been derived, a single scientist can produce milligram quantities of ~24 antigens per week. The Structural Genomics Consortium and other structural genomics operations have already produced ~2,000 human proteins that meet these criteria and are adding hundreds more each year. In the near term, access to antigen will not be limiting. Although antibodies can be made to unfolded proteins or peptides, the use of well-folded domains increases the likelihood of obtaining binders that recognize native proteins.

Each selection process reported here generated high-affinity selective binders, but success rates varied between approaches and libraries for individual antigens. This illustrates the value of a multidisciplinary approach to renewable antibody generation and validation. ELISAs were effective in eliminating nonspecific binders and SPR confirmed binder specificity and affinity. Twenty-five percent of binders tested had affinities below 20 nM, including all five binders that recognized endogenous protein in immunoprecipitation assays. Affinity clearly affects antibody performance; in preliminary studies, there is good correlation between binders with K_D values < 60 nM and with slow off rates and their ability to recognize endogenous protein by immunoprecipitation (unpublished data; M.R.D., K.C., K.P., B.K.K., T.P. and J.Mc.). A cut-off for monovalent binding affinity between 20 nM and 60 nM may evolve as the acceptance criterion to move binders into biological assays. However, not all binders with good binding kinetics (for example, monoclonal antibody and scFv to Crk and Grb2) recognized full-length protein in cells, presumably because the epitope is masked. The ability to screen multiple antibodies per target in a high-throughput format helps overcome this.

The high-throughput, automated monoclonal production has several advantages over classical methods. By using robotic platforms to perform many of the tissue culture operations, throughput was considerably increased. Additionally, the AMA assay allows screening of up to 100 targets simultaneously, identifies antibodies that recognize different structural and post-translational modifications, determines antibody isotypic subclasses and enables epitope mapping at an early stage in clone selection.

We found that recombinant antibodies were comparable to monoclonal antibodies in immunoprecipitation assays. Recombinant approaches offer unique advantages[28]. First, because the binder comes from an *in vitro* library, it can be cloned into any downstream vector, making it easily customizable by changing the tag or fusion partner[23], or even converting it to an IgG format[29,30]. Recombinant binders can be biotinylated *in vivo* and also multimerized to increase avidity[20]. In the future, different tags may be used for different applications and a challenge will be to identify optimal and standardized configurations for immunological assays. As an example, we found that recombinant binders function in immunofluorescence assays, but the secondary detection methods produced a high background. Future work should focus on identifying a format for this and other assays with improved signal-to-noise ratio and ease of use. Another advantage of *in vitro* systems is that competitors can be introduced during the selection process (for example, related family members) to enrich for antibodies that can distinguish between closely related molecules. Even without the addition of competitors in the selection process, binders from our initial screens could distinguish between Abl1 and Abl2 SH2 domains, which have 89% sequence identity[17].

Members in each of four laboratories produced many recombinant antibody fragments to the 20 targets (**Table 1**). For many targets, the number of clones was arbitrarily reduced for downstream applications. Thus, it seems reasonable to screen a predefined number of clones, as calculated by the attrition rates; we suggest one 96-well microtiter plate per target per binder type. The initial selection and primary ELISA screening in each laboratory could typically be performed within 1–2 person-months. With extended sequencing, specificity analysis and affinity maturation this could be increased to 12 person-months. The time frame in the hybridoma laboratory was routinely 12–16 weeks per target including immunization. Two staff members using robotic automation produced 165 cell lines in 6 months. The AMA method[8] reduced the initial 38,400 supernatants to a more manageable level for confirmatory ELISA.

Downstream biological validation would benefit most from optimization and streamlining. We propose that biological validation could be systematized by creating tagged full-length protein expression constructs to screen for binders active for immunoblotting, immunoprecipitation and immunofluorescence. Once candidate binders that recognize ectopically expressed targets are identified, they will be tested for their ability to recognize their endogenous target by probing wild-type cells, cell lines in which the endogenous protein is knocked down and after competition with the purified target. Confidence in any assay will increase if similar results are obtained with independent binders, ideally targeting different epitopes or domains.

Provided with purified targets, we believe that it is now possible to set up a pipeline to rapidly create and validate large panels of renewable antibodies with specific properties in a reasonable time frame. The ability to generate a core set of antibodies for each human protein will prove invaluable to researchers in multiple biological fields. We can already imagine adding other technologies into the pipeline, to generate binders that recognize different protein conformations, post-translational modifications, splice variants or protein complexes[11,28]. Looking to the future,

our ultimate aim is to create a pipeline, similar to DNA synthesizers in the 1980s, capable of supplying 'application-specific antibodies on demand'.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We acknowledge B. Liu and P. Nash for providing SH2 domain alignments; N. Bisson (Samuel Lunenfeld Research Institute) for providing a stable cell line encoding Flag-tagged Grb2; M. Taussig and M. Sundstrom for helpful discussions; and N. Bisson and B. Liu for critically reading the manuscript. The Structural Genomics Consortium is a registered charity (number 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research and the Wellcome Trust. This work was also supported by funds from Genome Canada through the Ontario Genomics Institute and Ontario Research Fund Global Leadership Round in Genomics and Life Sciences (to T.P.), The Human Frontier Science Program (O.R.), the European Commission 6th framework program coordination action 'Proteome Binders' (S.D.), from the Systematisch-Methodischen Platform Antibody Factory, within the German Nationales Genomforschungsnetz (S.D.), the Land Niedersachsen (S.D.), the Wellcome Trust (J.Mc.), the US National Institutes of Health (GM082288-09A1 and EY016094-01A1 to B.K.K. and R01-GM72688 and U54-GM74946 to A.A.K. and S.K.), to the Swedish Research Council (524-2008-617) (H.P.), and the Victorian State Government (The Department of Innovation, Industry and Regional Development), the Australian Federal Government (Bioplatforms Australia) and Monash University for financial support for the Monash Antibody Technologies Facility.

AUTHOR CONTRIBUTIONS

K.P., J.D.P., A.K.-V., D.J.S., N.E.J., A.W., J.Wo., A.K., M.M., D.M., J.M., S.H., S.B., A.F., M.S., K.W., A.D., K.H., M.S., R.S., J.M.S., A.S., J.O., S.H., L.-G.D., A.F., I.J., L.C., P.L., I.K., D.L. and F.A.F. designed and performed experiments; M.R.D., M.H., H.P., O.R., P.N., E.N. and D.C. conceived, designed and performed experiments and wrote the paper; J.We., C.H.A., B.K.K., A.A.K. and M.U. oversaw the project; J.Mc., S.K., S.S., S.D., A.S., T.P. and A.M.E. conceived and oversaw the project and wrote the paper; K.C. and S.G. conceived, designed and performed experiments, oversaw the project and wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Taussig, M.J. *et al.* ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome. *Nat. Methods* **4**, 13–17 (2007).
2. Bordeaux, J. *et al.* Antibody validation. *Biotechniques* **48**, 197–209 (2010).
3. Uhlén, M. *et al.* A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* **4**, 1920–1932 (2005).
4. Berglund, L. *et al.* A genecentric Human Protein Atlas for expression profiles based on antibodies. *Mol. Cell. Proteomics* **7**, 2019–2027 (2008).
5. Nilsson, P. *et al.* Towards a human proteome atlas: high-throughput generation of mono-specific antibodies for tissue profiling. *Proteomics* **5**, 4327–4337 (2005).
6. Ohara, R. *et al.* Antibodies for proteomic research: comparison of traditional immunization with recombinant antibody technology. *Proteomics* **6**, 2638–2646 (2006).
7. Schofield, D.J. *et al.* Application of phage display to high throughput antibody generation and characterization. *Genome Biol.* **8**, R254 (2007).
8. De Masi, F. *et al.* High throughput production of mouse monoclonal antibodies using antigen microarrays. *Proteomics* **5**, 4070–4081 (2005).
9. Hust, M. & Dübel, S. Mating antibody phage display with proteomics. *Trends Biotechnol.* **22**, 8–14 (2004).
10. Konthur, Z., Hust, M. & Dübel, S. Perspectives for systematic *in vitro* antibody generation. *Gene* **364**, 19–29 (2005).
11. Dübel, S., Stoevesand, O., Taussig, M.J. & Hust, M. Generating recombinant antibodies to the complete human proteome. *Trends Biotechnol.* **28**, 333–339 (2010).
12. Blow, N. Antibodies: the generation game. *Nature* **447**, 741–744 (2007).
13. Uhlén, M., Gräslund, S. & Sundstrom, M. A pilot project to generate affinity reagents to human proteins. *Nat. Methods* **5**, 854–855 (2008).
14. Scott, J.D. & Pawson, T. Cell signaling in space and time: where proteins come together and when they're apart. *Science* **326**, 1220–1224 (2009).
15. Seet, B.T., Dikic, I., Zhou, M.M. & Pawson, T. Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* **7**, 473–483 (2006).
16. Mersmann, M. *et al.* Towards proteome scale antibody selections using phage display. *New Biotechnol.* **27**, 118–128 (2010).
17. Pershad, K. *et al.* Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display. *Protein Eng. Des. Sel.* **23**, 279–288 (2010).
18. Hust, M. *et al.* A human scFv antibody generation pipeline for proteome research. *J. Biotechnol.* **152**, 159–170 (2010).
19. Nice, E.C. & Catimel, B. Instrumental biosensors: new perspectives for the analysis of biomolecular interactions. *Bioessays* **21**, 339–352 (1999).
20. Thie, H., Binius, S., Schirrmann, T., Hust, M. & Dübel, S. Multimerization domains for antibody phage display and antibody production. *New Biotechnol.* **26**, 314–321 (2009).
21. Martin, C.D. *et al.* A simple vector system to improve performance and utilisation of recombinant antibodies. *BMC Biotechnol.* **6**, 46 (2006).
22. Dufner, P., Jermutus, L. & Minter, R.R. Harnessing phage and ribosome display for antibody optimisation. *Trends Biotechnol.* **24**, 523–529 (2006).
23. Moutel, S. *et al.* A multi-Fc-species system for recombinant antibody production. *BMC Biotechnol.* **9**, 14 (2009).
24. Schütte, M. *et al.* Identification of a putative Crf splice variant and generation of recombinant antibodies for the specific detection of *Aspergillus fumigatus*. *PLoS ONE* **4**, e6625 (2009).
25. Luzi, L., Confalonieri, S., Di Fiore, P.P. & Pelicci, P.G. Evolution of Shc functions from nematode to human. *Curr. Opin. Genet. Dev.* **10**, 668–674 (2000).
26. Sorkin, A. Internalization of the epidermal growth factor receptor: role in signalling. *Biochem. Soc. Trans.* **29**, 480–484 (2001).
27. Hollinshead, M., Sanderson, J. & Vaux, D.J. Anti-biotin antibodies offer superior organelle-specific labeling of mitochondria over avidin or streptavidin. *J. Histochem. Cytochem.* **45**, 1053–1057 (1997).
28. Bradbury, A.R., Sidhu, S., Dübel, S. & McCafferty, J. Beyond natural antibodies: the power of *in vitro* display technologies. *Nat. Biotechnol.* **29**, 245–254 (2011).
29. Li, B. *et al.* Activation of the proapoptotic death receptor DR5 by oligomeric peptide and antibody agonists. *J. Mol. Biol.* **361**, 522–536 (2006).
30. Thie, H. *et al.* Rise and fall of an anti-MUC1 specific antibody. *PLoS ONE* **6**, e15921 (2011).

[2]The full list of authors and affiliations is as follows:

***In vitro* Antibody Consortium**

Helena Persson[4,13], Nicholas E Jarvik[4], Arkadiusz Wyrzucki[5], John Wojcik[5], Akiko Koide[5], Anthony A Kossiakoff[5], Shohei Koide[5], Sachdev Sidhu[4], Michael R Dyson[6], Kritika Pershad[7], John D Pavlovic[7], Aneesh Karatt-Vellatt[6], Darren J Schofield[6], Brian K Kay[7], John McCafferty[6], Michael Mersmann[8], Doris Meier[8], Jana Mersmann[8], Saskia Helmsing[8], Michael Hust[8], Stefan Dübel[8]

Monash Antibody Technologies Facility

Susie Berkowicz[9], Alexia Freemantle[9], Michael Spiegel[9], Alan Sawyer[9,13], Daniel Layton[9], Edouard Nice[9]

Pawson Laboratory

Anna Dai[1], Oliver Rocks[1], Kelly Williton[1], Frederic A Fellouse[1], Kadija Hersi[1], Tony Pawson[1,10]

Human Protein Atlas

Peter Nilsson[11], Mårten Sundberg[11], Ronald Sjöberg[11], Åsa Sivertsson[11], Jochen M Schwenk[11], Jenny Ottosson Takanen[11], Sophia Hober[11], Mathias Uhlén[11]

Structural Genomics Consortium

Lars-Göran Dahlgren[3,13], Alex Flores[3], Ida Johansson[3], Johan Weigelt[3], Lissette Crombet[12], Peter Loppnau[12], Ivona Kozieradzki[12], Doug Cossar[12], Cheryl H Arrowsmith[12], Aled M Edwards[12]

[4]Banting and Best Department of Medical Research, Department of Molecular Genetics, and the Terrence Donnelly Center for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario, Canada. [5]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois, USA. [6]Department of Biochemistry, University of Cambridge, Downing Site, Cambridge, UK. [7]Department of Biological Sciences, University of Illinois at Chicago, Chicago, Illinois, USA. [8]Technische Universität Braunschweig, Institute of Biochemistry and Biotechnology, Braunschweig, Germany. [9]Monash Antibody Technologies Facility, Faculty of Medicine, Nursing and Health Sciences, Monash University, Victoria, Australia. [10]Department of Molecular and Medical Genetics, University of Toronto, Ontario, Canada. [11]Department of Proteomics, School of Biotechnology, Royal Institute of Technology, Stockholm, Sweden. [12]Structural Genomics Consortium, Toronto, Ontario, Canada. [13]Present addresses: Department of Immunotechnology, Lund University, Lund, Sweden (H.P.); European Molecular Biology Laboratory Monoclonal Core Facility, Monterotondo-Scalo, Lazio, Italy (A.S.); School of Biological Sciences, Nanyang Technological University, Lab 7-01 Institute of Molecular and Cell Biology, Proteos, Singapore (L.-G.D.).

ONLINE METHODS

Subcloning. The SH2 domain alignment used to establish boundaries has been previously described[31]. Globplot was used to predict boundaries of globular regions[32]. Primers, sequences and initial templates for the SH2 domains are described in **Supplementary Table 3**. Vectors were tracked through the OpenFreezer reagent tracker (unpublished data; K.C., K.W., A.D. and T.P.) and assigned a unique identifier starting with V. The SH2 domains were amplified by PCR using Pfu ultra polymerase (Stratagene) and cloned into the AscI and PacI sites V2082 (pet28 SacB AscI PacI (AP)) by In Fusion (Clontech) (vector map and sequence in **Supplementary Fig. 5**). The clone collection is available through Open Biosystems (OHS4902).

To create mCitrine-tagged Lyn (V4652), we cloned the full-length open reading frame from template BC075001 into an N-terminal mCitrine vector (V2182) using the Gateway recombination system (Invitrogen). For Flag-tagged Grb2, the full-length open reading frame from template BC000631 with sequence encoding a triple Flag-tag epitope was cloned into pMSCVpuro and for myc-tagged Grb2 (V4908), the same open reading frame (ORF) was cloned into vector V516 using the Creator system[33]. The full-length ORF from template BC008506 was cloned into an mCitrine vector (V3522) to generate mCitrine-tagged Crk (V5811).

Antigen production. Antigens were produced as described previously[16,17]. Briefly, we expressed the 20 SH2 domains in *Escherichia coli* in terrific broth (TB) medium and purified them using a two-step protocol including immobilized metal-ion chromatography (IMAC) and gel filtration on an ÄKTA xpress system. The proteins were eluted in PBS (10 mM $NaPO_4$ and 154 mM NaCl; pH 7.5) and protein concentrations were measured. Resulting batches were quickly frozen in liquid nitrogen and stored at −80 °C. The identities of the proteins were confirmed by mass spectrometry[34], and 1 mg aliquots of the antigens were sent to labs where binders were generated.

Phage-display selection of single-chain Fvs and Fabs. Selection of scFv antibody fragments by phage display from the HAL libraries or 'McCafferty' library was performed as described previously[16,17]. Briefly, for the screening using HAL libraries, SH2 domains were coated onto either 96-well ELISA plates (Nunc) or immobilized on Dynabeads (Invitrogen)[16]. During the panning, 2.5×10^{11} scFv phage particles of HAL4 (Vκ repertoire) and 2.5×10^{11} scFv of HAL7 (Vλ repertoire) were added per well, and four wells were used per SH2 domain. In the next two rounds, antigen amount was reduced by 50%, M13K07 was introduced as a helper phage, phage input was reduced to ~1×10^{10} phage particles and washing cycles were increased. Panning conditions on the beads were identical to the 96-well-plate pannings. The 'McCafferty' library, with a diversity of 1.1×10^{10} clones[7], was screened against the 20 SH2 domain containing proteins[17] in 96-well-plate format. Only two rounds of selections were used, and the helper phage was cleaved by L-1-tosylamide-2-phenylethyl chloromethyl ketone (TPCK)-trypsin (Sigma). We created affinity-matured (chain-shuffled) binders by amplifying the variable heavy (V_H) inserts of selected scFvs by PCR and ligating the resulting inserts of the same target into the phagemid pSANG4 vector harboring naive variable light chain libraries[7] (unpublished data; M.R.D., K.C., K.P., B.K.K., T.P. and J.Mc.). The ligation reaction was transformed into TG1 cells by electroporation, and the resulting libraries were stored at −80 °C. Rescue of phage particles from the chain-shuffled libraries and two rounds of selections against the eight SH2 domains were done as described previously[7] except that in some selections we used biotinylated antigens (1–100 nM) captured with 25 µl of M-280 streptavidin dynabeads (Invitrogen).

The Fab libraries used synthetic hypervariable loops built on a humanized antibody scaffold and were constructed as described previously[35] but with greater diversity allowed in the third hypervariable loop of the light chain. Selections of Fab library E were performed essentially as described previously[35]. Selections of Fab library F were performed using a high-throughput selection method that enables parallel selection of a large number of antigens. This method was based on our previous work[36] but with a few important modifications. Briefly, four rounds of selections were carried out in 96-well microtiter plates, using four wells per antigen. After eight washes, the captured phages were used directly to infect bacteria without prior elution. We amplified phages overnight in a 96-well deep-well plates with no precipitation of phage before the next round of selection. For affinity maturation, additional diversity was introduced in the first and second hypervariable loops of the heavy chain or the light chain followed by another three rounds of panning of higher stringency.

Production and selection of monoclonal antibodies. Monoclonal antibodies were selected as previously described[8]. Briefly, BALB/c mice were each immunized with 10 µg of each antigen (two mice per target, 40 mice total) followed by boosters of 10 µg at 14-d intervals with a final prefusion boost 28 d after the third immunization. Serum ELISA identified the highest responder in each group, which was taken for hybridoma production (20 mice total). For hybridoma production, the splenocytes from each of these mice were then collected and fused to Sp2 myeloma fusion partners (American Type Culture Collection; ATCC). On day 11 after fusion, 40 µl of each supernatant was transferred to 384-well plates for AMA[8]. For AMA, we coated 20 µg of antigen onto aminosilane-modified microscope slides, and the hybridoma supernatants were spotted onto these slides using an Arrayjet Super Marathon inkjet microarrayer. Detection was by a mixture of Cy3-conjugated goat anti-mouse IgG and Cy5-conjugated goat anti-mouse IgM (Jackson Immunoresearch) at 1:1,000 dilution. IgG positives from the AMA assay were then validated by automated ELISA. Tecan EVO robotics were used to automate production and for automated ELISA screening.

ELISA screening of antibodies. We performed ELISAs as previously described[8,16,17,37,38]. Briefly, ELISAs were performed in 96- or 384-well formats with 100–200 ng of antigen coated onto each well. The antigens were incubated with varying amounts of phage or hybridoma supernatant and subsequently washed in PBS with 0.05 to 0.1% Tween. Detection methods varied depending on the type of technology and library used: detection was via a secondary antibody to Flag (Sigma) labeled with europium ("McCafferty" library), a mouse anti-Myc 9E10 followed by goat anti-mouse antibody conjugated to horseradish peroxidase (Sigma) (HAL library), a goat-anti-mouse antibody conjugated to horseradish peroxidase (Jackson Dianova) (hybridomas) and finally a horseradish peroxidase–conjugated anti-M13 monoclonal (GE Healthcare) (Fab libraries). After this initial validation,

scFvs and Fabs were subjected to restriction enzyme analysis or sequencing to determine unique clones.

Protein arrays. Microarrays were constructed by spotting protein onto epoxy-coated slides (Corning Life Sciences) with 14 identical sub-arrays on each slide, using a noncontact Nanoplotter 2.0E (GeSim). Each sub-array contained 432 protein fragments. This included 85 protein epitope signature tags (PrESTs) corresponding to 53 unique SH2-domain containing proteins, expressed as His6/albumin binding protein fusions and the 20 folded SH2 domains. The latter group was spotted at different concentrations ranging from 0.4 mg ml^{-1} to 4.0 mg ml^{-1} and also 1:1 dilutions of each. The PrESTs were diluted to 40 μg ml^{-1} in 0.1 M urea and 50 mM sodium carbonate–bicarbonate buffer (pH 9.6), complemented with 100 mg ml^{-1} BSA. Slides were blocked in 3% BSA in PBS with 0.1% Tween. All affinity binders were diluted between 1:10 up to 1:10,000 in PBST (some with addition of milk) depending on concentration and background signal. A 14-well silicon mask was put on the microarray slide and 60 ml of each of the diluted affinity binder was incubated for 1 h. The detection was done by either a fluorophore-labeled secondary antibody or a secondary biotin-labeled reagent and a fluorophore-labeled streptavidin, all incubated for 1 h.

Scale-up production of hybridoma monoclonal antibodies. Monoclonal antibodies were collected in the supernatant of hybridoma cell lines cultured to exhaustion in DMEM supplemented with 10% fetal bovine serum, 200 mM GlutaMAX and 1% HyBer. We centrifuged the supernatant at 2,000*g* and filtered with a 0.22 μm filter to remove cells and debris. The resultant supernatant was purified using a Profinia purification system (Bio-Rad) and a HiTrap Protein G column (GE Healthcare). The supernatant was applied to a column that had been pre-equilibrated with binding buffer (20 mM sodium phosphate; pH 7.0), at a flow rate of 1 ml min^{-1}. The column was washed with 10 ml binding buffer and the antibody was eluted with 4 ml elution buffer (0.1 M glycine-HCl; pH 2.7) and the pH was adjusted to pH 7 using NaOH (0.1 M; pH 9.0).

Scale-up production of single-chain antibodies. We subcloned sequences encoding the single-chain Fv variants into the pOPE101-Express (compatible restriction enzyme sites with the HAL-based libraries) or the pSANG10/pSANG10-TEV (compatible restriction enzyme sites with the 'McCafferty' library) vectors, respectively. Both contain a PelB leader sequence exporting the recombinant proteins to the periplasmic space and a hexahistidine tag to enable IMAC purification. pOPE expression clones were transformed to the *E. coli* strain JM109 and pSANG10-based clones to BL21 (DE3) R3 pRARE2 cells, respectively. Protein expression was performed in 1,500 ml TB medium essentially as described previously[16,17]. The resulting cell pellets, on average 27 g of cells per cultivation, were resuspended in ice-cold sucrose buffer (50 mm HEPES, 1 mM EDTA and 20% (w/v) sucrose (pH 8.0) supplemented with Complete EDTA-free protease inhibitor (Roche Applied Science; 1 tablet per 100 ml buffer)) at 1.5 ml g^{-1} of cells and placed on a shaking table in the cold room for 1 h. The cells were then collected by centrifugation at 6,000*g* for 20 min and the supernatants containing extracted antibodies were decanted into beakers on ice. The pellets were then resuspended in 5 mM $MgSO_4$ buffer (same volume as in the previous step) supplemented with 5 mM $MgSO_4$ and 25 U ml^{-1} benzonase (125 μl l^{-1}) and then incubated on ice for 30 min before being transferred to SA-800 tubes and centrifuged at 30,000*g* for 10 min. The $MgSO_4$ extracts were decanted into the same beakers (as for the first extraction) on ice and the total volume was estimated. One-half of the estimated volume was added of 3× IMAC bind and wash buffer 1 (150 mM HEPES, 1.5 M NaCl, 30% glycerol and 30 mM imidazole; pH 8.0) resulting in a threefold dilution to 50 mM HEPES, 0.5 M NaCl, 10% glycerol and 10 mM imidazole; pH 8.0. The supernatants were filtered with 0.45 μm filters and loaded onto the ÄKTA Xpress system. The proteins were purified first by IMAC (1 ml HisTrap HP columns, GE Healthcare) using wash buffer (50 mM HEPES, 500 mM NaCl, 10% glycerol and 25 mM imidazole; pH 7.5) and elution buffer (50 mM HEPES, 500 mM NaCl, 10% glycerol and 500 mM imidazole; pH 7.5). As a second step, the proteins were gel-filtered (HiLoad XK16/60 Superdex 200 columns; GE Healthcare) and finally eluted in gel filtration buffer (20 mM HEPES, 300 mM NaCl and 10% glycerol; pH 7.5). Peaks were analyzed by SDS-PAGE, and relevant fractions corresponding to monomeric proteins were pooled. Concentrations were measured on a Nanodrop ND-1000 (NanoDrop Technologies) spectrophotometer.

Scale-up production of Fabs. We transformed competent cells of *E. coli* (American Type Culture collection (ATCC) 55244) with a Fab expression plasmid and picked single colonies for large-scale growth. Overnight cultures in 12 ml 2× yeast extract, tryptone (YT) medium at 30 °C were collected by centrifugation (3,000*g* for 10 min) and the pellet was resuspended in an equal volume of fresh low-phosphate production medium (3.57 g $(NH_4)_2SO_4$; 0.71 g Na citrate.$2H_2O$; 1.07 g KCl; 5.36 g yeast extract; 5.36 g Hy-Case SF casein; 100 ml 1 M MOPS pH 7.3; 5 g glucose; 7 ml 1 M $MgSO_4$; and 100 mg ampicillin). The cell suspension was added to one liter of production medium in 2.8-l Fernbach flasks and incubated for 24 h at 25 °C at 230 r.p.m. in Innova 44 (New Brunswick Scientific) incubators. Cells were collected and stored frozen at −80 °C until extraction. Thawed cell pellet was resuspended in 30 ml PBS and cells lysed by passage through a microfluidizer. Lysate was clarified by centrifugation and the Fab was recovered by chromatography on protein A–sepharose (GE Biosciences). A 1 ml column was regenerated before use with 10 ml of 0.1 M H_3PO_4 and equilibrated with 10 ml of PBS (pH 7.4). After sample loading, the column was washed with 30 ml of PBS (pH 7.4) and the Fab eluted with elution buffer (50 mM NaH_2PO_4, 100 mM H_3PO_4 and 140 mM NaCl; pH 2.8). Eluted fractions were immediately neutralized with 25% (v/v) 1 M Na_2HPO_4, 140 mM NaCl (pH 8.6). Fab samples were diluted to between 0.2 and 0.5 mg ml^{-1} with PBS and stored at 4 °C.

Surface plasmon resonance analysis. We performed SPR analysis on a ProteOn XPR36 Protein Interaction Array System (BioRad). SH2 domains were immobilized in 50 mM phosphate buffer (pH 6 or 7 depending on domain pI) to a non-dilute EDAC, sulfo-NHS activated GLC surface using a flow rate of 30 μl min^{-1} for 5 min in the vertical direction. Immobilization levels were monitored to ensure immobilization of approximately 200–500 response units of each SH2 domain; a second injection was used if needed. The domains were then stabilized with PBS for 30 s and 0.85% H_3PO_4 for 18 s each at 100 μl min^{-1}.

doi:10.1038/nmeth.1607

Fabs were diluted in PBS plus 0.05% Tween 20 (PBST) at a starting concentration of 200 nM. For a subset of Fabs and all scFvs, the starting concentration was 400 nM. The binders were further diluted with PBST twofold in series to produce 5 concentrations of Fab. A PBST blank was also included. Fab injection parameters were: 100 µl min^{-1}, 60 s contact time and 600 s dissociation time, in the horizontal direction. SH2 domains were regenerated with an injection of 0.85% H3PO4 at a flow rate of 100 µl min^{-1} followed by a PBST wash of 30 s at 100 µl min^{-1} flow rate.

Immunoblotting. We propagated HEK 293T cells in DMEM media supplemented with 10% fetal bovine serum (FBS). Where indicated, HEK 293T cells were transfected overnight with mCitrine-tagged Lyn using polyethyleneimine (Sigma). Cells (~10^7) were lysed in NP40 lysis buffer (50 mM Hepes (pH 8), 100 mM KCl, 10% glycerol, 2 mM EDTA, 50 mM NaF, 0.5% NP40, 50 mM β-glycerol phosphate supplemented with 10 µg ml^{-1} aprotinin, 10 µg ml^{-1} leupeptin, 1 mM PMSF, 1 mM DTT, 0.1mM orthovanadate), and the lysates were clarified by centrifugation. The resulting supernatant was heated to 95 °C in 2× SDS sample buffer, resolved by 10% SDS-PAGE and transferred to FluoroTrans Membrane (Pall). The blots were probed with a polyclonal antibody to GFP (anti-GFP) for mCitrine recognition (1:5,000, Abcam), 2 µg ml^{-1} anti-Lyn Fab (LYN_SS_Fab_2) and anti-Flag M2 (1 µg ml^{-1}, Sigma) overnight as indicated, rinsed in PBS, and probed with goat anti-rabbit infrared dye (IR) 800 (1:15,000, Mandel) and goat anti-mouse IR 680 (1:15,000, Mandel). Blots were visualized on a LiCor Odyssey.

Immunoprecipitation. We transfected HEK 293T cells (~10^7 cells) with various constructs and lysed them in NP40 lysis buffer. For immunoprecipitates, myc-tagged anti-Crk-scFv (CRK_SD_scFv_2) was pre-bound to anti-Myc 9E10 (ref. 39) agarose beads (Sigma), flag-tagged anti-Grb2 scFV (GRB2_JM_scFv_3) was pre-bound to anti-Flag M2 agarose (Sigma) and monoclonal antibodies were pre-bound to mouse anti-IgG agarose (Sigma). anti-Grb2 Fab (GRB2_SK_Fab_3), anti-Crk Fab (CRK_SS_Fab_9) and anti-Shc1 (SHC1_JM_scFv_1) were biotinylated (Fabs with sulfo-NHS-SS-biotin and scFV with sulfo-NHS-LC-biotin) according to the EZ-Link biotinylation kit instructions (Pierce) and pre-bound to Streptavidin Magnesphere beads (Promega). The pre-bound antibodies were then added to clarified lysate. For the anti-Flag M2 immunoprecipitation, 5 µl of packed anti-Flag M2 agarose (Sigma) was added to the lysate. After 1 to 2 h of incubation, the immunoprecipitates were washed with NP40 lysis buffer and boiled in 2× SDS sample buffer. The samples were resolved by SDS-PAGE and immunoblot analysis was performed as described above using the following dilutions of antibodies: 1:1,000 dilution of anti-Grb2 (BD), 1:1,000 dilution of anti-Crk (BD), 1:2,000 dilution of anti-Shc1 (BD), 1:3,000 dilution of a polyclonal antibody to GFP for mCitrine recognition (Abcam), and 1:15,000 dilution of goat anti-mouse IR 680 and/or goat anti-rabbit IR 800 (Mandel). Blots were visualized on a LiCor Odyssey.

Immunofluorescence. We grew MDCK cells on 35 mm glass-bottom dishes (MatTekCorp), co-transfected overnight with CFP-tagged ErbB2 and YFP-tagged Shc1 by Effectene (Qiagen), serum-starved starved for 4 h and stimulated with EGF (100 ng ml^{-1}; Calbiochem) for the time indicated. Cells were fixed in 4% paraformaldehyde for 10 min, permeabilized in 0.1% Triton X-100 phosphate-buffered saline (PBS) for 10 min, washed with PBS and blocked with 4% BSA in PBS. For cell staining with biotinylated anti-Shc1 scFv (SHC1_JM_scFv_1), the cells were additionally pre-incubated with 0.1 mg ml^{-1} Streptavidin in PBS, washed with PBS, pre-incubated with 0.5 mg ml^{-1} biotin in PBS and then washed further with PBS. For non-biotinylated anti-Shc1 scFv (SHC1_JM_scFv_1), cells were incubated with 1 µg ml^{-1} anti-Shc1 scFv in PBS plus 4% BSA, rinsed with PBS, incubated with 5 µg ml^{-1} anti-Flag M2 monoclonal (Sigma), rinsed with PBS and then incubated with 2 µg ml^{-1} Alexa555 goat anti-mouse IgG (Molecular Probes) in PBS with 4% BSA, followed by extensive rinsing with PBS. For biotinylated anti-Shc1 scFv, the cells were incubated with 1 µg ml^{-1} of biotinylated anti-Shc1 scFv in PBS with 4% BSA, washed with PBS, incubated with 0.5 µg ml^{-1} of Cy3-conjugated streptavidin (Jackson ImmunoResearch), followed by extensive rinsing with PBS. The cells were imaged on an Olympus FV1000 confocal microscope equipped with a 63×, 1.2 NA water-immersion lens.

Renewable antibodies will be available upon request.

31. Liu, B.A. *et al.* The human and mouse complement of SH2 domain proteins-establishing the boundaries of phosphotyrosine signaling. *Mol. Cell* **22**, 851–868 (2006).
32. Linding, R., Russell, R.B., Neduva, V. & Gibson, T.J. GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res.* **31**, 3701–3708 (2003).
33. Colwill, K. *et al.* Modification of the Creator recombination system for proteomics applications–improved expression by addition of splice sites. *BMC Biotechnol.* **6**, 13 (2006).
34. Sundqvist, G., Stenvall, M., Berglund, H., Ottosson, J. & Brumer, H. A general, robust method for the quality control of intact proteins using LC-ESI-MS. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **852**, 188–194 (2007).
35. Fellouse, F.A. *et al.* High-throughput generation of synthetic antibodies from highly functional minimalist phage-displayed libraries. *J. Mol. Biol.* **373**, 924–940 (2007).
36. Sidhu, S.S. *et al.* Phage-displayed antibody libraries of synthetic heavy chain complementarity determining regions. *J. Mol. Biol.* **338**, 299–310 (2004).
37. Hust, M. *et al.* Improved microtitre plate production of single chain Fv fragments in *Escherichia coli*. *New Biotechnol.* **25**, 424–428 (2009).
38. Deshayes, K. *et al.* Rapid identification of small binding motifs with high-throughput phage display: discovery of peptidic antagonists of IGF-1 function. *Chem. Biol.* **9**, 495–505 (2002).
39. Evan, G.I., Lewis, G.K., Ramsay, G. & Bishop, J.M. Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell. Biol.* **5**, 3610–3616 (1985).