ARTICLES

A quantitative protein interaction network for the ErbB receptors using protein microarrays

Richard B. Jones[1]*, Andrew Gordus[1,2]*, Jordan A. Krall[1] & Gavin MacBeath[1]

Although epidermal growth factor receptor (EGFR; also called ErbB1) and its relatives initiate one of the most well-studied signalling networks, there is not yet a genome-wide view of even the earliest step in this pathway: recruitment of proteins to the activated receptors. Here we use protein microarrays comprising virtually every Src homology 2 (SH2) and phosphotyrosine binding (PTB) domain encoded in the human genome to measure the equilibrium dissociation constant of each domain for 61 peptides representing physiological sites of tyrosine phosphorylation on the four ErbB receptors. This involved 77,592 independent biochemical measurements and provided a quantitative protein interaction network that reveals many new interactions, including ones that fall outside of our current view of domain selectivity. By slicing through the network at different affinity thresholds, we found surprising differences between the receptors. Most notably, EGFR and ErbB2 become markedly more promiscuous as the threshold is lowered, whereas ErbB3 does not. Because EGFR and ErbB2 are overexpressed in many human cancers, our results suggest that the extent to which promiscuity changes with protein concentration may contribute to the oncogenic potential of receptor tyrosine kinases, and perhaps other signalling proteins as well.

The four human ErbB receptors induce a wide variety of cellular responses, ranging from migration to adhesion and from growth to apoptosis[1]. Ligand binding to the extracellular domain promotes receptor dimerization and activation of the intracellular tyrosine kinase domain. Activated receptors phosphorylate each other on a number of tyrosine residues, which serve as docking sites for the SH2 (ref. 2) or PTB (ref. 3) domains of downstream enzymes or adaptor proteins. We asked whether a purely biophysical analysis of protein recruitment, performed on a genome-wide scale, could provide insight into the nature of ErbB signalling at a system level. Our approach, which uses microarrays of SH2 and PTB domains to measure the affinity of each domain for peptides that represent physiological sites of tyrosine phosphorylation, yielded the following insights: (1) investigating interactions in a non-competitive format reveals high-affinity binding sites for SH2 and PTB domains, many of which do not conform to consensus recognition sequences[4,5]; (2) the recruitment sites on ErbB2 are much more promiscuous than those on the other receptors; (3) when only the highest affinity interactions are considered, the proteins that bind to EGFR constitute a small subset of those that bind to ErbB3; and (4) EGFR and ErbB2 become much more promiscuous when their concentration is raised, whereas ErbB3 does not. This, we propose, contributes to the high oncogenic potential of EGFR and ErbB2 and suggests alternative strategies for therapeutic intervention.

SH2 and PTB domains

To explore protein recruitment on a genome-wide scale, we began by cloning, expressing and purifying every SH2 and PTB domain encoded in the human genome. From an initial list of 109 SH2 domains and 44 PTB domains, we were able to obtain sequence-verified clones for 106 SH2 domains and 41 PTB domains (Supplementary Table 1 and Supplementary Fig. 1). Some human proteins contain two SH2 or PTB domains. When these tandem domains are close together, their adjacency can affect their recognition properties[6]. We therefore cloned not only the isolated domains, but also the ten tandem SH2 domains and three tandem PTB domains found in the human genome.

Because our goal was to generate high-quality, quantitative information, we chose to purify the domains from large-scale bacterial cultures, rather than to use high-throughput methods. After isolating each domain, we assessed its purity by electrophoresis (Supplementary Fig. 2) and its aggregation state by gel filtration (Supplementary Table 2). Soluble protein was obtained for 140 of the 160 constructs. All but one of the remaining constructs were produced as insoluble protein, purified under denaturing conditions, and subsequently refolded (Supplementary Table 1). Notably, 13 of the 14 SH2 domains derived from the STAT and SOCS families of proteins required refolding, and eight did not contain any monomeric protein. It is likely that any interactions observed with these aggregated domains are nonspecific in nature.

ErbB peptides

In order to focus on physiologically relevant interactions, we searched the literature for experimentally verified sites of tyrosine phosphorylation on the ErbB receptors. We uncovered 12 sites on EGFR, 6 on ErbB2 and 11 on ErbB3. No experimentally verified sites were found on ErbB4. We elected, however, to include four tyrosines that were predicted by similarity to be sites of autophosphorylation (http://us.expasy.org/sprot/). As surrogates for the activated receptors, we synthesized 17–19-residue, phosphotyrosine

[1]Department of Chemistry and Chemical Biology, and [2]Program in Biophysics, Harvard University, Cambridge, Massachusetts 02138, USA.
*These authors contributed equally to this work.

(pY)-containing peptides (Supplementary Table 3). In four instances, two sites of phosphorylation lie within two residues of each other, prompting us to synthesize doubly phosphorylated peptides in addition to the singly phosphorylated ones. We also prepared non-phosphorylated versions of each peptide to serve as controls. A fluorescent dye was appended to the amino terminus of each peptide to visualize binding.

Protein microarrays

While it is a daunting task to study the interaction of 159 proteins with 66 peptides, protein microarray technology facilitates such large-scale analyses[7,8]. We set out to fabricate microarrays of the purified domains that could then be queried with each fluorescent peptide. To reduce experimental variation and to expedite the processing of hundreds of arrays, we developed a strategy to produce protein microarrays in microtitre plates (see Methods). Samples were printed in duplicate and a small amount of cyanine-5 (Cy5)-labelled albumin was introduced into each sample to facilitate image analysis (Fig. 1a).

Although probing a protein array with a molecule of interest identifies a subset of interactions, the resulting information can be misleading. We have learned not to rely on data produced using a single concentration of a solution-phase probe, nor to rely on the intensity of individual spots, which does not always reflect the strength of the interaction. To circumvent these limitations, protein arrays were probed with eight concentrations of each peptide, ranging from 10 nM to 5 μM (Fig. 1b), yielding saturation binding curves for each peptide–protein pair. Under equilibrium conditions, the mean fluorescence of duplicate spots (F_{obs}) can be described by equation (1),

$$F_{obs} = \frac{F_{max}[pep]}{K_D + [pep]} \quad (1)$$

where F_{max} is the maximum fluorescence at saturation, [*pep*] is the total peptide concentration, and K_D is the apparent equilibrium dissociation constant. For each peptide, we fit all 159 curves, one for each single or tandem domain. Because nonspecific binding increases linearly with peptide concentration, whereas specific binding saturates, we scored as 'specific' those interactions that fit well to equation (1) ($R^2 \geq 0.9$), with a K_D below 2 μM and an F_{max} at least twofold higher than the mean fluorescence of control spots. The curves that met these criteria for ErbB2 pY1139 are shown in Fig. 1c. We also identified weaker interactions, which we recorded as '> 2 μM'. Following this strategy, we performed the quantitative analysis illustrated in Fig. 1 for all 66 peptides. Five of the peptides exhibited high levels of background binding (Supplementary Table 3), precluding an analysis of their interactions. High quality data were obtained for the other 61 peptides (Supplementary Table 4). These data are also available on our website (http://www.cgr.harvard.edu/macbeath/data/data.html) and will be updated as additional sites are identified. Of the 5,247 interactions that we measured with pY-containing peptides, 353 (6.7%) exhibit K_D values below 2 μM. If we consider 'active' a domain that recognizes one or more phosphopeptides, we found that at least 102 of the 115 SH2-containing constructs (89%) and 27 of the 44 PTB-containing constructs (61%) are active on the arrays. It is likely many of the 'inactive' PTB domains are functional, but their role is not to bind to sites of tyrosine phosphorylation[9].

To assess the accuracy of our method, we measured K_D values for eight domain–peptide pairs using surface plasmon resonance (SPR). In each case, the free energy of binding (ΔG) calculated from our microarray experiments matched within 5% the value measured by SPR (Fig. 1d).

Quantitative protein interaction networks

Using the data derived from this large-scale analysis, we constructed a

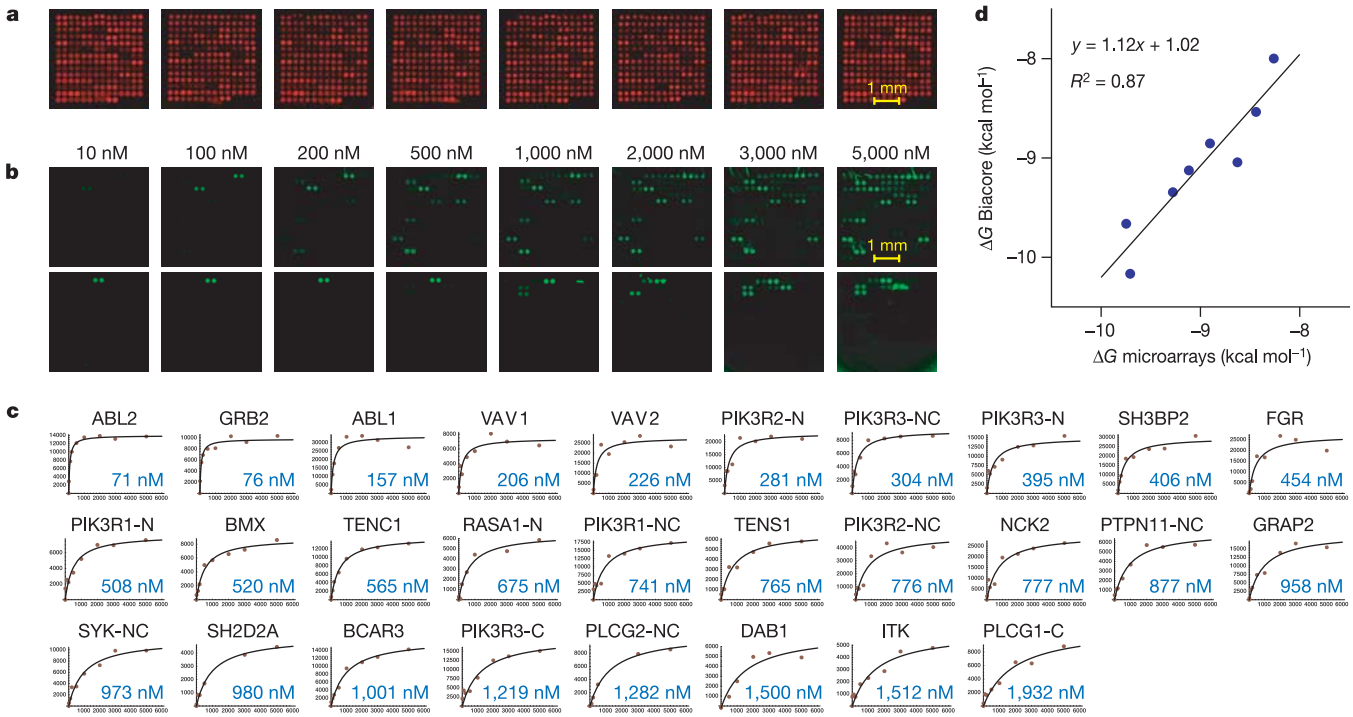

Figure 1 | Measuring the binding affinity of SH2/PTB domains for phosphopeptides derived from the ErbB receptors using protein microarrays. **a**, Fluorescent images of eight identical SH2/PTB microarrays in separate wells of a 96-well microtitre plate. The fluorescence arises from a trace amount of Cy5-labelled BSA that was added to each protein before arraying. **b**, Fluorescent images of SH2/PTB microarrays, probed with eight different concentrations of a 5(6)-TAMRA-labelled phosphopeptide derived from ErbB2 (pY1139). **c**, Plots showing fluorescence as a function of peptide concentration for 28 high-affinity interactions. The data were fit to equation (1) to determine the apparent K_D. **d**, Comparison of the free energy of binding for eight domain–peptide interactions measured using protein microarrays with those measured using SPR (Biacore).

graphical representation of pY-mediated recruitment (Fig. 2). These diagrams provide a system-level view of the ErbB receptors, showing biophysical interactions between signalling proteins and known sites of tyrosine phosphorylation. Which proteins are actually recruited in a given cell will depend on many factors, including the effective concentrations of both the activated receptors and the signalling proteins. These diagrams should therefore be viewed as quantitative maps of the receptors, rather than a depiction of protein recruitment in any specific cell type or state.

To evaluate how well our microarray experiments recapitulate known interactions, we compiled a list of previously reported interactions between SH2/PTB-containing proteins and the ErbB receptors (Supplementary Table 5). For interactions with EGFR and ErbB2, we relied on hand-curated databases (ref. 10 and http://proteome.incyte.com/); for ErbB3 and ErbB4, we surveyed the literature ourselves. Overall, our arrays detected 43 of the 65 previously reported interactions. For example, we observed that peptides derived from EGFR were able to bind strongly ($K_D < 2\,\mu M$) to the SH2/PTB domains of Crk, Grb2, Nck1, PI3Kα (also known as PIK3R1), PI3Kβ (also known as PIK3R2), PLC-γ1 (also known as PLCG1), PLC-γ2 (also known as PLCG2), Shp2 (also known as PTPN11), RasGAP (also known as RASA1), Shc1, Shc3, Syk and Vav1, and weakly to the SH2 domains of Grb10, Grb7, Nck2, Shp1 (also known as PTPN6), Nsp1 (also known as SH2D3A), Socs1, Stat1, Stat3, Vav2 and Vav3. Many of the known interactions that were not detected were members of the STAT and SOCS families of

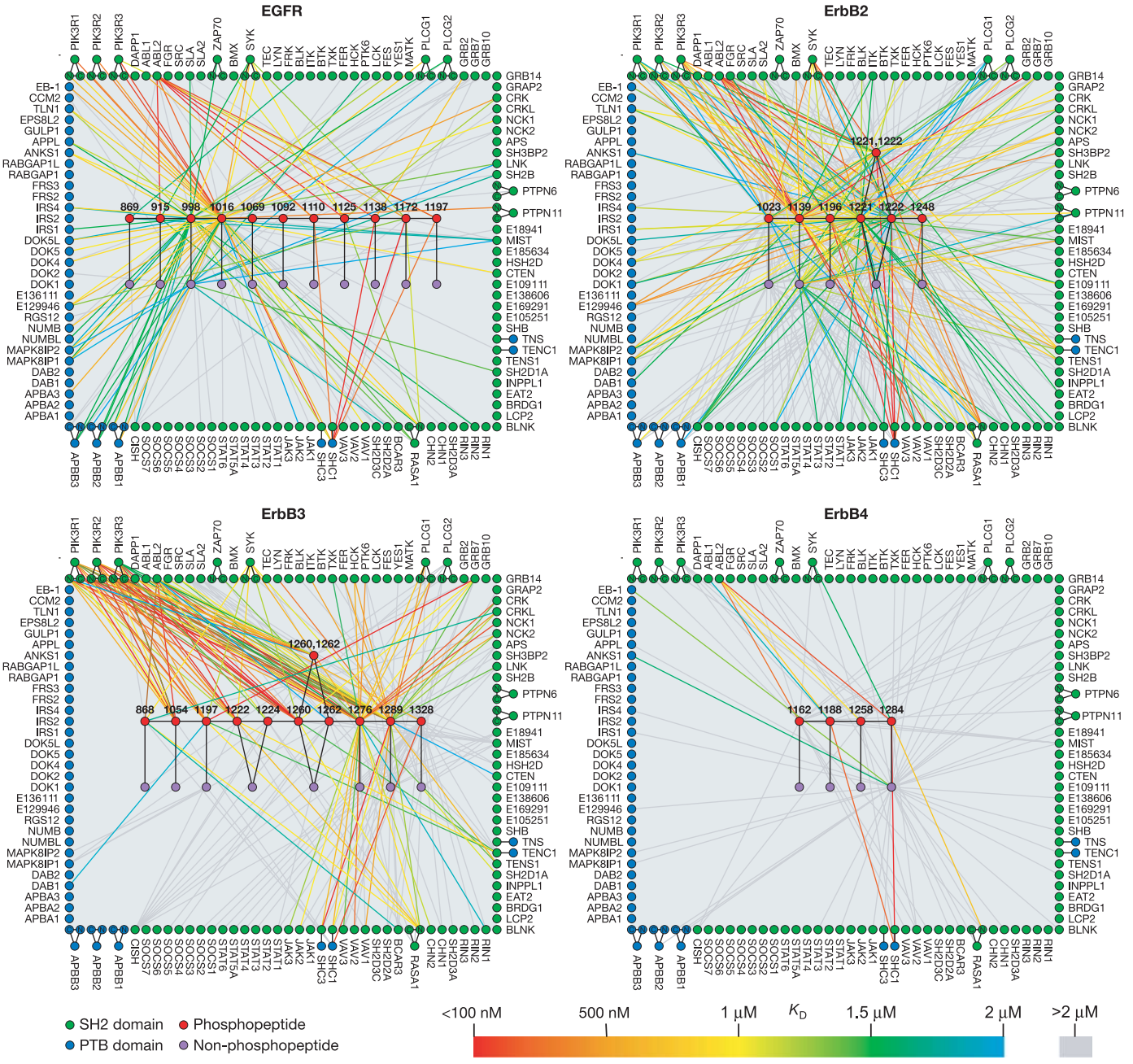

Figure 2 | Quantitative protein interaction networks for the four human ErbB receptors. Red circles represent phosphopeptides; purple circles represent the non-phosphorylated version of each phosphopeptide; green circles represent SH2 domains; and blue circles represent PTB domains. Lines connecting peptides to domains indicate observed interactions, coloured according to the affinity of the interaction (see legend). Red circles labelled with two numbers represent doubly phosphorylated peptides. The green or blue circles that lie outside the rectangle of individual domains represent tandem domains. Black lines connect the tandem domains to their corresponding individual domains.

proteins that contained a high percentage of aggregated protein. In addition, because many reported interactions are based on co-purification experiments, some may not depend on SH2/PTB domains or may be mediated by bridging proteins.

We also compared our observed interactions with those predicted by Scansite 2.0 (refs 11, 12), a program that uses the consensus binding information provided by oriented-peptide library screens[4,5] to predict interactions between domains and sequences of interest. We scanned the 14 SH2 domains and one PTB domain available in Scansite against the peptide sequences used in our study. For comparison purposes, we considered only the 11 domains that bind more than one peptide on the arrays and are predicted by Scansite to recognize more than one peptide in our study. When run on its lowest stringency setting, Scansite predicts 56% of the strong ($K_D < 2\,\mu M$) interactions that we observe, and 47% of all interactions, indicating that many peptides with sequences that do not conform to consensus motifs are recognized by these domains. For example, Scansite does not predict an interaction between the SH2 domain of Abl1 and ErbB2 pY1139, yet the arrays show a strong interaction ($K_D = 157\,nM$), which we also confirm by SPR ($K_D = 140\,nM$). Conversely, 51% of the Scansite hits are not observed on the arrays. Overall, we find that Scansite predictions closely match the microarray data for the PTB domain of Shc1 and the SH2 domains of PI3K and Grb2, but the overlap is much less for other domains (Fig. 3). We attribute this difference to the non-competitive format of the microarray experiment in which only physiologically relevant sequences are assessed without interference from tight-binding but irrelevant peptides.

Not surprisingly, the SH2/PTB microarrays uncovered many strong interactions ($K_D < 2\,\mu M$) that have not been reported previously: 32 with EGFR, 48 with ErbB2, 33 with ErbB3 and 3 with ErbB4 (Supplementary Table 5). For example, the arrays revealed that the SH2 domain of v-crk avian sarcoma virus CT10-homologue-like protein (CrkL) recognizes phosphopeptides derived from ErbB2 and ErbB3. Consistent with this observation, we found that CrkL becomes phosphorylated on Y207 in A431 squamous carcinoma cells and in MDA-MB-468 breast carcinoma cells within 1 min of stimulation with EGF, and in MDA-MB-468 and T-47D breast carcinoma cells within 5 min of stimulation with heregulin-β1 (HRG-β1, a ligand for ErbB3; Supplementary Fig. 3a, b). The slower phosphorylation of CrkL in response to HRG-β1 stimulation is probably due to the substantially lower number of ErbB3 molecules in MDA-MB-468 and T-47D cells relative to the number of EGFR and ErbB2 molecules in A431 and MDA-MB-468 cells[13].

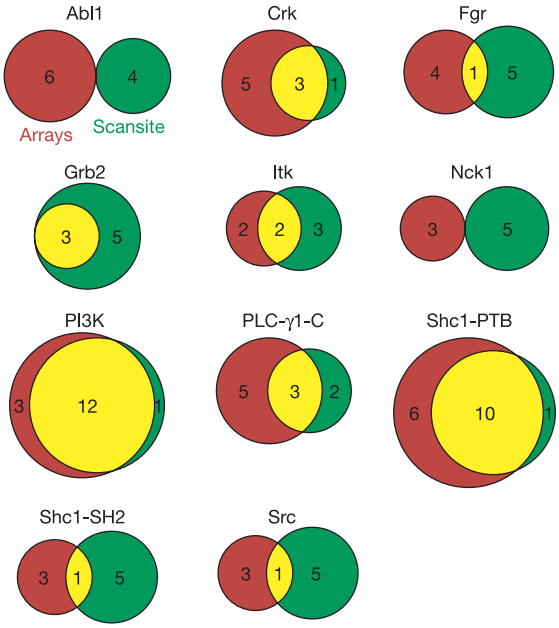

Figure 3 | Venn diagrams for 11 of the 15 SH2/PTB domains available in Scansite 2.0. For each domain, the red circle represents the phosphopeptides that are observed to bind to that domain using protein microarrays and the green circle represents the phosphopeptides that are predicted to bind to that domain by Scansite 2.0 (run on its lowest stringency setting).

System-level properties of ErbB receptors

Whereas our arrays provide a list of previously unrecognized biochemical interactions, as well as data that can fuel efforts to build computational models of signal transduction[14], they also offer an unbiased, system-level view of the ErbB receptors that was previously unavailable. The networks of Fig. 2 show that EGFR and ErbB3 each have two sites (Y998 and Y1016 on EGFR, and Y1276 and Y1289 on ErbB3) that can engage in many high-affinity interactions. These tyrosines may serve as 'multifunctional docking sites'[15] that have different roles depending on the relative concentrations of their target proteins. The other phosphotyrosines on these receptors are markedly more selective and presumably serve specialized functions.

In contrast, ErbB2 features many highly promiscuous sites. When considering only strong interactions ($K_D < 2\,\mu M$), the sites on ErbB2 bind over 17 different proteins on average, whereas those on EGFR, ErbB3 and ErbB4 bind 7.2, 8.8 and 2.3 proteins on average, respectively. Unlike the other receptors, ErbB2 does not recognize an extracellular ligand, but instead functions primarily as a heterodimerization partner. It has been shown that ErbB2 quantitatively increases both the amplitude and duration of EGFR signalling by increasing the ratio of active kinase to ligand and by inhibiting the downregulation of EGFR[16]. Our data suggest that ErbB2 may also qualitatively expand the diversity of signalling by recruiting proteins that the other receptors cannot. The sparse connections to ErbB4 may indicate that this receptor serves a more specialized function than the others. It is likely, however, that we are missing sites of tyrosine phosphorylation on ErbB4, and that even these four sites may not be physiologically relevant.

The networks of Fig. 2 show that, in principle, some phosphotyrosines can recruit many different proteins. For example, pY1139 of ErbB2 interacts strongly ($K_D < 2\,\mu M$) with 24 different proteins. Although SH2/PTB-containing proteins are produced at different levels in different cells, many that are able to bind to the same site are co-expressed[17]. Given the high connectivity of the network, it is possible that some sites serve more than one function within the same cell at the same time. The six tightest binders of ErbB2 pY1139 have K_D values that fall within less than fourfold of each other. If two of these proteins are present in the vicinity of the activated receptor at similar concentrations, both may be recruited and activated.

It is tempting to speculate that some sites perform different functions based on the location of the receptor in the cell. All of the interactions that we have investigated by SPR display half-lives of less than 10 s, which indicates that interactions based solely on SH2 domains are highly dynamic[18]. When a receptor is initially activated, proteins that function at the cell surface (such as PI3K and PLC-γ) are recruited. As the receptor is internalized, the repertoire of proteins it encounters changes, leading to the recruitment of a different set of proteins. A single site could thus serve more than one function in the same cell, at the same time, but in different locations. It is now clear that internalized receptors continue to signal[19] and that some proteins preferentially associate with surface receptors, whereas others bind almost exclusively to intracellular receptors[20]. The protein recruitment profile of a cell is therefore likely to depend not only on which proteins are co-expressed, but on receptor trafficking and partitioning as well.

Networks at different affinity thresholds

Most interaction networks reported to date are boolean: proteins

either 'interact' or 'don't interact'[21–24]. Such networks represent a single slice through interaction space, where the slice is made at an affinity threshold defined by the assay. Because we quantified every interaction, we can view different slices through the network at defined affinity thresholds. To a first approximation, these slices highlight interactions that are relevant at different concentrations of activated receptor. At low levels of receptor activation, or in cells in which the receptor is expressed at low levels, only the highest affinity interactions are important. At higher levels of activated receptor, however, the lower affinity interactions are also relevant (assuming the receptor is present in excess). Although the protein recruitment profile varies from one cell to the next based on affinities and concentrations, the general principles that arise from this analysis provide insight into the intrinsic properties of each receptor. Figure 4 shows four slices through the ErbB network at different affinity thresholds. ErbB4 was excluded due to the paucity of data. EGFR, ErbB2 and ErbB3, however, exhibit intriguing differences. Notably, the ErbB3 network changes very little as the threshold is lowered, whereas EGFR and ErbB2 become much more promiscuous (Fig. 4a, b). The biological implication of this finding is that cells should be less sensitive to changes in the levels of ErbB3 relative to EGFR and ErbB2. Consistent with this prediction, EGFR and ErbB2 expression levels vary more across normal human tissues than do those of ErbB3 (ref. 17).

The oncological implication is equally intriguing. The threshold slices of Fig. 4 suggest that elevated levels of ErbB3 should primarily induce stronger signalling through pathways that are normally activated by low levels of ErbB3. In contrast, elevated levels of EGFR or ErbB2 should induce signalling through alternative pathways that are not activated at lower levels. Interestingly, overexpression, gene amplification, or overactivation of EGFR and ErbB2 are frequently observed in human cancers[25–30], whereas there is no evidence for gene amplification of ErbB3, and overexpression is limited[1]. We propose that the high oncogenic potential of EGFR and ErbB2 arises, at least in part, from their ability to turn on different pathways when overexpressed, rather than simply from stronger signalling through their primary pathways. Secondary signalling proteins that are only recruited by overexpressed receptors may also engage in cross-talk with primary signalling proteins, further altering the state of the cell. This hypothesis suggests that proteins involved in the most critical secondary pathways may serve as more selective targets for cancer chemotherapy.

We also asked what each receptor can do that the others cannot, and how this changes with receptor levels. Figure 4c shows Venn diagrams indicating the number of proteins that can be recruited to each receptor at each affinity threshold. When only the tightest interactions are considered ($K_D < 500$ nM), there is only one protein that binds solely to EGFR (Shp2/PTPN11). With this one exception,

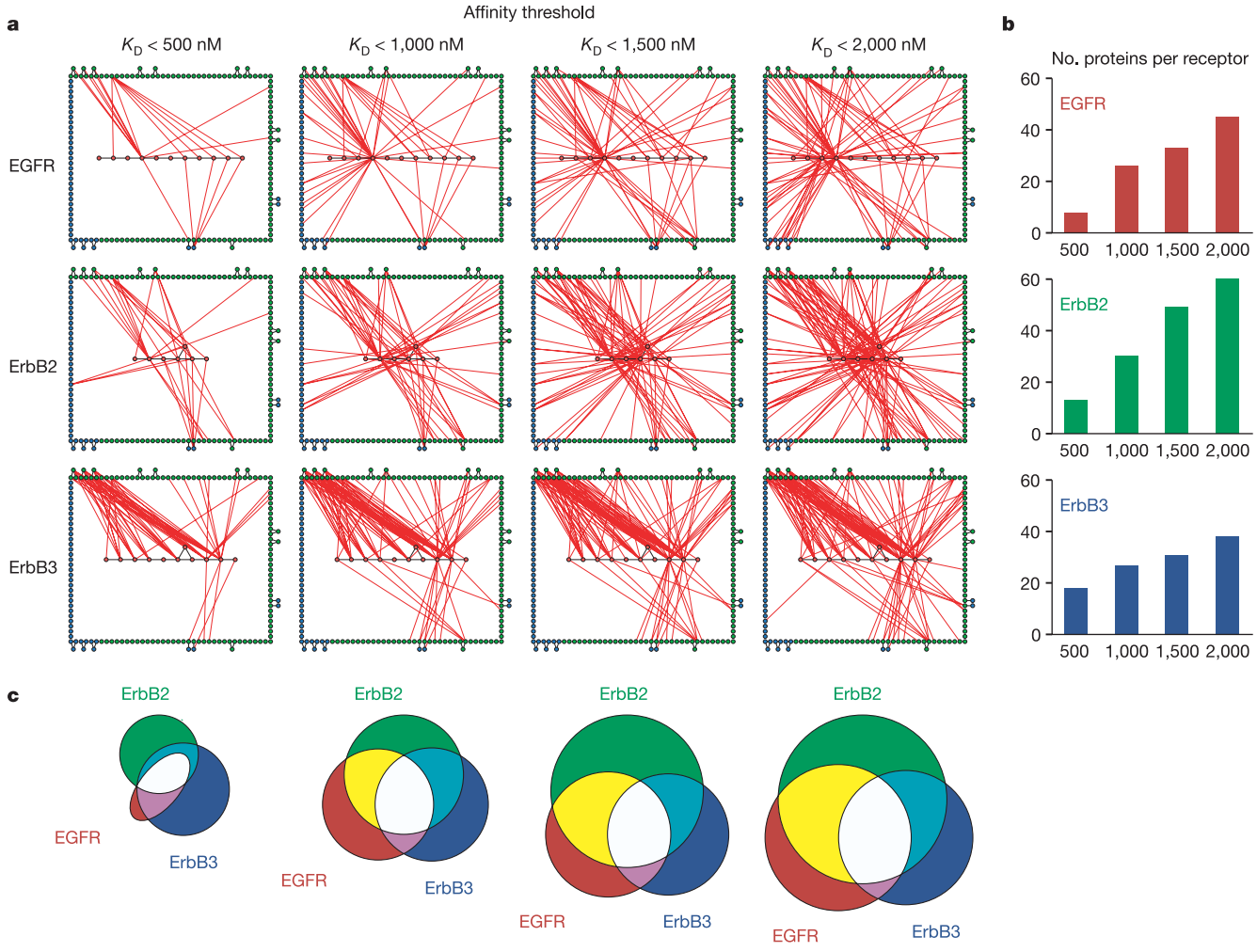

Figure 4 | A system-level view of EGFR, ErbB2 and ErbB3 at different affinity thresholds. **a**, Four views of the ErbB interaction networks, each at a different affinity threshold. At each threshold, only interactions with a K_D below the indicated value are shown. **b**, Plots illustrating the number of proteins that bind to each receptor at each affinity threshold. The *y* axis shows the number of different proteins that contain at least one SH2 or PTB domain with a K_D value below that indicated on the *x* axis (nM). **c**, Venn diagrams, drawn to scale, illustrating the number of proteins that are recruited to each receptor at each affinity threshold.

the proteins that bind to EGFR constitute a small subset of those that bind to ErbB3. It is only at high receptor levels that EGFR is able to recruit proteins that ErbB2 and ErbB3 cannot. Even at low concentration, on the other hand, ErbB2 recruits proteins that ErbB3 cannot. Our results indicate that, under stringent conditions, the ErbB2–ErbB3 complex is broadest in scope, followed by EGFR–ErbB3, EGFR–ErbB2 and finally EGFR–EGFR. Consistent with this observation, ErbB2–ErbB3 is the most transforming receptor complex[31,32] and the mitogenicity of these homo- and heterodimers tracks perfectly with their breadth of signalling[33].

What proteins are common to all three receptors? In general, they are the ones that initiate canonical signalling pathways: PI3K, which regulates cell cycle progression, cell growth, cytoskeletal rearrangements and vesicular transport; Shc1, Crk and RasGAP, which feed into the MAPK cascade to control proliferation and motility; and Syk and PLC-γ, which regulate cytoskeleton–membrane interactions. We were surprised, however, to find that Abl1 and Abl2 bind with high affinity to sites on all four receptors, even though they are not typically considered in the context of ErbB signalling. Nevertheless, we found that Abl1 is phosphorylated on Y412 in A431 cells within 1 min of treatment with EGF (Supplementary Fig. 3a, c). Abl1 normally controls cytoskeletal rearrangements[34] and thus may regulate cell migration or adhesion in response to EGF and its relatives.

Discussion

By performing a comprehensive analysis of SH2/PTB-mediated interactions with the ErbB receptors, we uncovered a quantitative network that reveals the ability of each receptor to recruit signalling proteins upon activation. In addition to confirming 43 previously recognized interactions, we identified 116 new biophysical interactions and provided evidence for the physiological relevance of several of them. Most importantly, the network defined by this effort provides system-level insight into ErbB function and reveals a surprising property of receptor tyrosine kinases: they differ in the extent to which they become more promiscuous when overexpressed. This observation highlights the importance of collecting quantitative information on protein–protein interactions. It is our hope that modelling studies based on a marriage of our biophysical measurements with quantitative cell biological data will prove useful in predicting normal cellular behaviour, as well as how best to intervene when signalling goes awry.

METHODS

Cloning of SH2 and PTB domains. When we began this effort, a search of three databases—SMART[35], Pfam[36] and Ensembl (http://www.ensembl.org/)—yielded a list of 109 SH2 domains, 44 PTB domains, 10 tandem SH2 domains and 3 tandem PTB domains. Sequences encoding these domains were amplified from human cDNA and transferred into pENTR/D-TOPO (Invitrogen) by topoisomerase I-mediated directional cloning. Each clone was verified by DNA sequencing.

Production and purification of recombinant proteins. SH2/PTB coding sequences were transferred into a Gateway-compatible *Escherichia coli* expression vector (pET-32-DEST). Recombinant proteins were produced in 500-ml cultures and purified by immobilized metal affinity chromatography. Purified proteins were dialysed against buffer A (300 mM NaCl, 50 mM Na_2PO_4, pH 8) and glycerol was added to a final concentration of 20% (v/v).

Peptide synthesis. Peptides were synthesized on the solid phase (50 μmol scale) using Fmoc chemistry. Peptides were labelled on their amino termini with 5-(and-6)-carboxytetramethylrhodamine (5(6)-TAMRA) before deprotection and cleavage. All peptides were purified by reverse phase high-performance liquid chromatography (HPLC).

Fabrication and processing of protein microarrays. Purified SH2/PTB domains were spotted in duplicate at a concentration of 40 μM onto aldehyde-modified glass substrates (112.5 mm × 74.5 mm × 1 mm) using a piezoelectric microarrayer. Ninety-six identical arrays were fabricated in a 12 × 8 pattern to match the spacing of a microtitre plate. Each array consisted of a 14 × 14 pattern of spots, with a 250 μm pitch. After a 1 h incubation, the glass was attached to a bottomless 96-well plate using an intervening silicone gasket. Immediately before use, the plates were quenched with buffer B (20 mM HEPES, 100 mM KCl, 0.1% Tween-20, pH 7.8) containing 1% BSA (w/v). Arrays were probed with 5(6)-TAMRA-labelled peptides, dissolved in buffer B. After a 30-min incubation, the arrays were washed with buffer B, rinsed with ddH_2O, and spun upside down to remove residual water.

Scanning and analysis of microarrays. Protein microarrays were scanned at 10-μm resolution using an LS400 scanner (Tecan). Spots were defined using the Cy-5 image and the mean fluorescence of each spot was calculated from the 5(6)-TAMRA image. Concentration-dependent measurements were fit to equation (1) for each domain–peptide pair and the resulting data were displayed graphically using Cytoscape 2.1 (http://www.cytoscape.org/).

Surface plasmon resonance. SPR studies were performed using a Biacore 3000. SH2 domains were produced as glutathione *S*-transferase (GST) fusion proteins and captured on the chip by an immobilized anti-GST antibody. Peptides were introduced in the solution phase and binding was measured under equilibrium conditions.

Cell culture and immunoblots. Cells were grown to approximately 70% confluence, serum-starved for 24 h, treated with either 100 ng ml^{-1} EGF or 10 nM heregulin-β1, and analysed by immunoblotting. Technical details of all experiments are provided in the Supplementary Methods.

Received 7 July; accepted 30 August 2005.
Published online 6 November 2005.

1. Yarden, Y. & Sliwkowski, M. X. Untangling the ErbB signalling network. *Nature Rev. Mol. Cell Biol.* **2,** 127–137 (2001).
2. Sadowski, I., Stone, J. C. & Pawson, T. A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus P130gag-fps. *Mol. Cell. Biol.* **6,** 4396–4408 (1986).
3. Kavanaugh, W. M. & Williams, L. T. An alternative to SH2 domains for binding tyrosine-phosphorylated proteins. *Science* **266,** 1862–1865 (1994).
4. Songyang, Z. *et al.* SH2 domains recognize specific phosphopeptide sequences. *Cell* **72,** 767–778 (1993).
5. Songyang, Z. *et al.* Specific motifs recognized by the SH2 domains of Csk, 3BP2, fps/fes, GRB-2, HCP, SHC, Syk, and Vav. *Mol. Cell. Biol.* **14,** 2777–2785 (1994).
6. Hatada, M. H. *et al.* Molecular basis for interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor. *Nature* **377,** 32–38 (1995).
7. MacBeath, G. & Schreiber, S. L. Printing proteins as microarrays for high-throughput function determination. *Science* **289,** 1760–1763 (2000).
8. Zhu, H. *et al.* Global analysis of protein activities using proteome chips. *Science* **293,** 2101–2105 (2001).
9. Yan, K. S., Kuti, M. & Zhou, M. M. PTB or not PTB—that is the question. *FEBS Lett.* **513,** 67–70 (2002).
10. Peri, S. *et al.* Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res.* **13,** 2363–2371 (2003).
11. Yaffe, M. B. *et al.* A motif-based profile scanning approach for genome-wide prediction of signalling pathways. *Nature Biotechnol.* **19,** 348–353 (2001).
12. Obenauer, J. C., Cantley, L. C. & Yaffe, M. B. Scansite 2.0: Proteome-wide prediction of cell signalling interactions using short sequence motifs. *Nucleic Acids Res.* **31,** 3635–3641 (2003).
13. Bowers, G. *et al.* The relative role of ErbB1–4 receptor tyrosine kinases in radiation signal transduction responses of human carcinoma cells. *Oncogene* **20,** 1388–1397 (2001).
14. Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D. & Muller, G. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nature Biotechnol.* **20,** 370–375 (2002).
15. Ponzetto, C. *et al.* A multifunctional docking site mediates signalling and transformation by the hepatocyte growth factor/scatter factor receptor family. *Cell* **77,** 261–271 (1994).
16. Hendriks, B. S., Orr, G., Wells, A., Wiley, H. S. & Lauffenburger, D. A. Parsing ERK activation reveals quantitatively equivalent contributions from epidermal growth factor receptor and HER2 in human mammary epithelial cells. *J. Biol. Chem.* **280,** 6157–6169 (2005).
17. Su, A. I. *et al.* Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl Acad. Sci. USA* **99,** 4465–4470 (2002).
18. Felder, S. *et al.* SH2 domains exhibit high-affinity binding to tyrosine-phosphorylated peptides yet also exhibit rapid dissociation and exchange. *Mol. Cell. Biol.* **13,** 1449–1455 (1993).
19. Haugh, J. M. Localization of receptor-mediated signal transduction pathways: the inside story. *Mol. Interv.* **2,** 292–307 (2002).
20. Burke, P., Schooler, K. & Wiley, H. S. Regulation of epidermal growth factor receptor signalling by endocytosis and intracellular trafficking. *Mol. Biol. Cell* **12,** 1897–1910 (2001).
21. Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415,** 141–147 (2002).
22. Ho, Y. *et al.* Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415,** 180–183 (2002).

23. Li, S. *et al.* A map of the interactome network of the metazoan *C. elegans*. *Science* **303,** 540–543 (2004).
24. Uetz, P. *et al.* A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* **403,** 623–627 (2000).
25. Gorgoulis, V. *et al.* Expression of EGF, TGF-α and EGFR in squamous cell lung carcinomas. *Anticancer Res.* **12,** 1183–1187 (1992).
26. Irish, J. C. & Bernstein, A. Oncogenes in head and neck cancer. *Laryngoscope* **103,** 42–52 (1993).
27. Moscatello, D. K. *et al.* Frequent expression of a mutant epidermal growth factor receptor in multiple human tumors. *Cancer Res.* **55,** 5536–5539 (1995).
28. Ross, J. S. & Fletcher, J. A. The HER-2/neu oncogene in breast cancer: prognostic factor, predictive factor, and target for therapy. *Stem Cells* **16,** 413–428 (1998).
29. Slamon, D. J. *et al.* Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **235,** 177–182 (1987).
30. Wong, A. J. *et al.* Structural alterations of the epidermal growth factor receptor gene in human gliomas. *Proc. Natl Acad. Sci. USA* **89,** 2965–2969 (1992).
31. Alimandi, M. *et al.* Cooperative signalling of ErbB3 and ErbB2 in neoplastic transformation and human mammary carcinomas. *Oncogene* **10,** 1813–1821 (1995).
32. Wallasch, C. *et al.* Heregulin-dependent regulation of HER2/neu oncogenic signalling by heterodimerization with HER3. *EMBO J.* **14,** 4267–4275 (1995).
33. Pinkas-Kramarski, R. *et al.* Diversification of Neu differentiation factor and epidermal growth factor signalling by combinatorial receptor interactions. *EMBO J.* **15,** 2452–2467 (1996).
34. Woodring, P. J., Hunter, T. & Wang, J. Y. Regulation of F-actin-dependent processes by the Abl family of tyrosine kinases. *J. Cell Sci.* **116,** 2613–2626 (2003).
35. Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P. SMART, a simple modular architecture research tool: identification of signalling domains. *Proc. Natl Acad. Sci. USA* **95,** 5857–5864 (1998).
36. Bateman, A. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **28,** 263–266 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank B. Schoeberl for comments on the manuscript, J. Grudzien for development work on the silicone gaskets, and the Bauer Center for Genomics Research at Harvard University for support with instrumentation and automation. This work was supported by awards from the W. M. Keck Foundation and the Arnold and Mabel Beckman Foundation. R.B.J. is the recipient of a Ruth L. Kirschstein National Research Service Award (NIH), A.G. is the recipient of an NSF Graduate Research Fellowship, and J.A.K. is the recipient of a Howard Hughes Medical Institute Predoctoral Fellowship in the Biological Sciences.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to G.M. (macbeath@chemistry.harvard.edu).