nature biotechnology

Direct identification of ligand-receptor interactions on living cells and tissues

Andreas P Frei[1–3], Ock-Youm Jeon[4,10], Samuel Kilcher[3,5,10], Hansjoerg Moest[1,2,6], Lisa M Henning[1,2], Christian Jost[3,7], Andreas Plückthun[7,8], Jason Mercer[5], Ruedi Aebersold[1,8,9], Erick M Carreira[4] & Bernd Wollscheid[1,2]

Many cellular responses are triggered by proteins, drugs or pathogens binding to cell-surface receptors, but it can be challenging to identify which receptors are bound by a given ligand. Here we describe TRICEPS, a chemoproteomic reagent with three moieties—one that binds ligands containing an amino group, a second that binds glycosylated receptors on living cells and a biotin tag for purifying the receptor peptides for identification by quantitative mass spectrometry. We validated this ligand-based, receptor-capture (LRC) technology using insulin, transferrin, apelin, epidermal growth factor, the therapeutic antibody trastuzumab and two DARPins targeting ErbB2. In some cases, we could also determine the approximate ligand-binding sites on the receptors. Using TRICEPS to label intact mature vaccinia viruses, we identified the cell surface proteins AXL, M6PR, DAG1, CSPG4 and CDH13 as binding factors on human cells. This technology enables the identification of receptors for many types of ligands under near-physiological conditions and without the need for genetic manipulations.

Identifying the receptors of key ligands can often lead to new research directions and provide valuable mechanistic information about signal transduction, drug action or off-target effects. However, the unbiased and unambiguous identification of ligand-receptor interactions remains a daunting task despite the emergence of mass spectrometry–based technologies for the identification of protein-protein[1,2] and small molecule–protein[3–5] interactions in cell lysates. This is mostly due to the transient nature of such interactions and the problems inherent in the analysis of plasma membrane proteins that are often hydrophobic and present in relatively low abundance[6,7]. Furthermore, cell surface proteins typically need to be embedded in their natural environment, that is, within living cells or tissues, to exhibit their characteristic binding properties[8,9].

We and others have previously used bifunctional reagents containing hydrazide and aminooxy active groups to label carbohydrate-containing proteins on the surface of live cells with biotin[10,11]. Subsequently, the purification and mass spectrometric analysis of corresponding peptides allows for the characterization of the cell surface proteome[11,12]. As an extension of this strategy, we hypothesized that specifically designed trifunctional reagents would allow the capture of ligand-based receptors bound to ligands on the surface of live cells followed by identification of the captured receptors using quantitative mass spectrometry. We designed and synthesized biocompatible chemoproteomic reagents (**Fig. 1a** and **Supplementary Note**) incorporating three independent functionalities: an *N*-hydroxysuccinimide ester for conjugating the reagent to a ligand containing a free amino group; a protected hydrazine that reacts with aldehydes in carbohydrates on glycoprotein receptors; and a biotin group for the affinity purification of the captured glycopeptides for mass spectrometric analysis[13]. Special emphasis was also placed on water solubility, conformational flexibility, spacer length and compatibility with gentle yet selective conditions for the ligand coupling and receptor-capture reactions. We call these trifunctional chemoproteomic reagents 'TRICEPS'.

The *N*-hydroxysuccinimide ester on TRICEPS nonspecifically couples the reagent to primary amines, the α-NH_2 group of the N terminus or the ε-NH_2 group of lysine side chains, in protein or peptide ligands of interest, under physiological conditions[14]. Typically, several primary amines are present in ligands. Simultaneous coupling to all or most available groups might interfere with subsequent receptor binding, whereas lower labeling ratios are expected to minimally impair the majority of ligands (**Supplementary Fig. 1**). After reaction of the *N*-hydroxysuccinimide ester, TRICEPS retained the ability to efficiently capture aldehydes on the surface of living cells (**Fig. 1b**). Aldehydes are generally absent from cell surfaces but can be introduced into the carbohydrates of cell surface glycoproteins through exposure of cells to gentle oxidizing conditions with sodium-metaperiodate[10,11,15]. In contrast to other combinations of reactive and protecting groups, the trifluoroacetylated hydrazines used in TRICEPS can undergo condensation with carbohydrate-derived aldehydes under essentially neutral conditions without prior removal of the protecting group. This is one of the key features of TRICEPS as the trifluoroacetamide protective group prevents unwanted side reactions

[1]Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, Switzerland. [2]NCCR Neuro Center for Proteomics, University of Zurich, Switzerland. [3]PhD Program in Molecular Life Sciences (MLS), University of Zurich/ETH Zurich, Switzerland. [4]Department of Chemistry and Applied Biosciences, Laboratory of Organic Chemistry, ETH Zurich, Switzerland. [5]Department of Biology, Institute of Biochemistry, ETH Zurich, Switzerland. [6]Department of Health and Technology, Institute of Food, Nutrition and Health, ETH Zurich, Switzerland. [7]Department of Biochemistry, University of Zurich, Switzerland. [8]Faculty of Science, University of Zurich, Switzerland. [9]Competence Center for Systems Physiology and Metabolic Diseases, Zurich, Switzerland. [10]These authors contributed equally to this work. Correspondence should be addressed to B.W. (bernd.wollscheid@imsb.biol.ethz.ch).

Received 6 April; accepted 8 August; published online 16 September 2012; doi:10.1038/nbt.2354

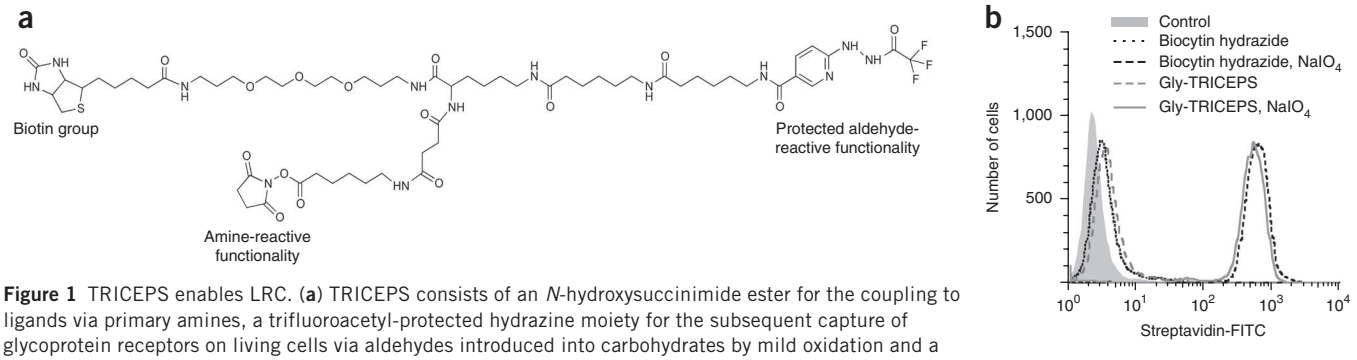

Figure 1 TRICEPS enables LRC. (**a**) TRICEPS consists of an *N*-hydroxysuccinimide ester for the coupling to ligands via primary amines, a trifluoroacetyl-protected hydrazine moiety for the subsequent capture of glycoprotein receptors on living cells via aldehydes introduced into carbohydrates by mild oxidation and a biotin moiety for the affinity purification of captured glycopeptides. Flexibility is provided by the stochastic coupling of TRICEPS to topologically different sites on ligands, the long (~48 Å) and flexible spacer between the reactive moieties and the presence of typically more than one aldehyde in a single carbohydrate structure after mild oxidation. Alternative ligand-reactive functionalities may be implemented to further broaden the possible application (e.g., for small-molecule ligands). (**b**) Labeling of living cells with TRICEPS through the aldehyde-reactive functionality. Commercially available biocytin hydrazide (an unprotected aldehyde-reactive biotin derivative) was used as a reference. Amine-reactive moieties of TRICEPS were quenched with glycine and reagents were reacted with Jurkat T lymphocytes with or without prior oxidation of the cells with sodium metaperiodate. Biotinylation was detected by flow cytometric analysis of ungated cell populations after staining with streptavidin-FITC.

between the nucleophilic hydrazine and electrophiles such as active ester groups in the reagent itself. As the condensation reaction can occur under neutral conditions, this eliminates the need to remove the protective groups under harsh conditions, which would be incompatible with protein ligands maintaining their native structures.

We next established a TRICEPS-based workflow to identify receptors for ligands using two samples (**Fig. 2**). In one sample, the TRICEPS-labeled ligand predominantly directed the carbohydrate-reactive element of the reagents toward ligand-specific glycoprotein receptors. In parallel, an equimolar amount of TRICEPS was either quenched with glycine or coupled to a control ligand with known or no binding preferences. This control allowed us to account for nonspecific binding based on the relative abundance of a given target protein on a cell line or tissue. The two TRICEPS-containing samples were then incubated with previously oxidized cells or tissues under near-physiological conditions. During this phase, transient and stable ligand-receptor interactions were expected to lead to increased covalent capture events between TRICEPS and nearby carbohydrates. In general, the reactivity between hydrazine and carbohydrates is advantageous because it targets TRICEPS away from polypeptide domains that are important for ligand binding. After the receptor-capture reaction, the cells were lysed and enzymatically digested with trypsin, and TRICEPS-labeled peptides were isolated using streptavidin beads. This circumvents problems inherent in the affinity purification of intact plasma membrane proteins, such as limited solubility and unspecific interactions through exposed hydrophobic domains. Captured N-glycopeptides were specifically released from the beads through an enzymatic cleavage with PNGase F. This endoglycosidase cleaves between the proximal *N*-acetylglucosamine and the asparagine of the glycopeptide in the N-X-S/T glycosylation site motif (where N stands for asparagine, X stands for any amino acid except proline and S/T for serine or threonine, respectively). Besides releasing the glycosylated peptides, PNGase F treatment also results in the enzymatic deamidation of the glycosylated asparagine to aspartic acid. This generates a +0.984 Da mass shift at the asparagine residue and thus introduces the specific N[115]-X-S/T signature (where 115 is the mass of the deamidated Asn residue) in formerly glycosylated cell surface peptides. This protocol leads to the generation of two virtually identical peptide samples that can be analyzed with a high mass accuracy mass spectrometer. The identified peptides were filtered for the presence of N[115]-X-S/T motifs, and the relative concentrations of cell surface glycopeptides in the ligand sample were compared to those in the control sample using MS1-based label-free quantification. Identified peptides of random cell surface proteins were expected to have equal concentrations in both samples, whereas peptides of the corresponding receptors were specifically enriched in the ligand sample.

In the first application of this approach, we coupled human insulin to TRICEPS and then added it to murine adipocytes at a concentration routinely used in cell culture applications (1.7 μM). In the parallel control reaction, an equimolar amount of TRICEPS was quenched with glycine and added to an equal number of cells. Mass spectrometric analysis led to the identification of 117 formerly glycosylated peptides with similar abundance in both samples. These were derived from 69 random, but highly abundant, cell surface glycoproteins and represented nonspecific capture events. Owing to the comparatively low abundance of the insulin receptor (INSR) on the target cells, INSR peptide signals were weak or not detectable above background in the control sample. In contrast, INSR peptides were highly enriched in the sample that had been generated with insulin providing selectivity for the capture reaction (**Fig. 3a** and **Supplementary Table 1**). These results show that TRICEPS-bound ligands can be used to guide the hydrazine reaction toward specific receptors on living cells. Furthermore, this experiment led to the immediate and unbiased detection of the interaction of human insulin with the murine INSR on the target cells. Notably, the INSR was identified with six unique enriched peptides. This added confidence to the specific receptor identification because every peptide provided independent enrichment information owing to the peptide-level affinity purification. In a similar experiment using human Jurkat T lymphocytes, the enrichment of the human INSR was prevented by pre-incubating cells with a tenfold molar excess of uncoupled insulin (**Supplementary Fig. 2** and **Supplementary Table 2**), again highlighting the robustness and specificity of the approach.

To demonstrate the versatility of the LRC technology, we used transferrin (a secreted glycoprotein) and apelin-17 (a G protein–coupled receptor (GPCR) binding peptide) as ligands to capture receptors on human U-2 OS osteosarcoma cells. This experiment simultaneously identified the transferrin receptor protein 1 and the apelin receptor as interaction partners of the respective ligands (**Fig. 3b** and **Supplementary Table 3**). The identification of the transferrin receptor shows that LRC technology is also compatible with glycosylated

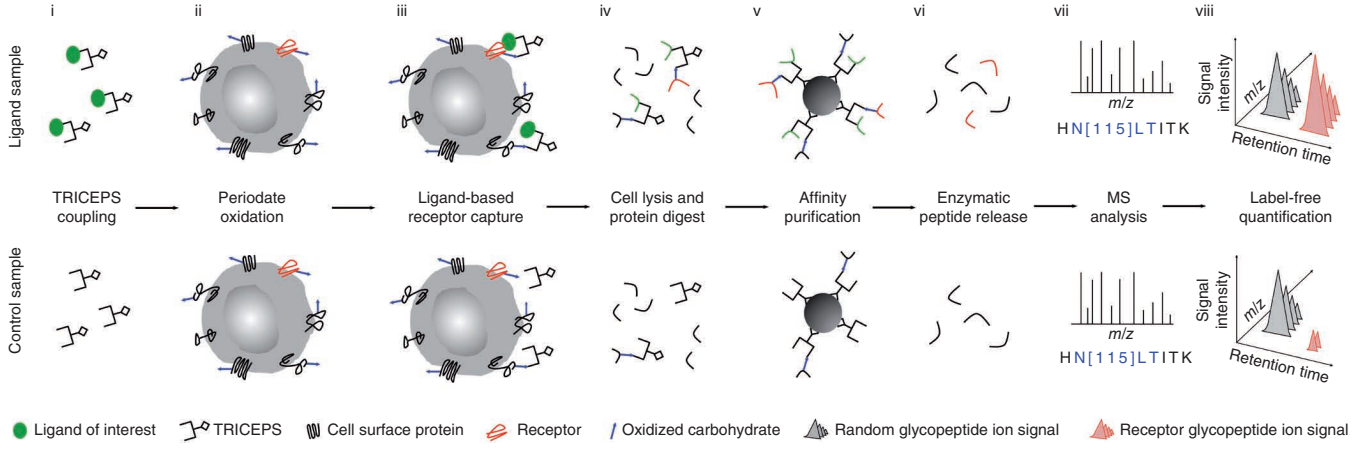

Figure 2 Workflow for the LRC and identification on living cells. (i) Conjugation of a purified ligand of interest to TRICEPS and quenching of an equimolar amount of TRICEPS with glycine (or coupling to a control ligand with known or no binding preferences) in the control reaction. (ii) Mild periodate oxidation of target cells or tissue to introduce aldehydes into carbohydrate structures for TRICEPS capture. (iii) LRC and stochastic biotinylation of random cell surface glycoproteins according to their abundance. (iv) Cell lysis and tryptic digest of proteins. (v) Biotin-mediated affinity enrichment of captured glycopeptides on streptavidin beads. (vi) Cell surface N-glycopeptide release by PNGase F treatment, which introduces the N[115]-X-S/T motif (N[115], deamidated asparagine; X, any amino acid except proline; S/T, serine or threonine, respectively) in formerly N-glycosylated peptides. (vii) Glycopeptide identification by high mass accuracy MS and peptide filtering for the presence of the N[115]-X-S/T motif. (viii) Relative label-free quantification of formerly glycosylated cell surface peptides to identify specific LRC events.

ligands such as transferrin. Furthermore, this experiment demonstrates the potential of LRC technology to identify interactions of peptide ligands with GPCRs. Such interactions are particularly difficult to capture by other means owing to the small size of peptide ligands and the typically low abundance and integral membrane nature of the corresponding seven-transmembrane domain receptors.

Next, we explored the utility of LRC technology to identify pharmaceutically interesting targets of growth factors, therapeutic antibodies and engineered affinity binders. The first experiment was performed with the therapeutic antibody trastuzumab (Herceptin), which binds to ErbB2, and the epidermal growth factor (EGF) on U251 human glioblastoma cells. To allow for statistical evaluation, we performed these experiments in biological triplicates. This application was of particular interest as the two ligands are very different in nature (trastuzumab is a 150-kDa glycoprotein, whereas EGF is a small 6-kDa protein), and they each bind to a single site on different members of the EGF receptor (EGFR) superfamily[16,17]. Using trastuzumab as a ligand, ErbB2 was unambiguously identified through the consistent enrichment of corresponding ErbB2 glycopeptides (**Fig. 3c** and **Supplementary Table 4**). As another positive control, the EGF experiment yielded enriched glycopeptides from domain III of the EGFR. Because the molecular size of the EGFR by far exceeds that of its ligand and the enriched peptides matched only two of the more than ten EGFR glycosylation sites, this result also suggests that the approximate localization of the EGF binding site is on this particular domain of the receptor.

We further explored the capability of LRC technology to map binding sites in a subsequent experiment with two 18 kDa designed ankyrin repeat proteins (DARPins), DARPin 9.01 and DARPin H14. These two DARPins were selected by phage display as antibody mimetic proteins targeting the full-length ectodomain of the ErbB2 protein[18]. We applied LRC to BT-474 human breast cancer cells and confirmed binding of both DARPins to the ErbB2 protein in its native environment on living cells. Furthermore, the specific enrichment of corresponding glycosylation sites confirmed that DARPin 9.01 binds to domain I in ErbB2, whereas DARPin H14 binds to domain IV (**Fig. 3d** and **Supplementary Table 5**). These epitopes were independently determined *in vitro* using enzyme-linked immunosorbent assays (ELISA) with different recombinant extracellular domains of ErbB2 (**Supplementary Fig. 3**). These results demonstrate the potential of the LRC technology to identify targets and binding sites at domain-level resolution for established and prospective clinical binding proteins. Small molecule–based LRC applications are also technically feasible through direct derivatization of ligands or the prior attachment of chemical functionalities to TRICEPS. Because such strategies typically lead to a uniform modification of coupled ligands at a particular site, ligands should be tested for functionality after derivatization. In addition, the number of receptor glycosylation sites that can be reached on average is likely to be reduced compared to LRC applications with protein ligands.

In contrast to homogenous cell populations, such as immortalized clonal cell lines, complex primary tissues contain different cell types, each of which is expected to express a quantitatively and qualitatively different set of cell surface proteins. We thus did an LRC experiment with trastuzumab on primary breast cancer tissue. This tissue had been classified as immunohistochemically ErbB2 negative with a diagnostic antibody and was therefore expected to contain only low levels of the target protein (**Supplementary Fig. 4**). Nevertheless, a single LRC experiment unambiguously confirmed the ErbB2 protein as the primary target and several high- and low-affinity immunoglobulin Fcγ receptors as additional interaction partners of this IgG antibody in the tissue (**Fig. 3e** and **Supplementary Table 6**). These results underscore the sensitivity of the approach and indicate that trastuzumab binds to multiple proteins on subpopulations of cells within complex target tissues, in contrast to the comparatively simple scenario on U251 cells. Thus, the LRC technology has the potential to detect multiple intended, as well as off-target, interactions of almost any specific binding protein on clinically relevant primary tissues inaccessible to genetic engineering.

This particular capability of the LRC technology also holds great potential for the identification of receptor panels for more complex ligands. As biological particles, viruses often depend on multiple host cell surface proteins to mediate binding and subsequent internalization[19], thus representing a unique class of complex ligands.

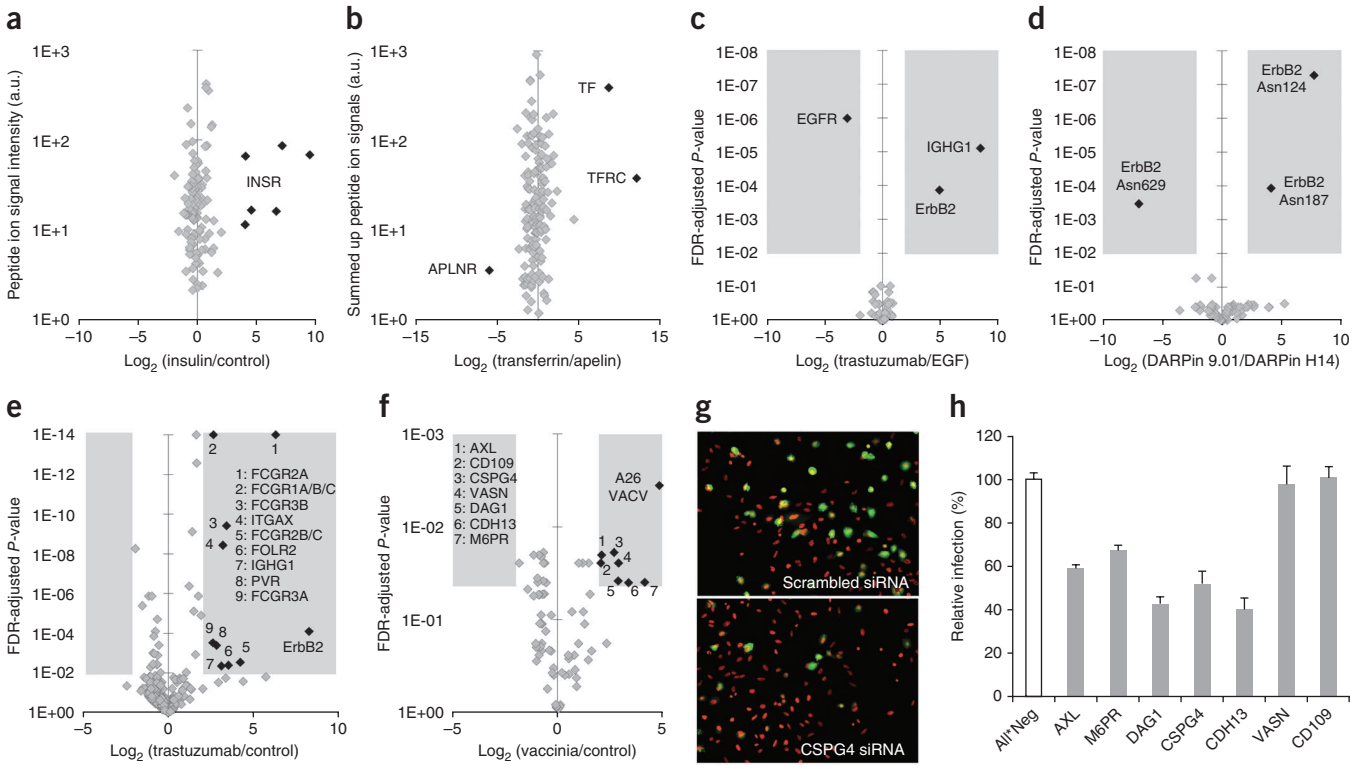

Figure 3 LRC identifies receptors and receptor panels for ligands ranging from peptides to intact viruses on living cells and tissues. (**a**) LRC with human insulin on differentiated murine visceral adipocytes. Data are shown on the peptide level. (**b**) LRC with transferrin and apelin on U-2 OS cells. Data are shown on the protein level. Two glycopeptides of transferrin were captured and identified as well without prior exposure of the ligand to oxidative conditions, which was also observed with other glycoprotein ligands used later on. APLNR, apelin receptor; TF, transferrin; TFRC, transferrin receptor. (**c**) LRC with epidermal growth factor and trastuzumab on U251 cells in biological triplicates. The receptor candidate space highlighted in gray is defined by an enrichment factor of fourfold or greater and an FDR-adjusted *P*-value less than or equal to 0.01. IGHG1, immunoglobulin heavy constant gamma 1. (**d**) LRC with DARPins on BT-474 cells in biological triplicates. Data are shown on the glycosylation site level. Identified ErbB2 glycosylation sites: Asn124, Asn187 (ErbB2 domain I), Asn629 (ErbB2 domain IV). (**e**) LRC with trastuzumab on primary breast cancer tissue in technical (multiple LC-MS/MS runs of the same sample) triplicates. (**f**) LRC with intact vaccinia viruses on HeLa CCL2 cells in biological triplicates. Receptor candidates with a fold enrichment ≥4 and a *P* value ≤ 0.05 were tested for effects on viral infectivity in follow-up investigations. Identified peptides and proteins are shown in gray. Ligands, receptors and receptor candidates are shown in black. (**g**) Representative images of siRNA-treated HeLa cells infected with vaccinia virus (red, nuclei; green, vaccinia virus early protein expression). (**h**) Percentage of vaccinia virus–infected HeLa cells after siRNA knockdown of the receptor candidates identified in the LRC experiment with intact viruses. Infection rates were calculated relative to control cells transfected with AllStars Negative Control siRNAs (All*Neg). Error bars indicate s.d.

As such, virus receptor identification remains a daunting task, with the vast majority of specific virus receptors remaining undefined. For instance, vaccinia virus (VACV) was used for the eradication of smallpox and is the most intensively studied poxvirus, yet little is known about the cell surface receptors used by VACV for infection. Competition assays suggest that VACV mature virions attach to host cells through interactions between viral proteins (A27 and D8) and glycosaminoglycans (heparin and chondroitin sulfates)[20,21], or A26 and the extracellular matrix protein laminin[22]. For internalization, phosphatidylserine in the viral membrane[23], EGFR[24], the lipid raft component CD98 (ref. 25) and the phosphatidylserine receptor AXL[26] have all been implicated. However, no direct *in vivo* interaction between mature virions and these proposed binding and internalization factors has been demonstrated. Thus, to better understand the complex interactions between these viruses and their host cells, we used LRC technology to identify the cell surface factors to which VACV mature virions bind.

Preliminary experiments demonstrated that only very high TRICEPS coupling ratios affected mature virion infectivity (**Supplementary Fig. 5**). Subsequently, LRC with intact mature virions on HeLa CCL2 cells revealed a defined set of seven cellular VACV binding factor candidates: AXL, M6PR, DAG1, CSPG4, CDH13, CD109 and VASN (**Fig. 3f** and **Supplementary Table 7**). Identification of chondroitin sulfate proteoglycan 4 (CSPG4), laminin binding protein dystroglycan 1 (DAG1) and AXL was consistent with previous independent findings[21,22,26]. The candidate VACV receptors were then subjected to short interfering RNA (siRNA)-mediated silencing to determine their functional relevance with regard to virus infection (**Supplementary Fig. 6** and **Supplementary Table 8**). Of the seven candidates tested, five (AXL, M6PR, DAG1, CSPG4 and CDH13) reduced VACV infection by 40–60% (**Fig. 3g**,**h**). That silencing of a single factor did not completely attenuate infection was not surprising. This likely reflects the complex nature of VACV-host cell interactions in which multiple virus and cellular factors each contribute to successful and potentially cooperative binding and infection. Collectively, these five factors are likely to represent the earliest set of cell surface proteins that the virus binds to and uses to infect a cell. Although LRC technology provided a unique opportunity to identify this defined set of host receptor candidates, the enrichments of individual receptors were generally weaker than in the comparatively simple one-to-one or one-to-many types of interactions addressed previously. This was expected owing to the comparatively large size of this promiscuous ligand and the many-to-many

types of interactions involved between multiple viral and host factors. Nevertheless, this discovery-driven application with a complex ligand demonstrated the potential of LRC technology to define a set of receptor candidates, after which the functional relevance can be investigated in targeted and combinatorial follow-up experiments. In summary, we are confident that LRC technology can be used to map the mostly uncharted territory of cell surface interactions.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession code. The data associated with this manuscript may be downloaded from https://proteomecommons.org/ Tranche using the following hash:4rCLZMX5zYWApQjYpivFhKVrxCjTJji6V/CCyxjZYqqs1d6AMwIAR0bq1utGVuQEoj9+B8C64rQbHBNjCHWhkT464A4AAAAAAAACSw = =.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We greatly acknowledge T. Clough and O. Vitek at Purdue University for help with statistical data analysis. We are grateful to A. Hofmann, T. Bock, D. Bausch-Fluck, F. Cerciello, A. Jacobs and A. Leitner for suggestions and support at all stages of the project. We acknowledge S. Dettwiler, P. Schraml, M. Tinguely, H. Moch and the Laboratory for *In situ* Technologies, University Hospital Zurich, for preparation and staining of breast cancer tissues. This work was supported by funding from National Center of Competence in Research (NCCR) Neural Plasticity and Repair (to B.W.), Swiss National Science Foundation (SNSF) (to B.W.), SystemsX.ch/InfectX (to B.W.), SNSF Ambizione (to J.M.), SystemsX.ch and European Research Council (ERC) (to R.A.) and SystemsX.ch/InfectX and ERC (to S.K. on behalf of A. Helenius). Immortalized murine pre-adipocytes were kindly provided by M. Rosenwald and C. Wolfrum (ETH Zurich). ErbB2-negative breast carcinoma tissue cut into 50 µm slices was kindly provided by the tissue biobank of the Institute of Surgical Pathology, University Hospital Zurich.

AUTHOR CONTRIBUTIONS

A.P.F. and B.W. designed the project and wrote the paper. A.P.F. performed experiments and analyzed all data. A.P.F., B.W., O.-Y.J. and E.M.C. designed TRICEPS and O.-Y.J. synthesized the reagents. J.M. and S.K. designed and performed vaccinia virus experiments and J.M. edited the manuscript. C.J. and A.P. designed DARPin experiments and performed ELISAs. R.A., H.M. and L.M.H. contributed ideas and performed experiments. All authors discussed the results and implications and commented on the manuscript at all stages.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/nbt.2354.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Hubner, N.C. *et al.* Quantitative proteomics combined with BAC TransgeneOmics reveals *in vivo* protein interactions. *J. Cell Biol.* **189**, 739–754 (2010).
2. Glatter, T., Wepf, A., Aebersold, R. & Gstaiger, M. An integrated workflow for charting the human interaction proteome: insights into the PP2A system. *Mol. Syst. Biol.* **5**, 237 (2009).
3. Bantscheff, M. & Drewes, G. Chemoproteomic approaches to drug target identification and drug profiling. *Bioorg. Med. Chem.* **20**, 1973–1978 (2012).
4. Lenz, T., Fischer, J.J. & Dreger, M. Probing small molecule-protein interactions: A new perspective for functional proteomics. *J. Proteomics* **75**, 100–115 (2011).
5. Barglow, K.T. & Cravatt, B.F. Activity-based protein profiling for the functional annotation of enzymes. *Nat. Methods* **4**, 822–827 (2007).
6. Elschenbroich, S., Kim, Y., Medin, J.A. & Kislinger, T. Isolation of cell surface proteins for mass spectrometry-based proteomics. *Expert Rev. Proteomics* **7**, 141–154 (2010).
7. Helbig, A.O., Heck, A.J.R. & Slijper, M. Exploring the membrane proteome–challenges and analytical strategies. *J. Proteomics* **73**, 868–878 (2010).
8. Savas, J.N., Stein, B.D., Wu, C.C. & Yates, J.R. Mass spectrometry accelerates membrane protein analysis. *Trends Biochem. Sci.* **36**, 388–396 (2011).
9. Lee, A. How lipids affect the activities of integral membrane proteins. *Biochim. Biophys. Acta.* **1666**, 62–87 (2004).
10. Zeng, Y., Ramya, T.N.C., Dirksen, A., Dawson, P.E. & Paulson, J.C. High-efficiency labeling of sialylated glycoproteins on living cells. *Nat. Methods* **6**, 207–209 (2009).
11. Wollscheid, B. *et al.* Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat. Biotechnol.* **27**, 378–386 (2009).
12. Hofmann, A. *et al.* Proteomic cell surface phenotyping of differentiating acute myeloid leukemia cells. *Blood* **116**, e26–e34 (2010).
13. Frei, A., Jeon, O.Y., Carreira, E. & Wollscheid, B. Trifunctional crosslinking reagents. European Patent Application No. 11000731 (2012).
14. Mädler, S., Bich, C., Touboul, D. & Zenobi, R. Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J. Mass Spectrom.* **44**, 694–706 (2009).
15. Sletten, E.M. & Bertozzi, C.R. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Edn Engl.* **48**, 6974–6998 (2009).
16. Ferguson, K.M. Structure-based view of epidermal growth factor receptor regulation. *Annu. Rev. Biophys.* **37**, 353–373 (2008).
17. Cho, H.-S. *et al.* Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* **421**, 756–760 (2003).
18. Steiner, D., Forrer, P. & Plückthun, A. Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display. *J. Mol. Biol.* **382**, 1211–1227 (2008).
19. Marsh, M. & Helenius, A. Virus entry: open sesame. *Cell* **124**, 729–740 (2006).
20. Chung, C., Hsiao, J., Chang, Y. & Chang, W. A27L protein mediates vaccinia virus interaction with cell surface heparan sulfate. *J. Virol.* **72**, 1577–1585 (1998).
21. Hsiao, J.C., Chung, C.S. & Chang, W. Vaccinia virus envelope D8L protein binds to cell surface chondroitin sulfate and mediates the adsorption of intracellular mature virions to cells. *J. Virol.* **73**, 8750–8761 (1999).
22. Chiu, W.-L., Lin, C.-L., Yang, M.-H., Tzou, D.-L.M. & Chang, W. Vaccinia virus 4c (A26L) protein on intracellular mature virus binds to the extracellular cellular matrix laminin. *J. Virol.* **81**, 2149–2157 (2007).
23. Mercer, J. & Helenius, A. Vaccinia virus uses macropinocytosis and apoptotic mimicry to enter host cells. *Science* **320**, 531–535 (2008).
24. Mercer, J. *et al.* Vaccinia virus strains use distinct forms of macropinocytosis for host-cell entry. *Proc. Natl. Acad. Sci. USA* **107**, 9346–9351 (2010).
25. Schroeder, N., Chung, C.-S., Chen, C.-H., Liao, C.-L. & Chang, W. The lipid raft-associated protein CD98 is required for vaccinia virus endocytosis. *J. Virol.* **86**, 4868–4882 (2012).
26. Morizono, K. *et al.* The soluble serum protein Gas6 bridges virion envelope phosphatidylserine to the TAM receptor tyrosine kinase Axl to mediate viral entry. *Cell Host Microbe* **9**, 286–298 (2011).

ONLINE METHODS

Mammalian cell cultures. All cells were grown at 37 °C and 5% ambient CO_2. Adherent U-2 OS osteosarcoma cells, U251 glioblastoma cells and HeLa CCL2 adenocarcinoma cells were grown to near 100% confluence from initial addition to 140 × 20 mm Nunclon dishes (Thermo Scientific) in Dulbecco's Modified Eagle's Medium (DMEM, Sigma-Aldrich) completed with 10% Fetal Bovine Serum (Sigma-Aldrich) and penicillin-streptomycin-glutamine (Invitrogen). BT-474 human breast cancer cells were grown accordingly in complete RPMI medium (Sigma-Aldrich). Jurkat T suspension cells were grown in 175 cm^2 Nunclon flasks (Thermo Scientific) to an approximate concentration of 400,000 cells/ml in complete RPMI medium (Sigma-Aldrich). Immortalized murine pre-adipocytes were kindly provided by Matthias Rosenwald and Christian Wolfrum (ETH Zurich) and originally obtained by Johannes Klein (University of Lübeck). Murine cells were grown in DMEM 4.5 g/l glucose (Invitrogen) completed with 20% fetal bovine serum and penicillin-streptomycin-glutamine. Pre-adipocytes were differentiated into white adipocytes by the addition of a differentiation cocktail (115 µg/ml 3-isobutyl-1-methylxanthine (Sigma-Aldrich), 1 µM dexamethasone (Sigma-Aldrich), 1 µg/ml insulin (Sigma-Aldrich) in medium) on day 0, removal of medium and addition of new medium containing the same cocktail on day 1, removal of medium and addition of new medium containing insulin on day 2, and removal of medium and addition of standard medium on day 4. Cells were harvested on day 7 after they had adopted a rounded phenotype and accumulated lipids in the form of intracellular lipid droplets.

Flow cytometric analyses. *Aldehyde capture with TRICEPS.* Reagents were incubated with a tenfold molar excess of glycine in 25 mM HEPES pH 8.2 for 60 min to quench the amine reactive moieties of the reagents. Jurkat T cells were collected and resuspended in labeling buffer (PBS, pH adjusted to 6.5 with H_3PO_4) and cooled down to 4 °C for all of the following steps. Half of the cells were subjected to mild periodate oxidation with 1.5 mM $NaIO_4$ (Pierce) for 15 min in the dark and cell pellets were washed once with labeling buffer. Subsequently, cells were reacted with 100 µM TRICEPS or biocytin hydrazide (Biotium) in labeling buffer for 60 min on a slow rotator. Cells were washed twice and stained with streptavidin-FITC (BD Biosciences, 554060) in FACS buffer (1% Fetal Bovine Serum in PBS) for 20 min. Labeled cells were washed once and analyzed with a FACSCalibur flow cytometer (BD Biosciences) in combination with Flowjo (version 8.8.4) software.

Transferrin uptake. 50 µg AlexaFluor-488 labeled transferrin (Invitrogen) was coupled to TRICEPS or incubated in coupling buffer (25 mM HEPES pH 8.2) without reagent. After coupling, the labeled protein was diluted in DMEM to a final concentration of 8.3 µg/ml. Confluent 12 wells of HeLa ATCC cells were starved in DMEM for 2 h. Transferrin was bound to cells for 30 min at 4 °C (750 µl/well) and cells were shifted to 37 °C for 15 min (water-bath). Cells were washed once with cold PBS and treated with acid washing buffer (0.1 M NaCl, 0.1 M glycine, pH 3.0) for 2 min. Samples were detached with 0.05% trypsin, fixed (4% formaldehyde in PBS, 20 min) and analyzed by flow cytometry.

VACV infection. Confluent 24 wells of HeLa ATCC cells were infected with VACV strain Western Reserve expressing EGFP from an early viral promoter (WR Early-EGFP) at different multiplicities of infection (MOI) in 250 µl DMEM. After 30 min at 37 °C, DMEM was replaced with full medium (DMEM, 10% FBS, NEAA and Glutamax). Five hours after infection, cells were detached with 0.05% trypsin, fixed (4% formaldehyde in PBS, 20 min) and analyzed by flow cytometry.

LRC. 100 µg insulin (Sigma-Aldrich), 100 µg holo-transferrin (Sigma-Aldrich), 30 µg apelin-17 (Tocris Bioscience), 50 µg EGF (Sigma-Aldrich), 100 µg trastuzumab (Roche), 5×10^8 plaque-forming units of vaccinia virus or 100 µg glycine (Sigma-Aldrich) were reacted with 40 µg TRICEPS in 40–100 µl 25 mM HEPES pH 8.2 for 60 min. Typically, coupling reactions do not need to be optimized and we recommend using 40 µg TRICEPS per 100 µg ligand for all applications. Higher coupling ratios can be applied in cases where a limited amount of ligand is available. Hydrolyzed TRICEPS and the fraction of ligands that are compromised by the coupling to reagents will simply contribute to the random capture of cell surface proteins, so there is no need to purify coupled ligands. Using intact viruses as ligands in LRC experiments, true receptors will be captured and enriched by TRICEPS coupled to viral proteins that interact with these host factors. At the same time, reagents that are coupled to other viral proteins may contribute to the stochastic capture of cell surface proteins (the majority of capture events in every LRC experiment) in the vicinity of actual binding events. On HeLa CCL2 cells no binding of insulin to the INSR was detected in preliminary LRC experiments and the peptide hormone was used to quench reagents. 5×10^7 cells (murine adipocytes, U-2 OS, U251, BT-474, HeLa CCL2) per experiment were harvested by gentle scraping in PBS and cooled down to 4 °C for all of the following steps. Frozen ErbB2-negative breast carcinoma tissue cut into 50-µm slices was kindly provided by the tissue biobank of the Institute of Surgical Pathology, University Hospital Zurich (approved by the local ethics committee: StV12-2005). The procedure for freezing and cutting tissues has been previously described[27]. Tissue slices were processed without further dissociation to partly conserve the three-dimensional tissue structure during incubation with trastuzumab. Cells or tissue slices were resuspended in 50 ml labeling buffer (PBS pH 6.5) and subjected to mild periodate oxidation with 1.5 mM $NaIO_4$ (Pierce) on a slow rotator in the dark for 15 min (30 min for tissue slices). Subsequently, cells were washed once with 50 ml labeling buffer and resuspended in 10 ml labeling buffer for incubation with the ligands or quenched reagents. LRC was carried out for 60 min (90 min for tissue slices) on a slow rotator. Cells were collected by centrifugation and the cell pellet was resuspended in 1 ml 50 mM ammonium bicarbonate (Sigma-Aldrich). Cells were lysed by indirect sonication (100% amplitude, 0.8 cycle, 2 × 30 s) in a VialTweeter (Hielscher) and the subsequent addition of RapiGest surfactant (Waters). Tissue slices were subjected to additional sonication cycles and vigorous pipetting after the addition of RapiGest to completely dissociate the tissue. After centrifugation (2,500*g*, 10 min), the supernatant was collected after which proteins were reduced by adding 5 mM TCEP (Pierce) at 20 °C for 30 min and then alkylated by adding 10 mM iodoacetamide (Fluka) in the dark at 20 °C for 30 min. Trypsin (Sigma) was added to a 1:50 trypsin to protein ratio and protein digestion was carried out overnight at 37 °C. After digestion, peptide mixtures were heated to 96 °C for 10 min to inactivate proteases and undigested particles were removed by centrifugation (13,000*g* for 10 min). TRICEPS-captured cell surface glycopeptides were bound to streptavidin by addition of 60 µl washed Streptavidin Plus UltraLink Resin (Pierce) and incubated for 60 min on a slow rotator at 4 °C. Beads were washed extensively in Mobicols (Boca Scientific) connected to a Vac-Man Laboratory Vacuum Manifold (Promega) with a variety of buffers: 5 M NaCl, followed by 100 mM NaCl, 100 mM glycerol, 50 mM Tris, 1% Triton X-100, followed by 100 mM $NaHCO_3$ pH 11, followed by 50 mM ammonium bicarbonate. Washed beads were incubated with 300 µl 50 mM ammonium bicarbonate containing 1.5 µl PNGase F (New England Biolabs) in a head-over-head shaker overnight at 37 °C. Peptides were eluted by spinning of the Mobicol and subsequent addition of another 500 µl 50 mM ammonium bicarbonate. Combined eluates were acidified with 40 µl 10% formic acid and subjected to C18 purification using 3–30 µg UltraMicroSpin Columns (The Nest Group) according to the manufacturer's instructions.

Liquid chromatography–tandem mass spectrometry (LC-MS/MS) analysis. Peptide samples were separated by reversed-phase chromatography on a high-performance liquid chromatography (HPLC) column (75-µm inner diameter, New Objective) that was packed in-house with a 10-cm stationary phase (Magic C18AQ, 200 Å, 3 µm, Michrom Bioresources) and connected to a nano-flow HPLC (nanoLC-Ultra 1D plus, Eksigent) and an autosampler (nanoLC AS-2, Eksigent). The HPLC was coupled to a LTQ-Orbitrap XL mass spectrometer (Thermo Scientific) equipped with a nanoelectrospray ion source (Thermo Scientific). Peptides were loaded onto the column with 95% buffer A (98% H_2O, 2% acetonitrile, 0.1% formic acid) and eluted with 300 nl/min over a 40-min linear gradient from 5–35% buffer B (2% H_2O, 98% acetonitrile, 0.1% formic acid). After the gradient, the column was washed with 95% buffer B and re-equilibrated with buffer A. Mass spectra were acquired in a data-dependent manner, with an automatic switch between MS and MS/MS scans. High-resolution MS scans were acquired in the Orbitrap (60,000 FWHM, target value 10^6) to monitor peptide ions in the mass range of 350–1,650 m/z, followed by collision-induced dissociation MS/MS scans in the ion trap (minimum signal threshold 150, target value 10^4, isolation width 2 m/z) of the five most

doi:10.1038/nbt.2354

intense precursor ions. To avoid multiple scans of dominant ions, the precursor ion masses of scanned ions were dynamically excluded from MS/MS analysis for 10 s. Singly charged ions and ions with unassigned charge states were excluded from MS/MS fragmentation.

MS data analysis and quantification. Raw data were converted to the open mzXML format with ReAdW (version 4.3.1). mzXML files were searched by the SEQUEST search engine against UniProtKB/Swiss-Prot protein databases (version 57.15 of *Homo sapiens*, *Mus musculus* and vaccinia virus strain Western Reserve) concatenated with the sequences of common contaminants. Search parameters for the peptide identification included a precursor mass tolerance of 0.05 Da, a minimum of one tryptic terminus and a maximum of two internal trypsin cleavage sites. Cysteine carbamidomethylation (+57.021 Da) was set as a static amino acid modification and methionine oxidation (+15.995 Da) as well as asparagine deamidation (+0.9840 Da) were set as differential modifications. The PeptideProphet and ProteinProphet tools of the Trans-Proteomic Pipeline (TPP version 4.5) were used for the probability scoring of peptides and proteins, and protein identifications were filtered to a false-discovery rate of ≤1%. For label-free quantification of the identified peptides, Thermo. RAW files were imported into the Progenesis LC-MS software (Nonlinear Dynamics) to detect and extract ion signals corresponding to peptide features. Peptide features matched to ion signals were filtered for the presence of the N[115]-X-S/T motif (wherein N[115] stands for a deamidated asparagine residue, X stands for any amino acid except proline and S/T for serine or threonine, respectively). This filtering step assured that only formerly N-glycosylated cell surface peptides and proteins that had been specifically captured and purified were considered for the quantitative sample comparison (typically 40–70% of the identified proteins). Peptide ion signals with high fold changes between samples were manually inspected to assure correct feature picking and extraction. The different enrichment factors obtained for potential receptor peptides can typically be explained by differences in the efficiency of the LRC reaction; for example, due to the spatial arrangement of the captured carbohydrates in relation to the ligand binding site. Furthermore, the virtual absence of some receptor peptide signals in the control sample can make ratios susceptible to the response of individual peptides in the mass spectrometer and background integration in the quantification process. Peptide signal intensities were exported on the feature level for the statistical analysis.

Statistical analysis. As described above, the output of a LRC experiment is a list of quantified spectral features. For the statistical analysis, the intensities for peptide features with multiple charge states and differential modifications were summed up and constant normalization was performed on the log-transformed intensities across samples. To make protein-level conclusions based on those measurements, a previously described fixed effects Analysis of Variance (ANOVA) model[28] was applied using the publicly available R-based software MSStats (http://www.stat.purdue.edu/~ovitek/Software.html) according to instructions in the package documentation. This model views individual features as replicate measurements of a protein's abundance and explicitly accounts for this redundancy, which increases the sensitivity and specificity of testing. Thus, it was used to test each protein for differential abundance in all pairwise comparisons of ligand and control samples and to report the *P*-values of the tests. Subsequently, *P*-values were adjusted for multiple comparisons using the approach by Benjamini and Hochberg to control the experiment-wide false-discovery rate at a desired level.

ErbB2-binding DARPin selection and ELISA. Selection of ErbB2-binding DARPins has been described previously[18]. In brief, Signal Recognition Particle Phage Display was used to select DARPins recognizing the complete ectodomain (domains I-IV) of ErbB2. To determine the binding specificity at the domain level in ELISAs, recombinant ErbB2-ectodomains carrying an N-terminal melittin signal sequence (MKFLVNVALVFMVVYISYIYA) and an N-terminal His_6 tag were expressed in *Spodoptera frugiperda* (*Sf9*) cells using baculoviral vectors. *Sf9* cells were grown to a density of 4×10^6 cells/mL and co-infected with the respective virus at a MOI of 1. 72 h post-infection, cells were harvested by centrifugation (30 min, 5,000*g*, 4 °C) and the cleared medium was subjected to immobilized metal ion affinity chromatography (IMAC) purification with Ni-NTA Superflow (Qiagen) purification resin. The purified ErbB2 domains (200 nM, 100 µl/well) in PBS were immobilized on MaxiSorp plates (Thermo Scientific) by overnight incubation at 4 °C. For ELISAs, wells were blocked with 300 µl of PBSTB (PBS, 0.1% Tween-20, 0.2% BSA) for 1 h at room temperature. 50 nM of purified DARPins were incubated with the target domains for 1 h at room temperature followed by three washing steps with 300 µl of PBSTB. For detection of bound DARPins, an anti RGS-His IgG1 mouse antibody (Qiagen) was added (1:5,000 in PBSTB, 1 h at room temperature), which recognizes the N-terminal MRGS-His_6 tag of the DARPins, and wells were washed as described above. After incubation with a secondary anti-mouse-IgG antibody alkaline phosphatase conjugate (Sigma-Aldrich) (1:10,000 in PBSTB, 1 h at room temperature), pNPP substrate (Fluka) was added to measure alkaline phosphatase activity.

siRNA knockdown and infection assays. HeLa cells were reverse-transfected with pooled Qiagen siRNAs (3 siRNAs/gene, **Supplementary Table 8**) at a final concentration of 20 nM in 96-well flat-bottom microscopy plates (Greiner Bio-One). siRNAs in ddH_2O were mixed with Lipofectamine RNAiMax (Invitrogen) in DMEM to a final volume of 30 µl according to the manufacturer's instructions (0.1 µl RNAiMax/well). 1,800 cells were seeded onto the Lipofectamine mixture and transfected for 72 h. Transfected cells were infected with VACV Western Reserve Early-EGFP with a target infection index of ~20% to allow for a dynamic measurement of possible enhancing and inhibiting effects on infection. Virus was bound to cells at 37 °C in 50 µl DMEM for 30 min, inoculum was replaced with full medium and cells were infected for an additional 6 h. Samples were fixed (4% formaldehyde, 20 min) and permeabilized (0.5% Triton X-100), and nuclei were stained with Hoechst 33258 (Molecular Probes). EGFP was stained with rabbit antiserum and an Alexa-Fluor-488 goat anti-rabbit secondary antibody (Invitrogen) to increase the signal of inherently weak early viral gene expression levels. Primary antibody staining was done in permeabilization buffer, secondary staining in PBS. Automated image acquisition (16 images/well) was done with a 10× objective using an ImageXpress Micro high content screening system (Molecular Devices). Cell numbers and raw infection indices for each well were determined using a MATLAB-based infection-scoring program (InfectionCounter V, Thomas Heger, ETH Zurich). Infection rates were calculated relative to the control wells transfected with AllStars Negative (All*Neg) Control siRNAs (Qiagen). An internal checkerboard assay was performed for each plate to normalize for the negative correlation between cell number and infection index. Increasing cell numbers (500–1,800 cells, 18 data points/plate) were transfected (All*Neg) and infected as described. The correlation between the number of nuclei and infection index was determined using logarithmic regression (% infection = a*ln(nuclei)+b). Relative infection rates of target siRNAs were calculated in relation to the negative control at the corresponding nuclei number. For the verification of siRNA-mediated protein depletion, HeLa ATCC cells were transfected with 20 nM nontargeting All*Neg siRNAs or pooled siRNAs targeting receptor candidate proteins for 72 h. Cells were detached (500 µM EDTA in PBS), lysed in RIPA buffer (50 mM Tris pH7.4, 150 mM NaCl, 0.5% Na-deoxycholate, 1.0% Triton X-100) and protein levels were analyzed by western blot using primary AXL (Cell Signaling, 4566), CDH13 (R&D Systems, AF3264), M6PR (Novus Biologicals, NB100-93413) and α-tubulin (Sigma, T9026) antibodies. To monitor CD109 cell surface expression, detached cells were stained with a PE-coupled CD109 antibody (BD Biosciences, 556040) for 30 min, washed with PBS and analyzed by flow cytometry.

27. Steu, S. *et al.* A procedure for tissue freezing and processing applicable to both intra-operative frozen section diagnosis and tissue banking in surgical pathology. *Virchows Arch.* **452**, 305–312 (2008).
28. Clough, T. *et al.* Protein quantification in label-free LC-MS experiments. *J. Proteome Res.* **8**, 5275–5284 (2009).