Nucleic Acids Research, 2003, Vol. 31, No. 14 ***e75***
DOI: 10.1093/nar/gng075

Heteromeric MAPPIT: a novel strategy to study modification-dependent protein–protein interactions in mammalian cells

Irma Lemmens, Sven Eyckerman, Lennart Zabeau, Dominiek Catteeuw, Els Vertenten, Kristin Verschueren[1], Danny Huylebroeck[1], Joël Vandekerckhove and Jan Tavernier*

Department of Medical Protein Research VIB09, Faculty of Medicine and Health Sciences, Ghent University, Albert Baertsoenkaai 3, B-9000 Ghent, Belgium and [1]Department of Developmental Biology VIB07 and Laboratory of Molecular Biology (Celgen), University of Leuven, Herestraat 49, B-3000 Leuven, Belgium

Received March 19, 2003; Revised and Accepted May 19, 2003

ABSTRACT

We recently reported a two-hybrid trap for detecting protein–protein interactions in intact mammalian cells (MAPPIT). The bait protein was fused to a STAT recruitment-deficient, homodimeric cytokine receptor and the prey protein to functional STAT recruitment sites. In such a configuration, STAT-dependent responses can be used to monitor a given bait–prey interaction. Using this system, we were able to demonstrate both modification-independent and tyrosine phosphorylation-dependent interactions. Protein modification in this approach is, however, strictly dependent on the receptor-associated JAK tyrosine kinases. We have now extended this concept by using extracellular domains of the heteromeric granulocyte/macrophage colony-stimulating factor receptor (GM-CSFR). Herein, the bait was fused to the βc chain and its modifying enzyme to the GM-CSFRα chain (or vice versa). We demonstrate several serine phosphorylation-dependent interactions in the TGFβ/Smad pathway using the catalytic domains of the ALK4 or ALK6 serine/threonine kinase receptors. In all cases tested, STAT-dependent signaling was completely abolished when mutant baits were used wherein critical serine residues were replaced by alanines. This approach operates both in transient and stable expression systems and may not be limited to serine phosphorylation but has the potential for studying various different types of protein modification-dependent interactions in intact cells.

INTRODUCTION

Post-translational modifications of proteins play a critical role in cellular functions as diverse as gene transcription, protein synthesis and degradation, cell cycle control, signal transduction and apoptosis. Often these modifications affect protein–protein interactions and the mechanisms underlying these regulated processes are the focus of intensive research.

Genetic methods to study protein–protein interactions were first explored in yeast cells using the 'yeast two-hybrid' method and have become an extremely useful tool for analyzing protein–protein interactions *in vivo* (1). In the original method, reconstitution in the nucleus of the DNA-binding domain (fused to bait) and activation domain (fused to prey) of the yeast transcription factor GAL4 lead to activation of a GAL4-sensitive promoter directing a reporter/selector gene. To circumvent the requirement for nuclear translocation, alternative yeast-based methods were developed, including the Sos or Ras recruitment system, wherein the bait–prey interaction is artificially tethered to the yeast cell membrane (2,3), and the ubiquitin-based split sensor system (USPS), which detects interactions occurring in the cytosol (4). An intrinsic disadvantage of using yeast cells is that interactions are detected in a lower eukaryote, implying a sub-optimal context for studying protein interactions of higher organisms. In particular, post-translational modifications are often hard to reproduce in yeast, although a 'yeast tri-hybrid' system, whereby a modifying enzyme is co-expressed in the yeast cell, was developed. However, expression of the mammalian modifying enzyme in yeast cells is often cytotoxic, and it should be carefully verified whether the post-translational modification is carried out correctly inside such yeast cells (5). Some of these limitations can be overcome by the use of mammalian cell systems. As such, the incorporation of β-galactosidase (6) or dihydrofolate reductase mutants in hybrid proteins (7) and fluorescence/bioluminescence resonance energy transfer (FRET/BRET) (e.g. using chimeras containing green fluorescent protein variants; reviewed in 8) may be methods of choice to trace interactions that occur in the cytosol. Modification-dependent interactions have been monitored via these mammalian complementation assays, but they rely on the presence and activity of endogenous modifying enzymes (9–11).

*To whom correspondence should be addressed. Tel: +32 9 331 33 01; Fax: +32 9 331 35 99; Email: jan.tavernier@rug.ac.be

We recently reported a novel method to study modification-dependent protein interactions in intact human cells using type I cytokine receptors. Ligands for these transmembrane glycoproteins include many interleukins, colony stimulating factors and hormones. Type I cytokine receptors lack intrinsic enzymatic activity but are constitutively associated with cytosolic tyrosine kinases of the JAK family. Ligand-induced clustering and/or reorganization of receptor subunits leads to trans-phosphorylation and activation of the JAKs, which subsequently phosphorylate tyrosine residues in the intracytoplasmic tail of the receptors allowing recruitment and activation by phosphorylation of signaling molecules. Among these, signal transducers and activators of transcription (STATs) play an essential role in the direct transmission of signals to the nucleus (12). Based on these insights, we developed a mammalian protein–protein interaction trap (MAPPIT) wherein the bait is fused to a STAT recruitment-deficient receptor, while the prey is coupled to a receptor fragment containing functional STAT recruitment sites. Bait–prey interaction leads to ligand-dependent STAT activation, which can be detected by the use of a STAT-responsive reporter/selector gene. Using this approach we were able to demonstrate modification-independent as well as tyrosine phosphorylation-dependent interactions. In the latter case, tyrosine phosphorylation of the bait is, however, strictly dependent on the receptor-associated JAK tyrosine kinase activity (13). To broaden the application field of the MAPPIT method to other types of protein modifications, we explored the use of a more complex heteromeric receptor system wherein the bait can be fused to one of the receptor chains and its modifying enzyme to the other. In such a configuration, bait and enzyme are brought into close proximity, favoring modification of the bait. We apply this concept here to serine phosphorylation-dependent Smad interactions in transforming growth factor β (TGFβ) family signaling.

MATERIALS AND METHODS

Constructs

The basic MAPPIT bait, prey, reporter and pSVsport-EpoR-LR constructs were generated as previously described (13). The leptin receptor (LR) fragment obtained from pSVsport-EpoR-LR by PacI and XbaI digestion was used to replace the IFNAR-1 fragment in pSVsport-hβc/IFNAR-1 construct to give pSVsport-βc-LR. The hIL-5Rα-LR chimeric receptor was obtained by transfer of the leptin fragment derived from the erythropoietin receptor (EpoR)-LR construct (13) by PacI and NotI digestion into the pSVsport-hIL-5Rα/IFNAR-2-2 construct (14). The pSVsport-IL-3Rα-LR and pSVsport-GM-CSFRα-LR constructs were generated by amplifying the extracellular domains by PCR with Pfu polymerase and exchanging the EpoR domain in the EpoR-LR construct by KpnI and PacI digestion. Replacing the LR domain by the mutant LRF3 fragment by PacI and NotI digestion generated the chimeric GM-CSFRα-LRF3 and βc-LRF3 receptors. Full-length human Smad3 and Smad5 cDNA fragments (cloned into pcdef plasmids, gifts from Dr Miyazono) were obtained by EcoRI and XhoI digestion, treated with Klenow polymerase and cloned into the SalI and Klenow polymerase blunted βc-LRF3 or GM-CSFRα-LRF3 vector. cDNA fragments encoding the cytoplasmic tail of ALK4 or its constitutive active variant (T206D, generated by site-directed mutagenesis) were amplified by PCR with Pfu using the PTZ18R-ALK4 plasmid (15) as a template. These fragments were cloned into pGAD424 or pGBT9, respectively, and subsequently transferred into the GM-CSFRα-LRF3 or βc-LRF3 vectors by ligating blunted (Klenow) EcoRI–BamHI fragments into a blunted SalI restriction site. The constitutively active variant of ALK6 contains a Q203D substitution and the relevant fragment was amplified by PCR using pCS2-CA-ALK6 (a gift from Dr Tylzanowski) and was cloned by SalI and NotI digestion into the chimeric receptor chains. The pMG2 vector originates from the pMG1 vector (13) and contains amino acids 905–918 from the gp130 chain in duplicate. The Smad-binding domain (SBD) of Smad-interacting protein-1 (SIP1) was derived from a pCS3-SIP1SBD construct encoding a 183 amino acid long segment encompassing the SBD (16) corresponding to amino acids 315–498 in full-length mouse SIP1. The cDNA fragment encoding this polypeptide was isolated by SmaI and XhoI digestion of the plasmid and cloned into the pMG2-SVT and pMET7-Flag vectors with a blunted EcoRI and a XhoI site. Human Smad4 (missing the first 3 amino acids, a gift from Dr Miyazono) was obtained by EcoRI and XhoI digestion and cloned into EcoRI and XhoI digested pMG2-SVT and pMET7-flag vectors. The Smad3/5 S→A mutants and the insertion of extra leucine residues (1, 2 or 3 at position 851 within the predicted transmembrane region of the leptin receptor) in the transmembrane domains from the chimeric receptor chains were generated using the Quick Change site-directed mutagenesis procedure (Stratagene). The pRK5-JAK2 construct was a kind gift from Dr Constantinescu. The dual expression vector is based on the pcDNA5/FRT vector (Invitrogen). A CMV-newMCS-BGH polyadenylation signal was amplified and inserted into the BglII and MfeI digested pcDNA5/FRT vector, resulting in a pcDNA5/FRT vector with two CMV-MCS-BGH cassettes. The βc-LRF3+1L-Smad3 chain was inserted in the dual expression vector by KpnI and NotI digestion, the GM-CSFRα-LRF3+1L-CA-ALK4 chain using SbfI/NsiI and SnaBI/SmaI digestions.

Cell lines, transfections, reporter assays and FACS analysis

The dual expression vector pcDNA5/FRT containing the GM-CSFRα-LRF3+1L-CA-ALK4 and βc-LRF3+1L-Smad3 chains was stably integrated in a HEK293-16 cell line, as previously described (13), using the Flp-In recombinase reaction (Invitrogen) and after selection on hygromycin (100 μg/ml) for 10 days. Culture conditions, transfection procedures and luciferase and β-galactosidase assays for HEK293T cells are described elsewhere (17). For a typical luciferase experiment, 4×10^5 cells were transfected with the desired constructs in the presence of luciferase and β-gal reporter genes. After 48 h, cells were left untreated (UN) or were treated with 10 ng/ml granulocyte/macrophage colony-stimulating factor (GM-CSF) or with the corresponding ligands at doses as indicated. After another 24 h, luciferase activity from triplicate samples was determined and normalized against β-galactosidase activity. The results shown are representative of at least three independent transfection

experiments. Recombinant hGM-CSF and hIL-3 were purchased from PeproTech Inc. (Rocky Hill, NJ). hIL-5 was produced in COS-1 cells and purified using H30 affinity chromatography as described before (18). hEpo was obtained from R&D Systems. GM-CSF was typically used at 10 ng/ml; the concentrations of interleukin 3 (IL-3), interleukin 5 (IL-5) and Epo used are as indicated. GM-CSFRα-LRF3+1L-CA-ALK4 expression was monitored with 2 μg/ml of an anti-human GM-CFSRα monoclonal antibody (sc-456; SanverTech) and 4 μg/ml Alexa488-conjugated goat anti-mouse IgG (Molecular Probes). The βc-LRF3+1L-Smad3 chain was stained with 2 μg/ml of a biotin-conjugated mouse anti-human monoclonal antibody against the common β-chain (CDw131; PharMingen International) and 4 μg/ml Alexa488-conjugated goat anti-mouse IgG (Molecular Probes). FACS analysis was performed on a FACSCalibur (Becton Dickinson).

Immunoprecipitation and western blot analysis

To demonstrate Smad3 serine phosphorylation, we transiently transfected ~3 × 10^6 HEK293T cells with plasmid vectors encoding the GM-CSFRα-LRF3-CA-ALK4 and βc-LRF3-Smad3 receptor chains with one extra leucine in their transmembrane part or the GM-CSFRα-LRF3-CA-ALK4 and βc-LRF3-Smad3 S→A (critical serines of Smad3 mutated to alanines) receptor chains or empty vector. Sixty-five hours after transfection cells were starved in serum-free medium for 5 h and were left untreated or were stimulated with 10 ng/ml GM-CSF for 10 min. Cleared lysates [in modified RIPA buffer, 50 mM Tris–HCl pH 8.0, 200 mM NaCl, 1% NP-40, 0.5% DOC, 0.05% SDS, 2 mM EDTA, 1 mM Na_3VO_4, 1 mM NaF, 20 mM β-glycerophosphate and Complete protease inhibitor cocktail (Roche)] of the untreated and stimulated cells were incubated with 2 μg mouse anti-human-β common antibody (CDw131; PharMingen International) and Protein G-Sepharose (Amersham Pharmacia Biotech). After immunoprecipitation, polyacrylamide gel electrophoresis and western blotting, serine phosphorylation of Smad3 was detected with a polyclonal anti-phosphoserine Smad3 antibody (a gift of Dr ten Dijke). The presence of the chimeric chain was visualized using an anti-LR antibody directed against residues 942–953 of the cytoplasmic domain of the leptin receptor.

For STAT3 and JAK2 phosphorylation 4 × 10^5 HEK293T cells were transfected with the desired constructs. After 65 h the cells were starved in serum-free medium for 5 h, left untreated or were stimulated with 10 ng/ml GM-CSF for 15 or 10 min, respectively. Cells were lysed in 200 μl modified RIPA buffer, of which 50 μl was loaded on a 7.5% polyacrylamide gel for gel electrophoresis and subsequent western blotting. STAT3 phosphorylation was detected with the Phospho-STAT3-Tyr705 antibody (Cell Signaling) and expression levels were verified using an anti-STAT3 antibody (Transduction Laboratories). To detect JAK2 phosphorylation, extra JAK2 was transfected (0.001 μg pRK5-JAK2, except for the mock transfection) and was revealed using an anti-phospho-JAK2 antibody (Upstate Biotechnology) and expression levels were checked using an anti-JAK2 antibody (Upstate Biotechnology).

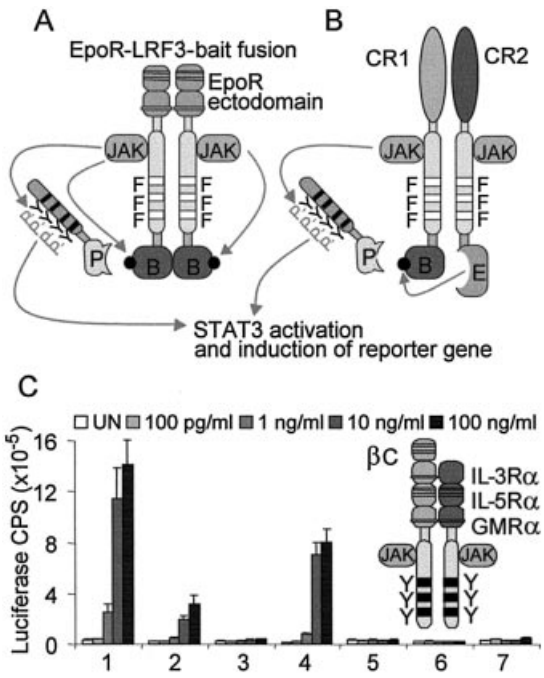

Figure 1. Layout of the heteromeric MAPPIT method. A representation of the different components and mechanism of the cytokine receptor-based interaction trap is shown. (**A**) The homodimeric MAPPIT approach is presented (see also 13). (**B**) The heteromeric MAPPIT principle. The ligand-binding domain in (B) is from a heteromeric cytokine receptor (with subunits CR1 and CR2), which replaces the erythropoietin receptor (EpoR) from (A). The bait, 'B' is fused to the CR1 chain whereas its modifying enzyme, 'E' is fused to the CR2 chain. Ligand-induced activation of the receptor-associated JAK tyrosine kinase puts the receptor complex in a 'stand-by' mode without induction of detectable reporter activity. No STAT3 recruitment and activation can occur due to the Y1138F mutation in the cytosolic domain of the LR. Additional Y985/1077F mutations eliminate adapter and/or negative feedback mechanisms. Complementation is induced upon interaction between the modified bait (with a black dot representing the modification) and prey, 'P', which leads to recruitment of the C-terminal part of gp130 containing four functional STAT3-binding sites. Subsequent STAT3 phosphorylation and activation induces luciferase activity under control of the rPAP1 promoter. Functional STAT3 recruitment sites are presented as black boxes; positions of the Y→F mutations within the LR-F3 are shown as open boxes. Hinge regions, drawn as shaded boxes, preceding the prey, bait and modifying enzyme provide additional flexibility in the chimeric polypeptides. (**C**) Dose–response curves of heteromeric chimeric receptors. HEK293T cells were transiently transfected with different receptor chimeras: 1, EpoR-LR; 2, IL3Rα-LR + βc-LR; 3, IL5Rα-LR + βc-LR; 4, GMRα-LR + βc-LR; 5, IL3Rα-LR; 6, IL5Rα-LR; 7, GMRα-LR. Bars and error bars represent mean values of the luciferase measurements and SD, respectively. LR, cytoplasmic domain of the wild-type leptin receptor. The βc-LR chain alone was stimulated with all the different ligands at a concentration of 10 ng/ml and showed no response in all cases tested (data not shown).

RESULTS

Design of the heteromeric MAPPIT approach

The MAPPIT system is outlined in Figure 1A. A chimeric cytokine receptor, composed of the extracellular part of the homodimeric EpoR fused to the transmembrane and cytosolic domains of the LR, can activate STAT3, but is made inactive by a Y1138F mutation, which is known to eliminate a critical STAT3 recruitment site. Ligand administration leads to JAK activation, but in the absence of a functional recruitment site

no STAT3-dependent signal transduction is initiated. In addition, similar Y→F mutations at positions Y985 and Y1077 prevent negative feedback and therefore can enhance signaling (17). This signaling-deficient LR variant is further referred to as LRF3. Upon co-expression of a chimeric gp130–prey construct, which contains four functional STAT3 recruitment sites in the gp130 moiety, complementation occurs leading to efficient STAT3 activation. These STAT3 proteins dimerize and subsequently translocate to the nucleus, allowing quantification of STAT3 activation using a STAT3-responsive rat pancreatitis associated protein 1 (rPAP1) promoter fused to the luciferase reporter gene (19). This approach was previously used for analytical purposes, but was also optimized to screen complex cDNA libraries (13). Importantly, the bait–prey interaction may be modification-dependent, but in this configuration bait modification is limited to tyrosine phosphorylation by the receptor-associated JAKs.

In order to extend this strategy to other modification-dependent bait–prey interactions, we investigated whether a heteromeric instead of a homodimeric receptor complex could be used. As shown in Figure 1B, this would facilitate the formation of a receptor complex containing at the cytoplasmic side both the bait and its modifying enzyme. We anticipated that in such a case modification of the bait would occur due to the enforced proximity of enzyme and its cognate substrate, induced by either formation of a preformed receptor complex or by ligand-induced receptor oligomerization.

Receptor complexes for IL-3, IL-5 and GM-CSF are composed of a ligand-specific α-subunit and a common β (βc) chain shared by all three cytokines (20). We first generated chimeric IL-3Rα, IL-5Rα, GM-CSFRα and βc receptors whereby the transmembrane and cytosolic domains were replaced by those of the wild-type LR. A clear difference in ligand-induced signaling output was observed in HEK293T cells transfected with the different IL-3Rα/βc, IL-5Rα/βc and GM-CSFRα/βc chimeric receptor combinations. The combination using the extracellular domain of GM-CSFRα and of βc was most efficient and was therefore selected for further analysis (Fig. 1C).

Detection of serine phosphorylation-dependent interactions

To test the heteromeric MAPPIT concept, serine phosphorylation-dependent Smad interactions were investigated. Smad proteins are components of the TGFβ signalling machinery, which regulate a wide range of biological and cellular processes. In one proposed and generally accepted model, signal transduction via TGFβ receptors is initiated when the ligand induces assembly of a heteromeric complex of type II and type I transmembrane serine/threonine kinase receptors. The type II kinase then phosphorylates the type I receptor in a conserved glycine-serine-rich (GS) domain. This activates the type I kinase, which subsequently recognizes and phosphorylates members of the intracellular Smad signal transduction pathway. The type I receptors specifically recognize the Smad subgroup known as receptor-activated Smads (R-Smads). These include Smad1, Smad2, Smad3, Smad5 and Smad8. The R-Smads have two conserved and well-characterized functional domains, i.e. the MH1 and MH2 domains, which are separated by a proline-rich domain. Receptor-mediated phosphorylation of the C-terminally located SSXS motif contributes to the relief of a mutually inhibitory interaction between these two domains and leads to R-Smad activation and subsequent accumulation in the nucleus. Prior to nuclear translocation, phosphorylated R-Smads bind to a common Smad, Smad4. Such R-Smad–Smad4 complexes participate in DNA binding and/or recruitment of transcriptional cofactors, thereby modulating gene transcription (21).

To create a signaling-deficient heteromeric receptor complex, the cytoplasmic tail from the wild-type LR in the GM-CSFRα/LR and βc/LR chimeras was replaced by its signaling-deficient LRF3 variant. These chimeras were used to probe interactions within the context of Smad signaling. Therefore, the cytoplasmic tail of a constitutively active form of the mouse activin-like kinase 6 type I receptor, carrying a Q203D mutation in its GS domain (CA-ALK6), and the bait, full-length human Smad5, were fused in-frame to the GM-CSFRα-LRF3 and βc-LRF3 receptor chains. Optimal bait–prey interaction-dependent signaling required two optimization steps. First, comparison of the two reciprocal combinations showed that more efficient signaling was obtained when the modifying enzyme CA-ALK6 was fused to the βc-LRF3 chimeric receptor and the bait, Smad5, to the GM-CSFRα-LRF3 receptor chain. Second, based on recent reports that optimal signaling by class I cytokine receptors may depend on the relative orientation between the extracellular and intracellular domains within a receptor complex (22,23), we inserted one, two or three extra leucine residues within the transmembrane parts of each of the receptor chimeras. The combination where both chimeric receptor chains contained two extra leucines appeared most favorable for the CA-ALK6/Smad5 combination, although other combinations also showed enhanced efficacy. Results for the optimal combination together with negative controls are shown in Figure 2A. Importantly, the negative control, whereby the serine residues in the SSXS phosphorylation motif of Smad5 were replaced by alanines, was completely signaling-deficient. Additional controls included absence of the ALK6 catalytic domain, absence of the Smad5 bait, absence of Smad4 (i.e. the presence of an irrelevant gp130–SV40 large T prey) and absence of the gp130 chain (wild-type Smad4). In all these cases no signal was obtained, indicating that STAT3 activation was strictly dependent on the interaction between the serine-phosphorylated bait and the gp130–Smad4 prey. Signaling was also strictly ligand-dependent.

Similarly, the cytoplasmic tail of a constitutively active form of the mouse activin-like kinase 4 type I receptor (15) (obtained by a T206D mutation in the GS domain, and further referred to as CA-ALK4) and full-length human Smad3 were fused in-frame to the GM-CSFRα-LRF3 and the βc-LRF3 chimeric receptors. Serving as prey were either full-length Smad4 or a polypeptide encompassing the SBD of the transcriptional repressor SIP1 (16), which were fused C-terminally to the STAT3 recruitment sites of the gp130 chain. Here, for both preys, the combination whereby CA-ALK4 was fused to the GM-CSFRα-LRF3 receptor chain and Smad3 to the βc-LRF3 chimeric receptor and whereby both receptor chimeras contained a single additional leucine residue within the transmembrane region gave best results (shown in Fig. 2B for the Smad3–Smad4 interaction, with the Smad3 S→A mutant as negative control).

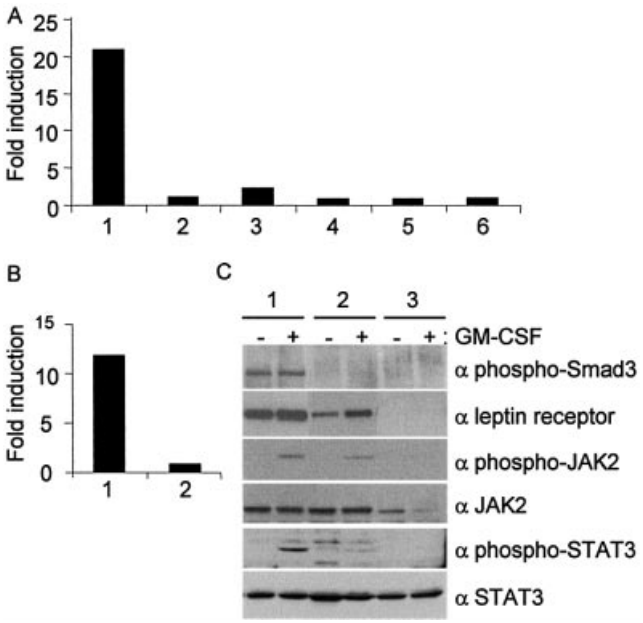

Figure 2. Detection of phosphoserine-dependent Smad interactions using the heteromeric MAPPIT procedure. (**A**) HEK293T cells were transfected with plasmids encoding: 1, βc-LRF3+2L-CA-ALK6, GM-CSFRα-LRF3+2L-Smad5 and gp130-Smad4; 2, βc-LRF3+2L-CA-ALK6, GM-CSFRα-LRF3+2L-Smad5 S→A (a mutant construct wherein critical serines of Smad5 are replaced by alanines) and gp130-Smad4; 3, βc-LRF3+2L (no CA-ALK6), GM-CSFRα-LRF3+2L-Smad5 and gp130-Smad4; 4, βc-LRF3+2L-CA-ALK6, GM-CSFRα-LRF3+2L (no Smad5) and gp130-Smad4; 5, βc-LRF3+2L-CA-ALK6, GM-CSFRα-LRF3+2L-Smad5 and gp130-SVT (no Smad4); 6, βc-LRF3+2L-CA-ALK6, GM-CSFRα-LRF3+2L-Smad5 and Smad4 (no gp130). Average values for relative luciferase activities (*x*-fold increase, luciferase values obtained from stimulated cells with respect to values derived from untreated cells) are shown. (**B**) HEK293T cells were transfected with plasmids encoding: 1, GM-CSFRα-LRF3+1L-CA-ALK4, βc-LRF3+1L-Smad3 and gp130-Smad4; 2, GM-CSFRα-LRF3+1L-CA-ALK4, βc-LRF3+1L-Smad3 S→A and gp130-Smad4. (**C**) HEK293T cells were transfected with plasmids encoding: 1, GM-CSFRα-LRF3+1L-CA-ALK4, βc-LRF3+1L-Smad3, and gp130-Smad4; 2, GM-CSFRα-LRF3+1L-CA-ALK4 and βc-LRF3+1L-Smad3 S→A (critical serines of Smad3 mutated to alanines) and gp130-Smad4; 3, mock control with empty vector. Cells were either left untreated (–) or were stimulated with 10 ng/ml GM-CSF (+). α phospho-Smad3, serine phosphorylated Smad3; α leptin receptor, the chimeric chain detected using an anti-leptin receptor antibody; α phospho-JAK2, phosphorylated JAK2; α JAK2, total JAK2; α phospho-STAT3, phosphorylated STAT3; α STAT3, total STAT3.

Analysis of bait and prey modifications

Serine phosphorylation of Smad3 by the CA-ALK4 kinase was checked by western blot analysis using an antibody specific for serine-phosphorylated Smad3 (Fig. 2C, α phospho-Smad3). Smad3 phosphorylation was ligand-independent, suggesting the existence of preformed receptor complexes. Similar ligand-independent serine phosphorylation of Smad3 was observed in an isogenic cell population stably expressing the chimeric receptor chains (as discussed below; data not shown). Importantly, although bait modification was independent of ligand, JAK activation as well as subsequent tyrosine phosphorylation of STAT3 was strictly dependent on GM-CSF administration (see Fig. 2C, α phospho-JAK2 and α phospho-STAT3).

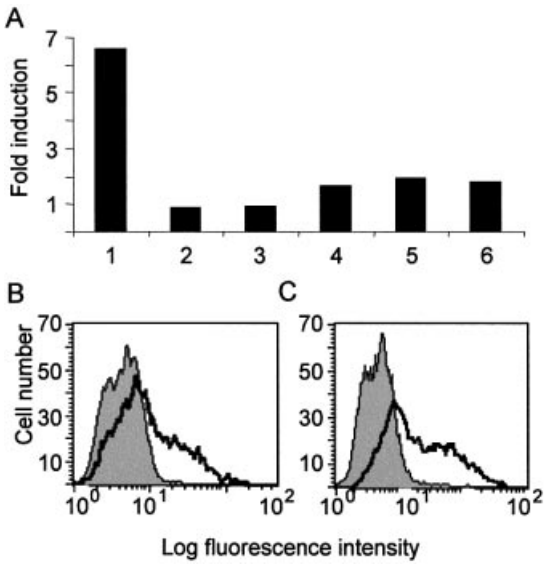

Figure 3. Heteromeric MAPPIT using cells stably expressing the chimeric receptor chains. (**A**) An isogenic cell pool stably expressing the GM-CSFRα-LRF3+1L-CA-ALK4 and βc-LRF3+1L-Smad3 receptor chains was selected and transfected with plasmids encoding: 1, gp130-Smad4; 2, Smad4 (no gp130); 3, gp130-SVT (no Smad4); 4, gp130-Smad4 combined with βc-LRF3+1L-Smad3 S→A mutant; 5, gp130-Smad4 combined with GM-CSFRα-LRF3+1L (no ALK4); 6, gp130-Smad4 combined with Smad4 (no gp130 chain). Average values for relative luciferase activities (*x*-fold increase, stimulated versus untreated) are shown. (**B** and **C**) FACS analyses, with filled curves representing parental HEK293-16 cells and open curves representing the selected isogenic cell pool expressing both receptor chimeras. (B) and (C) show the expression of the βc-LRF3+1L-Smad3 and the GM-CSFRα-LRF3+1L-CA-ALK4 chains, respectively.

Analysis of the Smad3–Smad4 interaction in stably transfected cells

Since the above mentioned experiments were performed by transient overexpression, we subsequently analyzed whether heteromeric MAPPIT could also be set up in a stable cell system with more physiological expression levels of the receptor chimeras. For this purpose a dual expression vector was generated based on pcDNA5/FRT (Invitrogen). This plasmid vector facilitates a Flp-In recombinase reaction, allowing fast and easy selection of a so-called isogenic cell population by stable integration into a transcriptionally active locus in the cell genome. This basic vector contains a CMV promoter followed by a multiple cloning site (MCS) and a bovine growth hormone polyadenylation signal (BGH). A second, similar expression cassette, but with a different MCS, was inserted, allowing expression of both heteromeric receptor chimeras from a single vector. As shown by FACS analysis in Figure 3B and C, stable expression of both receptor chimeras was obtained after recombinase-assisted integration. Upon transient expression of the gp130-Smad4 prey, a clear ligand-dependent signal was obtained. Significantly, this signal was inhibited by coexpressing a receptor chimera containing the S→A mutated Smad3 or a receptor chimera lacking the modifying enzyme and by the coexpression of wild-type Smad4 not linked to gp130 (Fig. 3A).

DISCUSSION

The complexity of the human proteome is estimated to exceed its encoding genome by at least one order of magnitude. This enhanced molecular variation is mostly attributed to post-translational modification of proteins, combined with alternative splicing. The best-studied type of protein modification is phosphorylation, whereby covalent attachment of phosphate groups can act as an activity modulator or even as an on/off switch controlling the activity of receptors, signaling molecules, enzymes or transcription factors. As such, protein phosphorylation is involved in a wide range of cellular processes, including growth, proliferation, differentiation, motility, membrane transport, and in their control by extracellular signals. Together, this ultimately results in a wide range of controlled physiological processes at the body level. Underscoring the importance of protein phosphorylation is the complexity of the human kinome, which with 518 known kinases represents ~1.7% of the complete protein encoding gene set of the genome, and the estimation that about one-third of all proteins can be a target for phosphorylation (24). Even further underscoring the importance and the complexity of controlled protein modifications is the fact that besides phosphorylation, many other types of functionally important post-translational modifications have been described, including proteolytic cleavage, ubiquitination, sumoylation, acetylation, acylation and methylation.

Here, we describe a novel method to facilitate the study of modification-dependent protein–protein interactions in the physiologically optimal context of intact mammalian cells. We previously reported a two-hybrid approach based on type I cytokine receptor signaling. Using this MAPPIT approach we were able to demonstrate tyrosine phosphorylation-dependent interactions between the pY402 motif of the EpoR and the cytokine-inducible SH2-containing protein (CIS) and suppressor of cytokine signaling-2 protein (SOCS-2) (13). Additional interaction-dependent signals were also demonstrated using other preys, including PLCγ, PI_3K, Grb2, SHP-2 and Vav (to be reported elsewhere). Given the homodimeric structure of the EpoR, MAPPIT applications are restricted to homodimeric baits. We therefore explored the use of the extracellular domains of the heteromeric GM-CSF receptor (GM-CSFRα and βc chains), whereby the bait is fused to one of the receptor chains and its modifying enzyme is fused to the other. In such a configuration, receptor complex formation brings bait and modifying enzyme into close proximity, favoring modification of the bait. In principle, any type of protein modification can be studied, as long as the modifying enzyme can be functionally coupled to the chimeric receptor chain. As model interactions, we selected several previously documented protein interactions in the TGFβ family signaling pathways, i.e. between Smad5 and Smad4, Smad3 and Smad4 and Smad3 and SIP1, which are all strictly dependent on correct serine phosphorylation of the R-Smad. Interestingly, as shown in Figure 2C, Smad3 serine phosphorylation is observed in the absence of ligand, suggesting that the GM-CSFRα/βc chimeric receptors likely exist as preformed complexes allowing spontaneous bait modification. Preformed complex formation in this cytokine receptor family is well documented for the EpoR (25–28). It is unclear at present which parts of the GM-CSFRα/LRF3 and βc/LRF3 receptor chimeras induce complex formation, however, this bait modification assay can be used as read-out for structure/function analyses. It should be stressed that the heteromeric MAPPIT read-out is completely independent of such preformed receptor complexes.

Given the availability of numerous constitutively active kinases, heteromeric MAPPIT is very well suited to study most, if not all, phosphorylation-dependent protein–protein interactions. It should be stressed that the substrate–enzyme proximity induced upon formation of the chimeric receptor complex is also likely to suffice for other types of protein modifications, e.g. in cases where this is controlled by regulated expression or localization of the modifying enzyme.

As has been shown for gp130 and for the EpoR, proper orientation of signaling subunits within a cytokine receptor complex is necessary for efficient activation and for triggering downstream signaling cascades. Constantinescu *et al*. concluded that orientation-dependent signaling is transmitted through the transmembrane domain toward the cytosolic juxtamembrane domain, suggesting that the transmembrane segment is important for preserving the orientation between the extracellular domain and the cytoplasmic domain (22). Greiser *et al*. reported two distinct requirements for efficient signaling through the gp130 chain: first, JAKs have to be located close to the membrane for full activation and, secondly, the cytoplasmic regions of the receptors have to be precisely orientated to allow STAT activation (23). Here we observed that inserting extra leucines into the transmembrane part of the chimeric GM-CSFRα-LRF3 and βc-LRF3 chains also clearly affected signaling efficiency. In the case of heteromeric MAPPIT, however, optimal receptor configurations appear to differ from case to case: insertion of two leucine residues in either receptor chain led to optimal signaling for the CA-ALK6/Smad5 set-up, whereas insertion of a single leucine residue was required for optimal signaling in the case of the CA-ALK4/Smad3 combination. It should be stressed that signaling in heteromeric MAPPIT depends on, besides optimal JAK activation, modification of the bait. A likely explanation for this bait-dependent signaling variability is therefore that the efficiency of bait modification may also depend on the relative orientations of the target serine residues in the Smad baits and the catalytic domains of the ALK modifying enzymes.

An important advantage of MAPPIT over other mammalian two-hybrid methods is the possibility to perform interactor hunts in complex cDNA libraries (13). We therefore explored the use of this heteromeric approach in stable cell systems. A dual expression vector was constructed, allowing rapid selection of a cell population stably expressing both receptor chimeras. Since both receptor chimeras are preferably expressed at comparable levels, we used a vector containing two expression cassettes based on the same CMV promoter. This is not the case for bi-cistronic vectors that make use of an internal ribosomal entry site (IRES) sequence since a lower expression level of the gene cloned behind the IRES sequence is often observed (29). In this isogenic cell pool, which stably expresses the GM-CSFRα-LRF3+1L-CA-ALK4 and βc-LRF3+1L-Smad3 chains, serine phosphorylation-dependent interaction between Smad3 and Smad4 was clearly detected. Furthermore, overexpression of βc-LRF3+1L-Smad3 S→A, of GM-CSFRα-LRF3+1L lacking the modifying enzyme or of

wild-type Smad4 inhibited this signal, showing that the read-out is strictly dependent on proper Smad3 serine phosphorylation and on JAK-dependent tyrosine phosphorylation of the gp130 prey. Clearly, screening of complex cDNA libraries may yield preys interacting with either the (modified) bait or with the modifying enzyme. Albeit the latter may prove interesting by themselves, preys specifically interacting with the bait can easily be discerned by the use of receptor chimeras lacking the bait or containing a mutant bait lacking the modification site, as was demonstrated before (13).

In conclusion, we have developed a novel method to study modification-dependent protein–protein interactions in intact mammalian cells. We demonstrated efficient detection of phosphoserine-dependent protein–protein interactions, however, the use of such engineered heteromeric receptor systems has the potential to be extended to many other types of modification-dependent protein–protein interactions.

ACKNOWLEDGEMENTS

We thank Dr ten Dijke for the anti-phosphoserine Smad3 antibody and Dr Constantinescu for the pRK5-JAK2 construct. We are indebted to Dr Miyazono for the Smad plasmids and Dr Tylzanowski for the ALK6 construct. This project was supported by het Vlaams Gewest-IWT (GBOU 010090) and VIB funding. Research at VIB07 was in addition to VIB funding supported by OT/00/44.

REFERENCES

1. Fields,S. and Song,O. (1989) A novel genetic system to detect protein-protein interactions. *Nature*, **340**, 245–246.
2. Aronheim,A., Zandi,E., Hennemann,H., Elledge,S.J. and Karin,M. (1997) Isolation of an AP-1 repressor by a novel method for detecting protein-protein interactions. *Mol. Cell. Biol.*, **17**, 3094–3102.
3. Broder,Y.C., Katz,S. and Aronheim,A. (1998) The ras recruitment system, a novel approach to the study of protein-protein interactions. *Curr. Biol.*, **8**, 1121–1124.
4. Johnsson,N. and Vershavsky,A. (1994) Split ubiquitin as a sensor of protein interactions *in vivo*. *Proc. Natl Acad. Sci. USA*, **19**, 10340–10344.
5. Osborne,M.A., Dalton,S. and Kochan,J.P. (1995) The yeast tribrid system—genetic detection of trans-phosphorylated ITAM-SH2-interactions. *Biotechnology*, **13**, 1474–1478.
6. Rossi,F., Charlton,C.A. and Blau,H.M. (1997) Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation. *Proc. Natl Acad. Sci. USA*, **94**, 8405–8410.
7. Joung,J.K., Ramm,E.I. and Pabo,C.O. (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc. Natl Acad. Sci. USA*, **97**, 7382–7387.
8. Pollok,B.A. and Heim,R. (1999) Using GFP in FRET-based applications. *Trends Cell Biol.*, **9**, 57–60.
9. Shioda,T., Andriole,S., Yahata,T. and Isselbacher,K.J. (2000) A green fluorescent protein-reporter mammalian two-hybrid system with extrachromosomal maintenance of a prey expression plasmid: application to interaction screening. *Proc. Natl Acad. Sci. USA*, **97**, 5220–5224.
10. Sorkin,A., McClure,M., Huang,F. and Carter,R. (2000) Interaction of EGF receptor and grb2 in living cells visualized by fluorescence resonance energy transfer (FRET) microscopy. *Curr. Biol.*, **10**, 1395–1398.
11. Ozawa,T., Kaihara,A., Sato,M., Tachihara,K. and Umezawa,Y. (2001) Split luciferase as an optical probe for detecting protein-protein interactions in mammalian cells based on protein splicing. *Anal. Chem.*, **73**, 2516–2521.
12. Bromberg,J. and Darnell,J.E. (2000) The role of STATs in transcriptional control and their impact on cellular function. *Oncogene*, **19**, 2468–2473.
13. Eyckerman,S., Verhee,A., Van der Heyden,J., Lemmens,I., Van Ostade,X., Vandekerckhove,J. and Tavernier,J. (2001) Design and application of a cytokine-receptor-based interaction trap. *Nature Cell Biol.*, **3**, 1114–1119.
14. Van Ostade,X., Schauvliege,L., Pattyn,E., Verhee,A., Vandekerckhove,J. and Tavernier,J. (2000) A sensitive and versatile bioassay for ligands that signal through receptor clustering. *J. Interferon Cytokine Res.*, **20**, 79–87.
15. Verschueren,K., Dewulf,N., Goumans,M.J., Lonnoy,O., Feijen,A., Grimsby,S., Vandi Spiegle,K., ten Dijke,P., Moren,A., Vanscheeuwijck,P. *et al*. (1995) Expression of type I and type IB receptors for activin in midgestation mouse embryos suggests distinct functions in organogenesis. *Mech. Dev.*, **52**, 109–123.
16. Verschueren,K., Remacle,J.E., Collart,C., Kraft,H., Baker,B.S., Tylzanowski,P., Nelles,L., Wuytens,G., Su,M.T., Bodmer,D. *et al*. (1999) SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5′-CACCT sequences in candidate target genes. *J. Biol. Chem.*, **274**, 20489–20498.
17. Eyckerman,S., Broekaert,D., Verhee,A., Vandekerckhove,J. and Tavernier,J. (2000) Identification of the Y985 and Y1077 motifs as SOCS3 recruitment sites in the murine leptin receptor. *FEBS Lett.*, **486**, 33–37.
18. Zabeau,L., Van der Heyden,J., Broekaert,D., Verhee,A., Vandekerckhove,J., Wu,S.J., Chaiken,I., Heinrich,P., Behrmann,I. and Tavernier,J. (2001) Neutralizing monoclonal antibodies can potentiate IL-5 signaling. *Eur. J. Immunol.*, **31**, 1087–1097.
19. Broekaert,D., Eyckerman,S., Lavens,D., Verhee,A., Waelput,W., Vandekerckhove,J. and Tavernier,J. (2002) Comparison of leptin- and interleukin-6-regulated expression of the rPAP gene family: evidence for differential co-regulatory signals. *Eur. Cytokine Netw.*, **13**, 78–85.
20. Carr,P.D., Gustin,S.E., Church,A.P., Murphy,J.M., Ford,S.C., Mann,D.A., Woltring,D.M., Walker,I., Ollis,D.L. and Young,I.G. (2001) Structure of the complete extracellular domain of the common beta subunit of the human GM-CSF, IL-3 and IL-5 receptors reveals a novel dimer configuration. *Cell*, **104**, 291–300.
21. Attisano,L. and Wrana,J.L. (2002) Signal transduction by the TGF-beta superfamily. *Science*, **296**, 1646–1647.
22. Constantinescu,S.N., Huang,L.J., Nam,H. and Lodish,H.F. (2001) The erythropoietin receptor cytosolic juxtamembrane domain contains an essential, precisely oriented, hydrophobic motif. *Mol. Cell*, **7**, 377–385.
23. Greiser,J.S., Stross,C., Heinrich,P.C., Behrmann,I. and Hermanns,H.M. (2002) Orientational constraints of the gp130 intracellular juxtamembrane domain for signaling. *J. Biol. Chem.*, **277**, 26959–26965.
24. Manning,G., Whyte,D.B., Martinez,R., Hunter,T. and Sudarsanam,S. (2002) The protein kinase complement of the human genome. *Science*, **298**, 1912–1934.
25. Constantinescu,S.N., Keren,T., Socolovsky,M., Nam,H., Henis,Y.I. and Lodish,H.F. (2001) Ligand-independent oligomerization of cell-surface erythropoietin receptor is mediated by the transmembrane domain. *Proc. Natl Acad. Sci. USA*, **98**, 4379–4384.
26. Livnah,O., Stura,E.A., Middleton,S.A., Johnson,D.L., Jolliffe,L.K. and Wilson,I.A. (1999) Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. *Science*, **283**, 987–990.
27. Remy,I., Wilson,I.A. and Michnick,S.W. (1999) Erythropoietin receptor activation by a ligand-induced conformation change. *Science*, **283**, 990–993.
28. Kubatzky,K.F., Ruan,W., Gurezka,R., Cohen,J., Ketteler,R., Watowich,S.S., Neumann,D., Langosch,D. and Klingmuller,U. (2001) Self assembly of the transmembrane domain promotes signal transduction through the erythropoietin receptor. *Curr. Biol.*, **11**, 110–115.
29. Mizuguchi,H., Xu,Z., Ishii-Watabe,A., Uchida,E. and Hayakawa,T. (2000) IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. *Mol. Ther.*, **1**, 376–382.