
Toward the semisynthesis of multidomain transmembrane receptors: Modification of Eph tyrosine kinases

NIKHIL SINGLA,^{1,2} JUHA PEKKA HIMANEN,^{1,2} TOM W. MUIR,^{2,3} AND DIMITAR B. NIKOLOV^{1,2}

¹Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA

²Tri-Institutional Chemical Biology Program, Weill Cornell Medical College, New York, New York 10021, USA

³Laboratory of Synthetic Protein Chemistry, The Rockefeller University, New York, New York 10021, USA

(RECEIVED April 6, 2008; FINAL REVISION June 29, 2008; ACCEPTED July 11, 2008)

Abstract

Expressed protein ligation (EPL) is a protein engineering approach that allows the modification or assembly of a target protein from multiple recombinant and synthetic polypeptides. EPL has been previously used to modify intracellular proteins and small integral membrane proteins for structural and functional studies. Here we describe the semisynthetic site-specific modification of the complete, multidomain extracellular regions of both A and B classes of Eph receptor tyrosine kinases. We show that the ectodomains of these receptors can be ligated to different peptides under carefully established experimental conditions, while their biological activity is retained. This work extends the boundaries of the EPL technique for semisynthesis of multidomain, extracellular, disulfide-bonded, and glycosylated proteins and highlights its potential application for reconstituting entire single-pass transmembrane proteins.

Keywords: expressed protein ligation; Eph receptor tyrosine kinase; multidomain extracellular region; inteins; peptides; ephrins

The protein semisynthesis technique expressed protein ligation (EPL) has been extensively utilized for the synthesis and semisynthesis of proteins that are otherwise difficult to obtain using purely recombinant methods (Muir 2003). The use of this strategy was also extended for the modification of proteins with a number of chemical tags (Lesachere et al. 2002). Though EPL has been applied successfully to a large number of intracellular proteins (and for some small secreted proteins, which can be expressed intracellularly and behave biochemically as intracellular ones), its application to

transmembrane proteins with large multidomain extracellular regions has yet to be documented. The difficulty here arises from the fact that oxidizing conditions are required to maintain disulfide-containing extracellular regions in their correctly folded state, while the intein domains used in EPL need a reducing environment to retain their correctly folded state and enzymatic activity. Furthermore, thiol molecules like dithiothreitol (DTT) or ethanethiol, with moderate to strong reducing properties, are required for the generation of the reactive C-terminal α -thioester moiety within the protein employed in the semisynthesis. Thus, the expression of recombinant proteins comprised of a disulfide-bond-containing, glycosylated extracellular region, as well as a functional intein, is challenging and, to the best of our knowledge, has not been reported.

Our interests are focused on the Eph single-pass membrane proteins, which are the largest family of receptor tyrosine kinases (RTKs). Ten membrane-attached ligands

Reprint requests to: Dimitar B. Nikolov, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, USA; e-mail: nikolovd@mskcc.org; fax: (212) 717-3135.

Abbreviations: EPL, expressed protein ligation; Eph, Eph receptor; GyrA, GyrA Intein, ECD, extracellular domain; MESNA, 2-mercaptoethanesulfonic acid.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.035659.108>.

called ephrins bind the 16 different Eph receptors and initiate unique bidirectional signaling cascades at cell–cell contact sites. The Eph receptors and ephrins are both divided into two subclasses, A and B, based on binding-partner preferences and sequence similarity. They trigger cellular responses to cues in the environment in both developing and mature tissues, regulating key processes including cell migration, tissue patterning, and axonal guidance (Wilkinson 2000; Pasquale 2005; Himanen et al. 2007). They also play critical roles in angiogenesis (Kullander and Klein 2002; Brantley-Sieders and Chen 2004) and have been implicated in angiogenesis-dependent diseases such as cancer (Pasquale 2005) and retinopathy. Both Eph receptor subclasses are expressed in many tumor types, where they regulate metastasis, tumor growth, and tumor-induced angiogenesis. It is therefore likely that blocking either the kinase activity of the Eph receptors or their interactions with the ephrins could potentially provide novel antitumor strategies (Levitzki 1999; Bange et al. 2001).

The Eph receptors have a conserved domain structure, including an N-terminal ephrin-binding domain and an adjacent cysteine-rich region followed by two fibronectin type III repeats. A membrane-spanning segment separates the ectodomain from the intracellular region, which consists of a juxtamembrane segment, a tyrosine kinase domain, and a SAM domain, often followed by a PDZ-binding motif. Recent structural studies have elucidated the molecular details of the initial Eph–ephrin recognition and binding (Himanen et al. 2001). However, these studies were limited to the minimal interaction domains, while evidence suggests that other regions of the receptor and the ligand are also important (Hock et al. 1998; Lackmann et al. 1998; Stapleton et al. 1999; Thanos et al. 1999). Most importantly, to date there is no structural information on how extracellular ligand binding results in

the activation of intracellular kinase domain in Eph receptors or, for that matter, in any other receptor kinase. Hence, intensive efforts are directed toward studying the structures of full-length, transmembrane receptor kinases and their activation by their ligands. Toward this goal, we report here the utilization of a semisynthetic, EPL-based approach to chemically join recombinant Eph ectodomains secreted from human cell lines with synthetic peptides (Fig. 1A). To our knowledge, this is the first reported chemical ligation of a multidomain, extracellular protein to a synthetic molecule. The ability to selectively and specifically modify the large glycosylated and disulfide-bonded Eph ectodomains for various in vitro studies paves the way toward understanding the fundamental steps in receptor tyrosine kinase activation and signaling at the molecular level. In addition, this work represents a critical first step toward the goal of preparing the full-length Eph receptor using semisynthesis.

Results and Discussion

Our semisynthetic strategy required the generation of an α -thioester derivative of the recombinant Eph extracellular domain (ECD), which can then be ligated to a synthetic peptide with a cysteine at its N terminus (Fig. 1A). Mammalian HEK293 cells and the baculovirus/insect cells were used for the recombinant protein expression of the complete secreted Eph ECD, corresponding approximately to residues 1–530, fused to an intein and a purification tag. C-terminal intein fusions can be used to directly introduce α -thioester groups into recombinant proteins through a thiolysis cleavage reaction (Muir et al. 1998; Severinov and Muir 1998). Accordingly, we decided to generate a “sandwich” fusion protein in which either the *Mycobacterium xenopi*

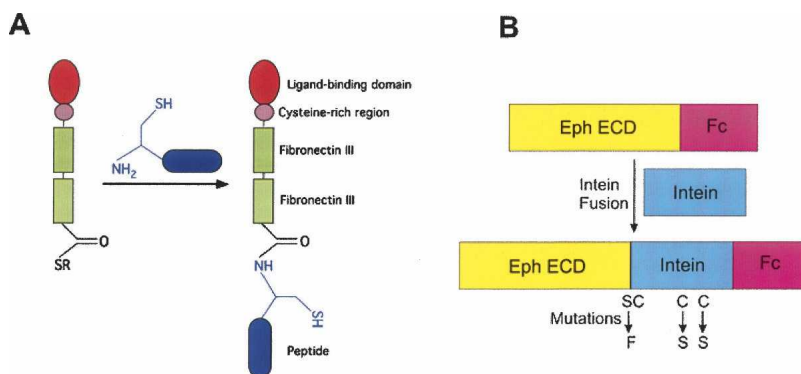


Figure 1. Expressed protein ligation. (A) Intein-mediated thiolysis generates a thioester group at the C terminus of the complete extracellular region of the Eph receptor. The cysteine residue at the N terminus of a synthetic peptide then reacts with the thioester group to generate the semisynthetic product. (B) Schematic representation of the fusion protein construct used for the expression. The fusion protein was generated by inserting the intein between the Eph ECD and the Fc purification tag. To increase the thiol cleavage efficiency, a Ser \rightarrow Phe and two Cys \rightarrow Ser mutations in the GyrA intein were created as described in the text.

(*Mxe*) GyrA intein or the *Saccharomyces cerevisiae* (*Sce*) VMA intein was inserted between the EphB4 ECD sequence and an Fc purification tag. Previous studies of the expression of extracellular proteins have shown that the Fc tag, in addition to facilitating the purification, increases the solubility of the recombinant protein (Pabbisetty et al. 2007). This fusion strategy leaves a serine residue next to the cysteine at the N terminus of the intein (Fig. 1B). The expression levels and solubility of the triple fusion proteins were evaluated in transiently transfected HEK293 cells. The *Sce* VMA intein fusion expressed poorly (data not shown); hence stably transfected cells were generated using only the GyrA-intein-containing construct (MW = ~110 kDa). The fusion protein was purified by affinity chromatography on Protein-A Sepharose (Fig. 2A). The intein cleavage was then induced using different concentrations of the sodium salt of 2-mercaptoethanesulfonic acid (MESNA). Protein-A Sepharose beads were used again to separate the intein-Fc fragment from the cleaved EphB4 ECD α -thioester fragment. The identities of all protein products were confirmed by N-terminal sequencing, Western blot, and tryptic mass-spectrometry analysis.

The HEK293-expressed protein preparations contained a large number of impurities (serum-derived immunoglobulins) from the mammalian cell growth media, and the thiolysis efficiency of the constructs was poor (Fig. 2B, right panel). We, therefore, decided to express the three-part fusion protein in the baculovirus/insect-cell system, which can utilize serum-free media for cell growth. A small-scale expression screen was conducted in both Sf9 and Hi5 insect cells for the secreted EphB4ECD-GyrA-Fc using Protein-A Sepharose for detection. Both the expression levels (~10–12 mg/L) and the purity of the protein

were higher when Hi5 cells and serum-free media were used for expression (Fig. 2A), and, consequently, all future experiments were performed using the baculovirus/Hi5 system. The insect-cell-derived protein was subjected to intein-mediated cleavage induced with various concentrations of MESNA. Although at this stage the cleavage was free of all other impurities, the overall efficiency of Eph ECD α -thioester formation was still poor (Fig. 2B, left panel).

For optimization of the thiol-induced cleavage, we rationalized that mutating the two nonactive site cysteine residues (Cys79 and Cys114) to serines in the *Mxe* GyrA intein might prevent any formation of nonnative disulfide bonds between these cysteines and the cysteines in the Eph ECD. This could potentially improve the fraction of correctly folded protein and thereby increase the cleavage efficiency. It has also been shown that the (-1) residue at the N-terminal splice junction (just before the intein N-terminal cysteine) plays an important role in the efficiency of thiol-induced cleavage (Southworth et al. 1999). For the *Mxe* GyrA intein, serine, proline, glutamic acid, and aspartic acid are often unfavorable at this location, while other residues such as methionine and phenylalanine usually work better: These “rules” are by no means hard and depend greatly on the nature of the protein fused to the intein (Muralidharan and Muir 2006). We therefore mutated the serine residue at the C terminus of the EphB4 ECD to phenylalanine. The EphB4ECD-GyrA-Fc protein, with the two Cys \rightarrow Ser and the Ser \rightarrow Phe substitutions, was expressed using the baculovirus system. The efficiency of the thiol-induced cleavage using this new construct was much higher, generating α -thioester protein fragments in sufficient purity for subsequent protein chemistry (>95%) (Fig. 3A). These

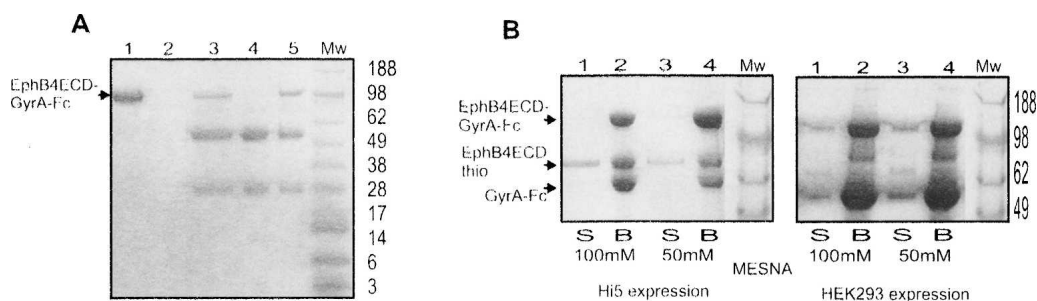


Figure 2. Generation of Eph-ECD α -thioester. (A) Expression of the EphB4ECD-GyrA-Fc fusion protein. The protein was purified from the expression media by affinity chromatography using Protein-A Sepharose beads and resolved on SDS-PAGE as shown. (Mw) Molecular weight marker (in kilodaltons). (Lanes 1,3,5) Expression of the fusion protein in Hi5, Sf9, and HEK293 cells, respectively; (lanes 2,4) control (no transfection) Hi5 and Sf9 cells. The *top* band (MW ~ 110 kDa) is the fusion protein. The two *lower* bands are impurities derived from the serum-rich cell growth media used in the HEK293 and Sf9 cell systems. (B) Generation of the EphB4 ECD α -thioester. The MESNA-induced cleavage of the EphB4ECD-GyrA-Fc fusion protein purified from Hi5 (*left* panel) and HEK293 (*right*) cells is resolved on SDS-PAGE. (Mw) Molecular weight marker (in kilodaltons). (Lanes 1,2) Cleavage with 100 mM MESNA; (lanes 3,4) cleavage with 50 mM MESNA. The uncleaved fusion protein and the cleaved intein-Fc fragment bind to the Protein-A Sepharose beads (B) as observed in lanes 2 and 4. The EphB4 ECD α -thioester fragment (EphB4ECDthio, MW ~ 62 kDa) stays in the supernatant (S) as observed in lanes 1 and 3.

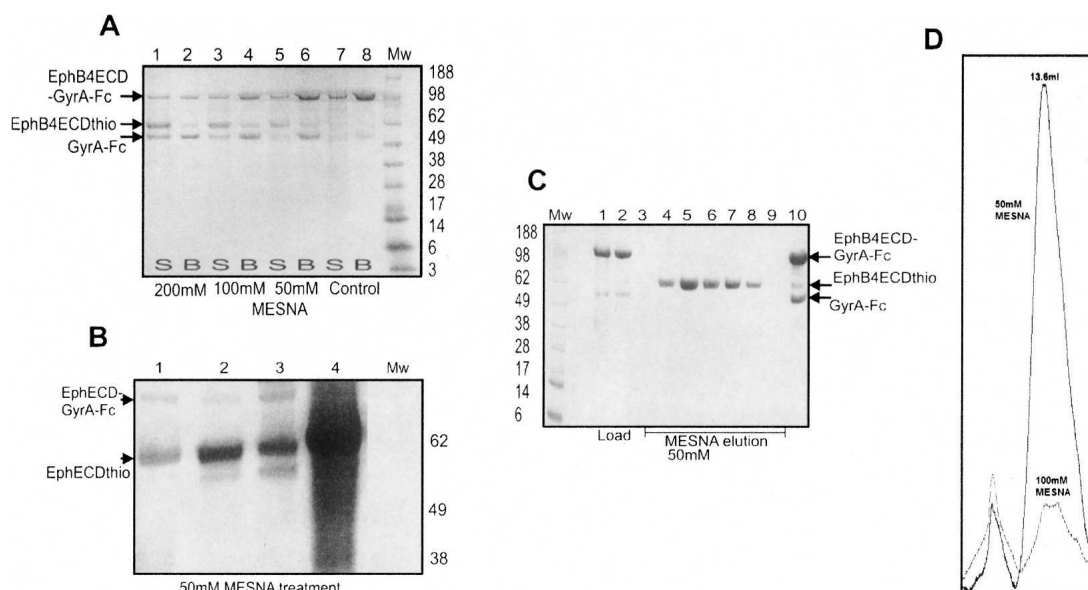


Figure 3. Optimization of the Eph-ECD-thioester yield. (A) α -Thioester generation in the ($2\times$ Cys \rightarrow Ser/Ser \rightarrow Phe) EphB4ECD-GyrA-Fc mutant. SDS-PAGE after the treatment of fusion protein mutant with increasing concentration of MESNA as described in the text. (Mw) Molecular weight marker (in kilodaltons). (Lanes 7,8) Control with no MESNA added. The uncleaved fusion protein and the cleaved intein-Fc fragment bind to the Protein-A Sepharose beads (B) as shown in lanes 2,4, and 6. The amount of EphB4 ECD α -thioester fragment (EphB4ECDthio, MW \sim 62kDa) generated in the supernatant (S) increases with increasing concentration of MESNA as shown in lanes 1,3, and 5. (B) Effect of mutations at the C terminus of Eph ECD on the thiol cleavage efficiency. SDS-PAGE showing the generation of the Eph ECD α -thioester from the different C-terminal mutants with 50 mM MESNA treatment. (Mw) Molecular weight marker (in kilodaltons). (Lanes 1–3) EphB4ECD α -thioester with glycine, phenylalanine, and alanine, respectively, at the C terminus. (Lane 4) EphA3 ECD α -thioester with phenylalanine at the C terminus. (C) Large-scale purification of EphB4 ECD with an active C-terminal α -thioester. MESNA induced on-column (Protein-A Sepharose column) cleavage of the mutated ($2\times$ Cys \rightarrow Ser/Ser \rightarrow Phe) EphB4ECD-GyrA-Fc fusion protein. (Mw) Molecular weight marker (in kilodaltons). (Lanes 1,2) Loaded fusion protein; (lanes 3–9) fractions containing the EphB4 ECD α -thioester eluted after 50 mM MESNA treatment; (lane 10) the uncleaved fusion protein and GyrA-Fc fragment remains bound to the Protein-A Sepharose column after cleavage. (D) Gel-filtration chromatography elution profile of the EphA3 ECD obtained after the cleavage of EphA3ECD-GyrA-Fc with 50 mM (solid line) and 100 mM (dotted line) MESNA. The protein is properly folded at the 50 mM thiol concentration, but becomes largely misfolded and aggregated as the concentration of MESNA is increased.

mutagenesis experiments furthermore document that the internal cysteine residues on the GyrA intein are not crucial for the splicing reaction. We also evaluated substitutions of the serine to alanine and glycine, but neither of these improved the cleavage efficiency beyond that obtained with the phenylalanine mutation (Fig. 3B).

Large-scale purification of the fusion protein was performed via affinity chromatography on a Protein-A Sepharose column. The intein-mediated cleavage of the column-bound protein was induced by incubating with MESNA (Fig. 3C), and the eluted Eph ECD- α -thioester protein was further purified on a gel-filtration column (Fig. 3D). It remains properly folded as judged by the fact that it shows a gel-filtration profile similar to that of the Eph ECD obtained after cleavage of the Fc tag with thrombin. The pooled fractions corresponding to the peak at \sim 13.0 mL migrate as a single band on SDS-PAGE.

Next, we sought to establish whether this intein fusion strategy also works for the generation of A-class Eph

receptor α -thioesters. Hence, a fusion protein comprised of EphA3 ECD with a phenylalanine mutation at its C terminus, a double Cys \rightarrow Ser mutant GyrA intein, as well as a Fc tag was constructed and expressed in the baculovirus/Hi5-cell system. The expression levels for the EphA3ECD-GyrA-Fc protein were higher than those for the EphB4 construct, as were the levels of the EphA3 ECD α -thioester fragment generated after thiol treatment (Fig. 3B).

As previously noted, the ECD region of Eph receptors is rich in disulfide bonds. This presents a potential problem for EPL reactions, which require the presence of mild reducing agents, such as MESNA. Accordingly, the concentration of the thiol agent had to be adjusted so that it did not reduce the cysteines in the ECD, while, at the same time, was high enough to facilitate the cleavage and the ligation reaction. As outlined in Materials and Methods, various reducing agents at different concentrations were evaluated for their ability to generate the Eph ECD α -thioester without reducing the disulfide bonds.

Gel-filtration chromatography (Fig. 3D) confirms that the Eph ECD obtained after treatment with 50 mM MESNA is properly folded and does not aggregate, whereas with increasing concentrations of MESNA to 100 mM and beyond, protein misfolding/aggregation becomes prominent. Hence, MESNA at a concentration of 50 mM was used in all further experiments; this provided sufficient cleavage efficiency while retaining the ligand binding activity of the protein.

The α -thioester-containing Eph ECDs were ligated to different peptides, each with an N-terminal cysteine resi-

due and a FLAG tag sequence for readout of the ligated product on Western blots. In each case, the (α -thioester) protein was incubated overnight at 4°C with the FLAG-tag peptides to generate the corresponding C-terminally modified Eph receptors. MESNA was included in these reactions to maintain the active α -thioester (principally by suppressing hydrolysis), while minimizing the disulfide-bond reduction within the Eph ECD fragment as judged by ligand binding. The ligation reactions were evaluated on Western blots (Fig. 4A). Different modifications, such as histidine and other tags (Fig. 4B), could be introduced via

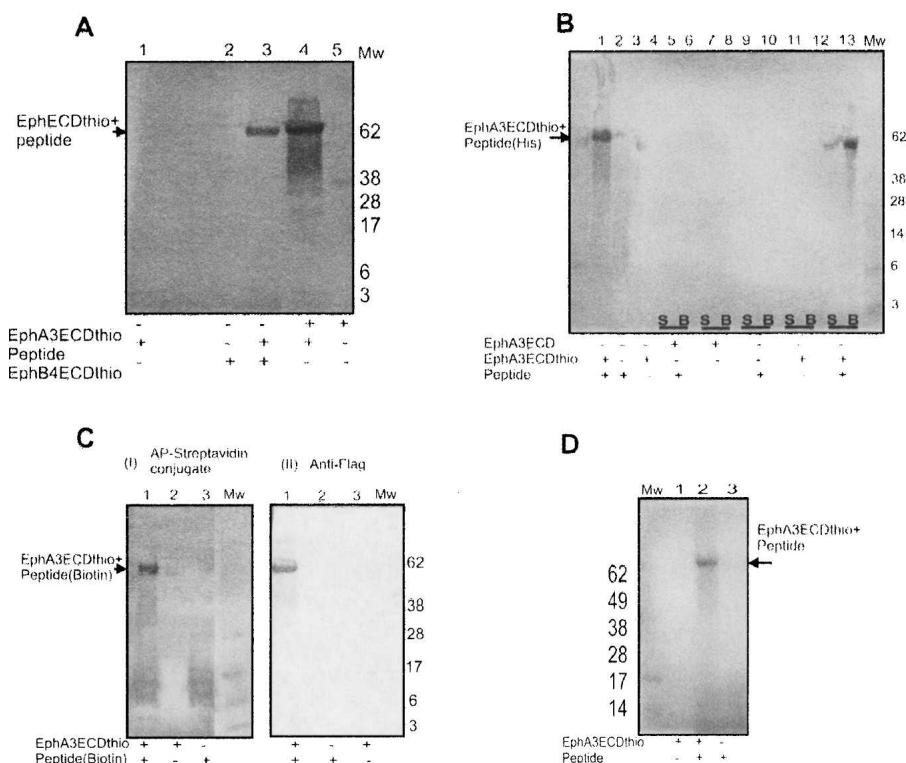


Figure 4. Ligation of the Eph ECD α -thioester with synthetic peptides. (A) The reactive α -thioester generated at the C terminus of the EphB4/A3 ECD was ligated to a peptide with FLAG-tag sequence and an N-terminal cysteine residue (CGGGDYKDDDDK). The ligation reaction was resolved on SDS-PAGE and visualized using anti-FLAG antibody on a Western blot as shown. (Mw) Molecular weight marker (in kilodaltons). The blot shows a signal for the ligated products but not for the controls, i.e., receptor thioesters before the ligation, and the small peptide by itself. (B) Modification of Eph ECD with Histidine-tag. EphA3 ECD α -thioester was ligated with a FLAG/His-tagged peptide (CGGGDYKDDDDKHHHHHH, MW = 2.1 kDa). The ligated product binds to the Ni^{2+} chelating beads (B) and was resolved on an SDS-PAGE. The results were visualized with an anti-FLAG antibody as shown on the Western blot. (Mw) Molecular weight marker (in kilodaltons). (Lane 1) EphA3 ECD α -thioester mixed with peptide. The ligated product is confirmed by signal on the blot. Lane 2 (peptide by itself) and lane 3 (EphA3 ECD α -thioester by itself) show no signal on blot. (Lanes 4–11) Controls (lacking different ligation components) showing supernatant (S) and beads (B) after binding to the Ni^{2+} chelating beads. Only EphA3 ECD α -thioester mixed with the peptide (lanes 12,13) gives a ligated product that binds to the beads and shows a signal on the blot. (C) Biotinylation of Eph ECD with hydrophilic peptide. EphA3 ECD α -thioester was ligated with FLAG-tag peptide that is biotinylated at its C terminus (CGGGDYKDDDD-K-Biotin). The ligated product was resolved on SDS-PAGE gel and analyzed with AP-conjugated Streptavidin (I) or anti-FLAG antibody (II) in the Western blots as shown. (Mw) Molecular weight marker (in kilodaltons). The ligated product is confirmed by signal on the blot with both methods of staining while controls give no signal. (D) Biotinylation of Eph ECD with hydrophobic peptide. EphA3 ECD α -thioester was ligated with a hydrophobic peptide, derived from the TM Eph region, that is biotinylated at its C terminus (CLALIAGTAVK-Biotin). The ligated product was resolved on SDS-PAGE and analyzed with AP-conjugated Streptavidin on Western blot. (Mw) Molecular weight marker (in kilodaltons). (Lane 2) The ligated product is confirmed by signal on the blot; (lanes 1,3) controls lacking either the peptide or the Eph thioester.

the peptides. The efficiency of this ligation reaction is ~99%, as documented in Figure 5A. In a similar fashion we site-specifically attached a single biotin at the C terminus of the Eph ECD both via a hydrophilic and a hydrophobic peptide as shown in Figure 4C,D, respectively. The short hydrophobic peptide (C-L-A-L-I-A-G-T-A-V-K-Biotin) was based on the transmembrane Eph receptor sequence and was used to document that the described method allows for the ligation of transmembrane-like sequences to the Eph ectodomains in the presence of appropriate detergent.

Eph-ephrin complex formation takes place via an extensive interaction interface. The glycosylation and disulfide bonds on the extracellular domains of both receptor and ligand are essential for their correct folding and the formation of their complexes (Himanen et al. 2001; Kullander and Klein 2002). The most critical issue in the

semisynthesis approach is to ensure that the proteins maintain their correct fold and native biological activity after the thiol-induced cleavage and the ligation reaction. Hence several *in vitro* binding assays, including pulldown, tryptophan fluorescence titration, and gel filtration, were performed to confirm specific Eph receptor binding to respective ephrin ligands before and after ligating the receptors to synthetic peptides. First EphB4 and EphA3 ECDs, obtained following thiol cleavage and gel filtration, were incubated with their respective Fc-tagged ligands, ephrin-B2 and ephrin-A5, and the complexes pulled down with Protein-A Sepharose beads (Himanen et al. 2004). The recombinant EphB4 and EphA3 ectodomains containing the reactive α -thioester are biologically active and able to form specific complexes with their respective ligands, ephrin-B2 and ephrin-A5. Similarly, the semisynthetic Eph receptor ectodomains ligated to synthetic peptides form specific

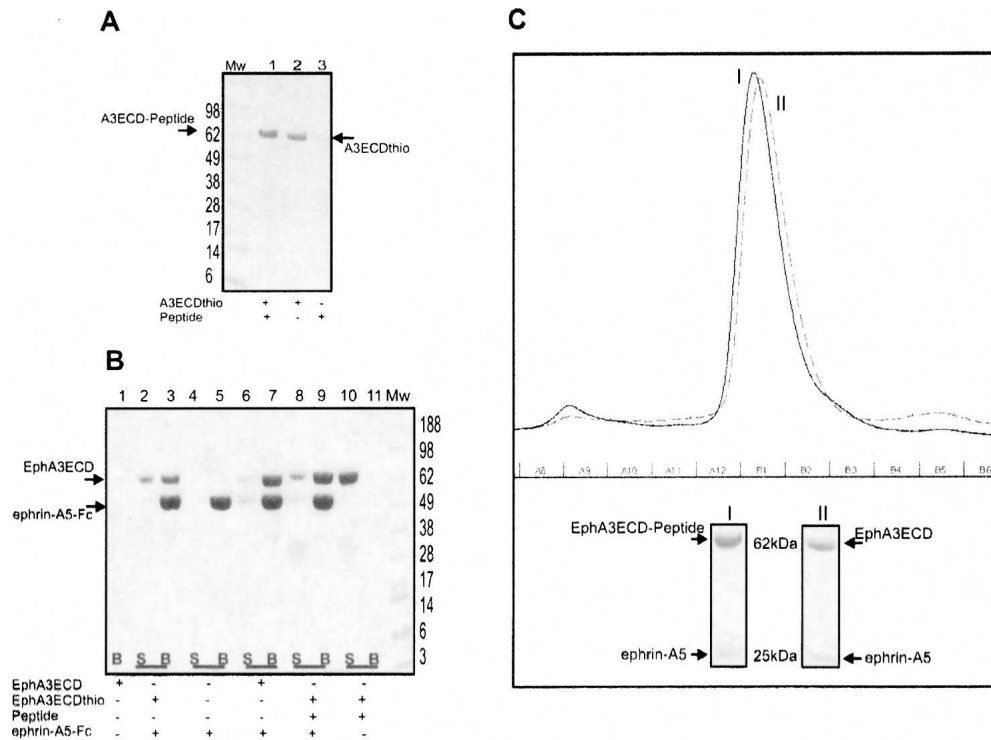


Figure 5. The peptide ligated Eph ECD retains full biological activity. (A) The ligation efficiency is ~99%. The EphA3 ECD α -thioester ligated with the Histidine-tagged peptide, MW = 2.1 kDa (described above) is resolved on Coomassie-stained SDS-PAGE. (Lane 1) The ligated product is confirmed by the increase in MW corresponding to that of the peptide used for ligation; (lane 2) the unligated EphA3 ECD α -thioester; (lane 3) the peptide by itself is too small and is not visible on the gel. (B) The Eph ECD retains its ligand-binding activity after thiol treatment and ligation. The EphA3 ECD α -thioester obtained after thiolysis and the peptide-ligated product were mixed with the Fc-tagged ephrin-A5 ligand. The Coomassie-stained SDS-PAGE shows that the recombinant ectodomain of EphA3, its α -thioester derivative, and the peptide-ligated product form a specific complex with the ephrin-A5 ligand, since they are pulled down by Protein-A Sepharose beads (B; lanes 3,7,9). Without the addition of the ephrin-A5-Fc, the peptide-ligated product of EphA3 ECD stays exclusively in the supernatant (S; lane 10). (Mw) Molecular weight marker (in kilodaltons). (C) The elution profile of the EphA3ECD/EphA3ECD-peptide in complex with the ligand ephrin-A5 resolved on a Superdex-200 10/30 gel filtration column. (I) The major peak corresponding to the ligated EphA3-ECD-Histidine-tag peptide and ephrin-A5 complex. (II) The major peak corresponding to the recombinant unligated EphA3 ECD and ephrin-A5 complex. Fraction B1, corresponding to the proteins in the major peaks of I and II, respectively, is resolved on Coomassie-stained SDS-PAGE gel. The proteins are labeled accordingly.

complexes with their respective ligands. Figure 5B, for example, shows that the peptide-ligated EphA3 ECD forms a specific complex with ephrin-A5. The results from the pulldown assay were confirmed using gel-filtration chromatography (Fig. 5C), documenting that the peptide-ligated EphA3 ECD forms a stable complex with ephrin-A5 in a fashion very similar to that of the unligated EphA3 ECD. Finally, fluorescence titration spectroscopy was used to estimate the ligand binding affinity of the EphA3 constructs. Ephrin-A5 bound the recombinant unligated EphA3 ECD, with essentially the same dissociation constant (K_d of ~ 100 nM) as the His-peptide ligated EphA3 ECD (K_d of ~ 90 nM; data not shown).

The findings described here document the extension of the expressed protein ligation technique for the site-specific modification of multidomain, disulfide-bond-containing, glycosylated extracellular regions of cell-surface receptors. Specifically, we generated functional EphECD-intein-Fc fusion proteins, using both A- and B-class Ephs that are secreted in large amounts in the growth media of the eukaryotic cells used for recombinant protein expression. Although the ECD is normally biologically active under oxidizing conditions, while the intein - in a reducing environment, we show that appropriate conditions can be identified that retain the activity of both. We also document that the two internal cysteine residues on the GyrA intein are not necessary for the splicing reaction. Using a thiol-induced reactive α -thioester, we are able to site-specifically modify the Eph ECD C terminus with different synthetic peptides. Neither the thiolysis nor the subsequent ligation reactions affect the ability of the Eph ECDs to specifically bind to their respective ephrin ligands. While our future efforts are predominantly geared toward extending the semi-synthetic protocols for the production of entire transmembrane Eph receptors, the methods developed herein will be applicable to the semisynthesis of a variety of other single-pass transmembrane proteins. Furthermore, the highly specific labeling of large transmembrane receptors with biophysical and biochemical probes could provide invaluable tools to study the molecular mechanisms that regulate the biological activity of these molecules and the signaling events that they initiate.

Materials and Methods

Construct design for protein expression in HEK293 cells

The extracellular domain of murine EphB4 (residues Glu17–Gln537), fused to the human IgG1 hinge and Fc regions (EphB4ECD-Fc), was cloned into a modified pcDNA3.1 vector (Invitrogen) and was constitutively expressed in a HEK293 (human embryonic kidney) cell line using the CD5 signal sequence. A thrombin cleavage site was introduced at the C-terminal site of the gene of interest. Mutant versions (C-terminal

Asn \rightarrow Ala substitution) of the *Mxe* GyrA intein (198 amino acids) and the *Sce* VMA intein (455 amino acids) are available in commercial vectors, pTXB1 and pTYB1 (New England Biolabs), respectively. These inteins, which promote only the first step of protein splicing (Xu and Perler 1996), were PCR amplified and, using a BamHI restriction site, were fused in frame to the C terminus of EphB4 ECD between the ECD and the Fc purification tag. The ectodomain of human ephrin-A5 (residues 1–228) was cloned N-terminally to a thrombin cleavage site and the Fc region of IgG1 in the same modified pcDNA3.1 vector and also expressed in HEK293 cells.

Construct design for protein expression in insect cells

The EphB4ECD-GyrA construct from the HEK293 expression vector was PCR amplified and cloned using the EcoRI and EagI restriction sites of a modified pAcGP67-B Baculovirus vector (BD Pharmingen). The EphA3ECD was cloned into the same vector using the BamHI and NotI restriction sites. The GyrA intein was PCR amplified as before and, using the NotI and EagI restriction sites, fused in frame to the C terminus of the EphA3 ECD. The vector is under the control of the heterologous GP67 signal sequence and contains a thrombin cleavage site just before the C-terminal Fc purification tag. The constructs were sequenced, and the recombinant baculovirus vector was cotransfected with BaculoGold DNA (Pharmingen) in Sf9 cells. After three rounds of viral amplification, passage-3 virus was used to infect Hi5 cells. Infected cells were grown at 27°C/100 rpm and harvested after 72 h.

Protein purification and generation of EphB4/A3 ECD α -thioester

The secreted fusion proteins, containing Fc purification tags, were extracted from the medium by affinity chromatography using Fast Flow Protein-A Sepharose (Amersham). The purified EphB4/A3ECD-GyrA-Fc protein was dialyzed into thiol cleavage buffer containing 100 mM HEPES (pH 7.5), 300 mM NaCl, and 1 mM EDTA (a number of different buffers and pH conditions were tried). Different thiol reagents, such as DTT, thiophenol, and MESNA, were evaluated at different concentrations, incubation periods, and temperature conditions. Finally, overnight incubation at 4°C was selected in the presence of 50 mM MESNA. Protein-A Sepharose beads washed with buffer were added to bind the intein-Fc fragment and the uncleaved protein, while the cleaved EphB4 ECD α -thioester fragment stays in the supernatant. The protein was further purified by gel-filtration chromatography using the Superdex-200 10/30 column (GE HealthCare) preequilibrated with a buffer containing 20 mM HEPES (pH 7.2), 150 mM NaCl, and 2.5 mM $MgCl_2$. The EphB4 ECD obtained after cleavage was identified on Western blots using goat anti-EphB4 primary antibody (SIGMA) and alkaline phosphatase (AP) conjugated anti-goat IgG (H+L) secondary antibody. The SDS-PAGE gel was electroblotted onto a polyvinylidene (PVDF) membrane (BioRad), blocked with 2% fat free milk in TBST (0.5% Tween 20, 1.5 M NaCl in 100 mM Tris-HCl at pH 8.0), and incubated overnight with the primary antibody. After rinsing with TBST, the membrane was incubated for 1 h with the secondary antibody. The AP was visualized with NBT/BCIP substrate tablets (Roche). Site-directed mutagenesis, aimed to improve the cleavage efficiency, was performed with the QuikChange

Mutagenesis Kit (Stratagene) following the manufacturer's instructions. The Eph ECD mutants were expressed and purified using the baculovirus/insect-cell system as described above.

Peptide ligation

The EphB4/A3 ECD containing an active C-terminal α -thioester obtained after thiol cleavage was used for ligation to commercially synthesized peptides. The peptides (synthesized by Bio-peptide Inc.) had an N-terminal cysteine residue and different modification at the C terminus (e.g., were biotinylated or His-tagged). They also contained a FLAG-tag sequence (Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys) for an easy readout of the ligation reaction. The purified peptides were dissolved in thiol cleavage buffer as above. For the hydrophobic peptide, 0.1% Triton X-100 was added to the thiol cleavage buffer. The ligation reactions were performed by incubating overnight at 4°C an excess of the peptide with the freshly obtained Eph ECD α -thioester in ligation buffer containing 100 mM HEPES (pH 8.4) and 10 mM MESNA. The ligation of FLAG-tagged peptides was confirmed by Western blot using a rabbit anti-FLAG (Sigma) polyclonal primary antibody and an AP-conjugated anti-rabbit IgG (H+L) (Promega) secondary antibody. The ligation product containing biotin was confirmed by Western blot using AP conjugated streptavidin (Zymed). The ligation of His-tag-containing peptides was confirmed by a pulldown assay of the ligated product with Ni²⁺-chelating Sepharose beads (Amersham).

Pulldown ligand binding assay

Eph ECD containing active C-terminal α -thioester

Approximately 10 μ g of EphA3 ECD, obtained following thiol cleavage, were incubated with Fc-tagged ephrin-A5 for an hour at 4°C, in 500 μ L of binding buffer containing 20 mM HEPES (pH 8.2), 150 mM KCl, and 2 mM MgCl₂. Protein-A Sepharose Fast Flow beads (Amersham) were washed with the binding buffer, added to the reaction mixture, and incubated at 4°C for 2 h. The beads were harvested by centrifugation and washed twice with 500 μ L of the binding buffer, and the bound proteins were separated on a SDS-PAGE (Fig. 5B) (Himanen et al. 2004). A similar protocol was used to document the retention of biological activity (binding to its corresponding ligand, ephrin-B2) for the EphB4 ECD obtained following thiol cleavage (data not shown).

Eph ECD ligated to peptides

Approximately 10 μ g of the 6 \times His-peptide-ligated EphA3 ECD, obtained after the ligation reaction with the synthetic peptide, were incubated with Fc-tagged ephrin-A5, and the pulldown Protein-A Sepharose experiment was performed as described above. The bead-bound proteins were resolved on SDS-PAGE (Fig. 5B), documenting that the modified recombinant EphA3 ectodomain is biologically active and able to form a specific complex with its ligand, ephrin-A5.

Gel-filtration and tryptophan fluorescence titration ligand binding assays

The purified EphA3 ECD obtained after thrombin cleavage and the EphA3 ECD α -thioester obtained after treatment with 50 mM MESNA and ligation to the His-tagged peptide were mixed with the purified ephrin-A5 ligand in a 1:1 molar ratio. The mixture

was incubated for an hour at 4°C, in 500 μ L of binding buffer containing 20 mM HEPES (pH 8.2), 150 mM KCl, and 2 mM MgCl₂. The complex was then loaded onto a Superdex-200 10/30 gel-filtration column (GE HealthCare) preequilibrated with a buffer containing 20 mM HEPES (pH 7.2), 150 mM NaCl. Tryptophan fluorescence titration spectroscopy experiments using the purified EphA3 ECD, ephrin-A5, and the peptide ligated EphA3 ECD were performed as described (Himanen et al. 2004).

Acknowledgment

This work was supported by the National Institutes of Health (grant NS38486 to D.B.N.).

References

- Bange, J., Zwick, E., and Ullrich, A. 2001. Molecular targets for breast cancer therapy and prevention. *Nat. Med.* **7**: 548–552.
- Brantley-Sieders, D.M. and Chen, J. 2004. Eph receptor tyrosine kinases in angiogenesis: From development to disease. *Angiogenesis* **7**: 17–28.
- Himanen, J.P., Rajashankar, K.R., Lackmann, M., Cowan, C.A., Henkemeyer, M., and Nikolov, D.B. 2001. Crystal structure of an Eph receptor–ephrin complex. *Nature* **414**: 933–938.
- Himanen, J.P., Chumley, M.J., Lackmann, M., Li, C., Barton, W.A., Jeffrey, P.D., Vearing, C., Geleick, D., Feldheim, D.A., Boyd, A.W., et al. 2004. Repelling class discrimination: Ephrin-A5 binds to and activates EphB2 receptor signaling. *Nat. Neurosci.* **7**: 501–509.
- Himanen, J.P., Saha, N., and Nikolov, D.B. 2007. Cell–cell signaling via Eph receptors and ephrins. *Curr. Opin. Cell Biol.* **19**: 534–542.
- Hock, B., Bohme, B., Karn, T., Yamamoto, T., Kaibuchi, K., Holtrich, U., Holland, S., Pawson, T., Rubsamen-Waigmann, H., and Strebhardt, K. 1998. PDZ-domain-mediated interaction of the Eph-related receptor tyrosine kinase EphB3 and the ras-binding protein AF6 depends on the kinase activity of the receptor. *Proc. Natl. Acad. Sci.* **95**: 9779–9784.
- Kullander, K. and Klein, R. 2002. Mechanisms and functions of Eph and ephrin signalling. *Nat. Rev. Mol. Cell Biol.* **3**: 475–486.
- Lackmann, M., Oates, A.C., Dottori, M., Smith, F.M., Do, C., Power, M., Kravets, L., and Boyd, A.W. 1998. Distinct subdomains of the EphA3 receptor mediate ligand binding and receptor dimerization. *J. Biol. Chem.* **273**: 20228–20237.
- Lesaigner, M.L., Lue, R.Y., Chen, G.Y., Zhu, Q., and Yao, S.Q. 2002. Intein-mediated biotinylation of proteins and its application in a protein microarray. *J. Am. Chem. Soc.* **124**: 8768–8769.
- Levitzi, A. 1999. Protein tyrosine kinase inhibitors as novel therapeutic agents. *Pharmacol. Ther.* **82**: 231–239.
- Muir, T.W. 2003. Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* **72**: 249–289.
- Muir, T.W., Sondhi, D., and Cole, P.A. 1998. Expressed protein ligation: A general method for protein engineering. *Proc. Natl. Acad. Sci.* **95**: 6705–6710.
- Muralidharan, V. and Muir, T.W. 2006. Protein ligation: An enabling technology for the biophysical analysis of proteins. *Nat. Methods* **3**: 429–438.
- Pabbisetty, K.B., Yue, X., Li, C., Himanen, J.P., Zhou, R., Nikolov, D.B., and Hu, L. 2007. Kinetic analysis of the binding of monomeric and dimeric ephrins to Eph receptors: Correlation to function in a growth cone collapse assay. *Protein Sci.* **16**: 355–361.
- Pasquale, E.B. 2005. Eph receptor signalling casts a wide net on cell behaviour. *Nat. Rev. Mol. Cell Biol.* **6**: 462–475.
- Severinov, K. and Muir, T.W. 1998. Expressed protein ligation, a novel method for studying protein–protein interactions in transcription. *J. Biol. Chem.* **273**: 16205–16209.
- Southworth, M.W., Amaya, K., Evans, T.C., Xu, M.Q., and Perler, F.B. 1999. Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase A intein. *Biotechniques* **27**: 110–120.
- Stapleton, D., Balan, I., Pawson, T., and Sicheri, F. 1999. The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. *Nat. Struct. Biol.* **6**: 44–49.
- Thanos, C.D., Goodwill, K.E., and Bowie, J.U. 1999. Oligomeric structure of the human EphB2 receptor SAM domain. *Science* **283**: 833–836.
- Wilkinson, D.G. 2000. Eph receptors and ephrins: Regulators of guidance and assembly. *Int. Rev. Cytol.* **196**: 177–244.
- Xu, M.Q. and Perler, F.B. 1996. The mechanism of protein splicing and its modulation by mutation. *EMBO J.* **15**: 5146–5153.