Published in final edited form as:

FEBS J. 2009 August ; 276(16): 4395–4404. doi:10.1111/j.1742-4658.2009.07147.x.

Structural Recognition of an Optimized Substrate for the Ephrin family of Receptor Tyrosine Kinases

Tara L. Davis^{1,2}, John R. Walker¹, Abdellah Allali-Hassani¹, Sirlester A. Parker³, Benjamin E. Turk³, and Sirano Dhe-Paganon^{1,2}

¹Structural Genomics Consortium, University of Toronto, 101 College Street, Toronto, Ontario M5G 1L7, Canada.

²Department of Physiology, University of Toronto, 101 College Street, Toronto, Ontario M5G 1L7, Canada.

³Department of Pharmacology, Yale University School of Medicine, New Haven, CT 06520.

Summary

Ephrin receptor tyrosine kinase A3 (EphA3, EC 2.7.10.1) is a member of a unique branch of the kinome in which downstream signaling occurs in both ligand- and receptor- expressing cells. Consequently the ephrins and ephrin RTKs often mediate processes involving cell:cell contact, including cellular adhesion or repulsion, developmental remodeling, and neuronal mapping. The receptor is also frequently overexpressed in invasive cancers, including breast, small-cell lung and gastrointestinal cancers. However, little is known about direct substrates of EphA3 kinase and no chemical probes are available. Using a library approach, we found a short peptide sequence that is a good substrate for EphA3 and that is suitable for cocrystallization studies. Complex structures show multiple contacts between kinase and substrates, and in particular two residues undergo conformational changes and by mutation are found to be important for substrate binding and turnover. In addition, a difference in catalytic efficiency between EPH kinase family members is observed. These results provide insight into the mechanism of substrate binding to these developmentally integral enzymes.

Keywords

Ephrin kinase; receptor tyrosine kinase; x-ray crystallography; substrate recognition

Introduction

The Ephrin receptor class of receptor tyrosine kinases (EPH RTKs) is the largest subgroup of RTKs in the kinome, and encodes a wide range of biological activities. Many of these activities relate directly to cell:cell communication - including signaling involved in cell morphology and cell movement - and also effect cell proliferation, differentiation, and survival [1–3]. The EPH RTKs are uniquely suited to these types of signaling pathways due to the distinctive mode of interaction between the RTK and the ephrin ligand; cells expressing and presenting the ligand interact with neighboring cells expressing the transmembrane RTK, and this contact induces “bidirectional” signaling in both the ephrin-expressing and kinase-expressing cell types [2,4,5]. It follows that both the ephrin ligand and the EPH RTKs are attractive drug targets for diseases intimately connected with pathological cell contact, including many types of

cancers; tumorigenic growth, invasiveness, and angiogenic pathways are clearly and directly impacted by ephrin and EPH expression levels in tumor cells [3,6,7].

Of the sixteen EPH RTKs encoded by the human genome, EphA3 has emerged as a novel target for therapeutics aimed at cancer and leukemia. EPHA3 is involved in neural and retinal development in mammals, and was originally described as a determinant of retinotectal mapping [8–10]. Surprisingly, EPHA3 knockouts showed a clear heart phenotype, developing abnormal atria that led to high post-natal mortality [11]. The molecular basis for these findings has not been elucidated. Later work has shown overpopulation of EPHA3 mutations in colorectal, lung, liver, and kidney cancers [12–14], and in glioblastoma, melanoma, and rhabdomyosarcoma cell lines among others [15,16] suggesting that EphA3 kinase domain is an attractive candidate for drug development in these highly aggressive tumors. EphA3 (along with most of the EPH A-class RTKs) is a highly promiscuous receptor for ephrins, which allows for crosstalk between four of the five ephrin A type ligands in addition to ephrin B2 [3,7,17–19]. Since EphA3 is widely expressed in tissues from placental stages and throughout development, as are many of the ephrin ligands, it is important to find pharmacological strategies for studying EphA3 that are as specifically targeted towards this isoform.

Along these lines our lab has previously studied the auto-regulatory mechanism of the EphA3 kinase domain by determining a group of EphA3 structures in various states of activation [20]. In the current work a *de novo* peptide has been developed, showing a marked increase in affinity for EphA3 over peptides derived from auto-phosphorylation sites in the juxtamembrane region of EphA3. Two structures in complex with peptide rationalize the increase in affinity observed in solution. Two residues contributed by the kinase domain in the structure seem to explain the high affinity towards substrate, and mutational analysis confirms their importance in the kinase:substrate interaction. Finally, the selectivity of this peptide for EphA3 over other ephrin receptor kinases gives insight into substrate specificity for this biologically relevant class of receptor tyrosine kinases and provides a valuable tool for future research.

Results/Discussion

The juxtamembrane region of the cytosolic domain is a validated auto-phosphorylation site for Eph kinases and was initially targeted for co-crystallization efforts. The juxtamembrane EphA3 peptide, D⁵⁹⁸PHTYEDPTQ⁶⁰⁶ - where the numbers correspond to the residue numbers of the EphA3 receptor - is a substrate for the EphA3 kinase domain with catalytic efficiency of 200 min⁻¹mM⁻¹ ($K_m = 1 \pm 0.02$ mM; $k_{cat} = 199 \pm 9$ min⁻¹) [20]. Unfortunately, extensive attempts to crystallize EphA3 with this peptide were unsuccessful, perhaps due to poor affinity for the kinase domain. To screen for more suitable substrates, a positional scanning peptide approach was utilized that evaluates phosphorylation of a set of arrayed degenerate peptides having fixed amino acids at one of the five preceding, or four succeeding, positions relative to the phospho-acceptor tyrosine (described as positions -5 through +4 throughout the text). In addition to the 20 unmodified amino acids, the array also included peptides containing phosphothreonine or phosphotyrosine at each fixed position. The results of this screen indicated that EphA3 was largely unselective at positions upstream of the phosphorylation site with the exception of the -1 position, where the kinase selects primarily acidic residues (including phosphotyrosine and phosphothreonine) and asparagine, and is also tolerant of hydrophobic residues such as leucine and isoleucine (Figure 1 and Supplemental Table S1). The positions following the substrate tyrosine in general showed greater stringency, with clear preferences for tryptophan at the +4; aliphatic residues (including proline) at the +3; and acidic residues at the +1 position. Strikingly, the enzyme strongly preferred phosphotyrosine at the +2 position of the array, with other polar residues being selected to a much lesser extent.

Based on the combinatorial peptide array results, the following peptide was synthesized: KQWDNYEpYIW (hereafter referred to as EPHOPT), where pY at the position +2 to the substrate tyrosine denotes a phosphotyrosine incorporated into the peptide during synthesis. This peptide was tested under similar conditions to the original juxtamembrane substrate and showed a remarkable augmentation in regards to both turnover and binding affinity; catalytic efficiency increased over 200-fold, with a drop in K_m almost two orders of magnitude (from 1 mM to $18 \pm 4 \mu\text{M}$) and a 4-fold increase in k_{cat} (from 199 to 850 ± 44 per minute) (Table 2, peptide EPHOPT). Co-crystals of the EphA3 kinase domain with AMP-PNP and EPHOPT were obtained under identical conditions as that of unliganded EphA3, and a $1.7\text{\AA}$ dataset was collected. Statistics for data collection and processing are provided in Table 1.

The complex structure between EphA3 and EPHOPT shows clear density for most of the substrate peptide including mainchain atoms for the -4 through to the $+4$ position, but partial or no density for the sidechains of the amino terminal three residues and carboxy terminal tryptophan residue. Density for the substrate tyrosine and the phosphorylated tyrosine at position $+2$ is clear and unambiguous (Figure 2B). Overall, the structure of the kinase domain is found to be in the activated form, as described previously (for instance an AMPPNP bound structure, PDB code 2QO9); the juxtamembrane region is mostly disordered, concomitant with a greater degree of order found in the activation loop region (Figure 2A). The orientation of the Tyr742:Ser768 residue pair, described previously as a marker of EPH kinase activation [20], is in the non-clashing “active” rotamer position. As expected, most structural rearrangements to accommodate AMP-PNP binding are accomplished by the N-terminal lobe, especially the $\beta 1$ - $\beta 2$ loop (G loop) and αC regions. The crystallization of the complex between the EphA3 kinase domain and the EPHOPT peptide did not result in the full ordering of the activation loop (AL); instead, the N-terminal part of the AL was found ordered to residue Asp774, while the C-terminal part of the AL was ordered to residue Gly784. This represents an appearance in density of only one residue on either end of the AL over our most ordered structure to date (PDB 2QOC, representing a kinase domain without the juxtamembrane segment and bound to AMP-PNP) [20]. Perhaps this is due to the relatively short peptide that was used for crystallization, or to apparent crystal contacts that place a symmetry-related molecule relatively close to where the AL order ends.

Interactions between the EphA3 kinase domain and the EPHOPT substrate do not effect large conformational changes in the N- or C-terminal lobes of the kinase (Figure 2A). There is a slight movement in the αF - αG loop in the C-terminal lobe, which has the effect of moving the loop residues Met828-Gln831 $\sim 0.9\text{\AA}$ closer to the substrate. There are, however, some conspicuous differences in the AL loop residues beginning at Gly784 and continuing through to Trp790 in the EPHOPT complex structure (Figure 2C and Figure 3). These residues will be described in greater detail below. In addition, the availability of models representing low-activity (PDB 2QOQ), intermediate (2QOB, 2QO9, 2GSF), and high-activity (PDB 2QOC) conformations of EphA3 can be used with the current models in order to directly compare conformational changes induced by the ATP analog to the effects of substrate interaction.

Several residues undergo significant changes in the substrate bound complex. Arg745, for instance, is found in three discrete positions in all EphA3 structures. In the substrate complex Arg745 is found moved further in towards the activation loop which can be compared to previously determined active conformations which also move slightly towards the AL (Figure 2C and Figure 3A). The conformer found in the EphA3:EPHOPT complex is quite similar to that found in activated insulin receptor kinase (IRK), where the corresponding residue interacts with a phosphotyrosine in the AL of that protein [21]. This Arg745 flip is not simply a consequence of phosphorylation of the AL tyrosine pTry799; in the substrate bound complex it is found in a unique position even compared to other EphA3 structures where this tyrosine is phosphorylated based on LC-MS/MS analysis [20]. Arg823, located in the αF - αG loop region

(Figure 3A), is also found in a unique rotamer position in the EphA3:EPHOPT complex relative to all other EphA3 structures. In the substrate complex, Arg823 moves to coordinate both the backbone of Asn-1 (2.96Å) and Oδ1 in Asp-2 (2.87Å) (Figure 2C and Figure 3A). This residue, like Arg745, is conserved amongst almost all EPH RTK isoforms, except the pseudokinase EPHA6 and a substitution from Arg745 to Lys in EPHA1. Similarly, Glu827 moves in the substrate complex and coordinates the backbone N and O of Lys-5 (2.81, 2.65Å) and also supports orientation of Arg823. Finally, Asn830 coordinates Oε1 of Glu+1 (2.65Å) (Figure 2C) and this moves αG towards the substrate in an orientation unique to the EphA3:EPHOPT complex (Figure 3A).

However, the most striking residue movement in the EphA3:EPHOPT complex is Lys785, in the C-terminal region of the kinase activation loop (Figure 2C). Other structures have a random orientation or disorder at this position, but in the substrate complex this residue is clearly ordered, flipped out towards solvent, and nestled in between the Y0/E+1/pY+2 sequence of peptide (Figure 3A and 3B). Although not making direct electrostatic interactions with the phosphotyrosine moiety - which might have been predicted based on the complementary charge of the lysine - the structure implies that the function of Lys785 could be to lock the C-term AL into position relative to substrate sequences.

Based on the EPHOPT complex, a series of variant peptides were synthesized to probe the relevance of the +2 substrate position in affinity and turnover efficiency. In contrast to the data from the *in vitro* peptide screen, the effect of changing the substrate +2 phosphotyrosine residue to a phenylalanine (peptide OPTYF) results in only minor changes in K_m and k_{cat} ($30 \pm 5.7 \mu M$ and $421 \pm 30 \text{ min}^{-1}$; relative change of ~2-fold in each parameter) (Table 2). In order to rationalize this finding, EphA3 was co-crystallized with the OPTYF peptide; the structure is quite similar to the EPHOPT complex structure, with an RMSD of 0.17 Å over all Ca atoms and the majority of EphA3 residue sidechains in conformations as described previously. The substrate tyrosine and the phenylalanine aromatic sidechain at the +2 position are superimposable with the substrate tyrosine and phosphorylated tyrosine in the EPHOPT complex (Figure 3C). The subtle difference in phosphorylation efficiency between the two peptides might be explained by the conformations of a few key kinase residues in the OPTYF complex -including Arg712, Arg823, and the Glu827 - Asn830 region - which are in a low to intermediate activity conformation, and do not coordinate with substrate as they do in the EPHOPT complex (Figure 3). Additionally, the backbone atoms of the Trp-3 - Asp-2 region of OPTYF substrate have moved relative to the EPHOPT peptide and are no longer coordinated by EphA3 kinase; and the -4 residue is disordered.

The glutamate at substrate position +1 is pointing away from kinase residue Asn830, a drastic change from the coordination seen in the EPHOPT complex (Figure 3). Finally, the AMP-PNP molecule included in the co-crystal trials is disordered in the OPTYF structure. Although the structural changes are subtle, the phosphotyrosine at the +2 plays an important role in reordering the structure of the region of the C-terminal lobe that interacts with these substrates. In the absence of the phosphate group, the sidechain of substrate residue Glu+1 has reoriented to point towards kinase residue Lys785 (a 2Å movement) and the N-terminal part of the substrate has moved out of the kinase subsite delineated by residues Arg712 and the region including residues Arg823-Asn830.

The sidechain of Lys785 in the EphA3:OPTYF complex is also flipped out in the same distinctive way as in the EPHOPT complex, implying that this ordered movement is concomitant with substrate binding and is perhaps minimally sufficient for substrate coordination. This would also explain why, even though there are several rearrangements in the OPTYF complex that result in fewer interactions with the C-terminal lobe, there is still significant affinity of this peptide for EphA3. In line with these findings - and also in agreement

with the peptide array data – replacement of the phosphotyrosine with a lysine (peptide OPTYK) leads to a decrease in catalytic efficiency of one order of magnitude, mainly due to K_m effects ($107 \pm 15.6 \mu\text{M}$, a ~10-fold effect) (Table 2). This is presumably because a lysine at the +2 position of the substrate would be expected to clash strongly with Lys785 from the kinase domain, again suggesting that the concerted movement of Lys785 is directly related to substrate coordination.

To test the relative importance of the Asn830 and Lys785 interactions with substrate, EphA3 mutants were generated and their catalytic efficiencies tested. EphA3(N830A) showed more than one order of magnitude decrease in catalytic efficiency against the EPHOPT substrate, due largely to k_{cat} effects (k_{cat}/K_m $0.835 \mu\text{M min}^{-1}$, a 56-fold difference) (Table 2). EphA3(N830A) also showed a 5-fold weaker affinity for the OPTYF peptide than for EPHOPT peptide. Based on the structural data, this result is likely due to the fact that the EPHOPT sequence forms interactions with the second substrate binding subsite comprised of residues Arg712 and the region including residues Arg823-Asn830, while the OPTYF peptide does not; therefore the loss of the interaction with Asn830 would be more significant for the OPTYF peptide. In comparison, the EphA3(K785E) mutation negatively affected both K_m and k_{cat} by about an order of magnitude relative to wild-type enzyme ($188 \pm 64 \mu\text{M}$ and $34 \pm 0.7 \text{ min}^{-1}$). The catalytic efficiency for EphA3(K785E) against EPHOPT was almost negligible (260-fold decrease). In line with the identical orientation of Lys785 seen in the OPTYF structure, the catalytic efficiency for EphA3(K785E) against OPTYF was equally low (K_m $148 \pm 10 \mu\text{M}$; k_{cat} $23 \pm 8 \text{ min}^{-1}$, k_{cat}/K_m $0.155 \mu\text{M min}^{-1}$) (Table 2). Both kinase mutants were competent for auto-phosphorylation (four sites verified by LC-MS; data not shown), so it is unlikely that the dramatic decreases in catalytic efficiency seen were due to trivial misfolding of the mutant EphA3 kinase domain. Finally, both Asn830 and Lys785 are completely conserved across EPH isoforms (excepting the pseudo-kinases EPHA10 and EPHB6) (Figure 4), suggesting that these residues are involved more generally in both binding and effective catalysis of substrate in the EPH RTK family. In fact, all of the residues that interact directly with the EPHOPT substrate based on the EphA3 complex structure are conserved across both the EPHA and EPHB kinase classes. However, there are neighboring residues that are poorly conserved, including the AL residue at position 782 in EphA3 (Figure 4). Although density for this residue has not been observed in the structures of EPH kinases, the sidechain would likely be found near the phosphotyrosine at position +2 and is a good candidate for substrate recognition. This residue is variously an arginine in EphA3 and either a serine, threonine, glutamine in the EPHA isoforms or a serine-leucine or alanine-leucine insert in EPHB isoforms (Figure 4).

To test whether the EPHOPT peptide is specific for EphA3, a group of four additional EPH kinase domains, including EphA5, EphA7, EphB3, EphB4, and EphB2, was analyzed. We found that the EPHOPT peptide was mildly to strongly selective for EphA3, with catalytic efficiencies decreasing from 3 to 88 –fold for the other isoforms tested (EphA3>EphA5≈EphB3>>EphA7≈EphB4>>EphB2) (Supplemental Table S2). Utilizing array technology, the *in vitro* substrate specificity for EphA4 was recently published and can be summarized as {not R, H, K, P}-Y-[E/D]-[E/D]-[PILF] [22]. These results are similar to our EphA3 motif, and would indicate that EphA4 should be active against the EPHOPT substrate as well. The identity of the EphA3 residue Arg782 in EphA7, B4, and B2 kinases are all non-arginine, and indeed lower catalytic efficiencies for our substrate against those isoforms was found. However, why EphB3 was nearly as efficient as EphA5 and had nearly identical substrate affinity as EphA3 is presently unclear.

In summary, we have identified a substrate with low micromolar affinity for EphA3, a target of interest because of its isoform specific participation in cancer pathologies. Complex structures revealed a binding conformation in the catalytic cleft that is likely adopted in the recognition of physiologically relevant substrates and provides a molecular basis for our

observed peptide affinities and enzyme isoform specificities. These results will facilitate future studies focused on the rational design of peptide-like chemical probes.

Experimental Procedures

Cloning and Expression

The construct used for expression of EphA3 kinase domain has been described previously [20]. For site-directed mutagenesis, plasmids were subjected to QuickChange (Stratagene) mutagenesis using mutagenic primers spanning the altered codons. Resultant plasmids were transformed into BL21 Gold (DE₃) cells (Stratagene) for large scale protein expression. Cells were grown in supplemented Terrific Broth media at 37 °C to an OD₆₀₀ of 5–6 and were induced overnight at 15 °C with 100 μM IPTG.

Purification

Cell pellets were resuspended in lysis buffer (50 mM Tris [pH 8.0], 500 mM NaCl, 1 mM PMSF, and 0.1 ml general protease inhibitor Sigma P2714), lysed by sonication at 4 °C, and mixed for 30 minutes with HisLink resin (Promega). Resin was washed by batch-method and loaded into gravity columns; protein was eluted with elution buffer (lysis buffer plus 250 mM imidazole and 10% glycerol). The tag was removed with thrombin (one unit added (Sigma T9681) per milligram of protein) by incubation overnight at 4 °C. The sample was subjected to size-exclusion chromatography using HiLoad Superdex 200 resin (GE Healthcare) pre-equilibrated with gel filtration buffer (lysis buffer plus 1 mM TCEP [Tris (2-carboxyethyl) phosphine hydrochloride] and 1 mM EDTA)). Protein was concentrated to 250 μM and incubated overnight at 4 °C with 10 mM MgCl₂ and 5–10 mM ATP in order to drive complete auto-phosphorylation of the kinase. Excess nucleotide or other reagents were removed by a HiTrapQ HP column (GE Healthcare). Purified protein was exchanged into gel filtration buffer by concentration and dilution and used at 10–20 mg/ml for crystallization studies.

Crystallization, Data Collection, and Structure Solution

As described previously, crystals of EphA3 form in multiple conditions but only after degradation to construct boundaries corresponding to Thr595–Thr912 [20]. For all crystallization experiments used in the current work, protein purified as above was exposed to 10 mM adenosine 5'-(β,γ-imido) triphosphate (AMP-PNP tetralithium salt, Sigma) and 10 mM MgCl₂, along with the peptide of interest, and incubated at 4 °C for at least 30 minutes prior to co-crystallization trials. EPHOPT peptide and OPTYK peptides were ordered from the peptide synthesis core facility at Tufts University; EPHOPT is soluble to 100 mM in aqueous buffer; OPTYF was used as a 60 mM stock in aqueous solution. Optimal conditions for co-crystallization were found to be 22%–28% Peg 3350, 50 mM Tris (pH 7.5), and 40 mM (NH₄)₂SO₄ using the hanging drop vapor diffusion method and 1 μL+1 μL drops. Crystals typically appeared 24 hr after incubation at 18 °C; the typical size of crystals was 400×200×200 microns. Crystals were harvested into cryoprotection buffer (1:1 mixture of glycerol and mother liquor; final concentration of glycerol was 15%) and frozen in liquid nitrogen. Diffraction data from co-crystals of EphA3 with peptide were collected on an FR-E generator equipped with an RAXIS-IV++ detector (Rigaku) and integrated and scaled using either the HKL2000 program package for the EPHOPT complex [23,24], or iMosflm and SCALA for the OPTYF complex [25,26]. PHASER was used with the coordinates of 2GSF as the starting model in order to obtain initial phasing [27]. Manual rebuilding was performed using WinCoot [28] and refined using REFMAC [29,30] in the CCP4i program suite [31]. The coordinates and structure factors for the structures of EphA3 described in the text have been deposited into the PDB with codes 3FXX (EPHOPT complex) and 3FY2 (OPTYF complex). All models have excellent stereochemistry as judged by PROCHECK [32] and MOLPROBITY [33], with no

residues in disallowed regions of Ramachandran space. Statistics of model refinement for both structures are provided in Table 1.

Kinase Specificity Determination

EphA3 phosphorylation site sequence specificity was determined by screening a 198 member positional scanning peptide library [34]. Unphosphorylated EphA3 (1.1 mg/ml), purified as described above, was activated by incubation in 20 mM Tris (pH 8.0), 10 mM MgCl₂, 100 mM NaCl, 2 mM DTT, 5% glycerol with 5 mM ATP for 30 min at ambient temperature. Peptides were arrayed at 50 μM in multiwell plates in 50 mM Tris (pH 7.5), 10 mM MgCl₂, 1 mM DTT, 0.1% Tween 20. Reactions were begun by adding activated EphA3 to 70 – 800 ng/ml and ATP to 50 μM (including 0.3 μCi/μL γ-[³³P]-ATP). Peptides had the general sequence GAXXXX-Y-XXXXAGKK(biotin), where X is a roughly equimolar mixture of the 18 amino acids excluding cysteine, and tyrosine. In each peptide, one of the X positions was replaced with 1 of 22 residues (one of the 20 unmodified amino acids, pSer or pTyr). After incubation at 30° C for 2 hr, aliquots of each reaction were simultaneously transferred to streptavidin membrane, which was processed as previously described [34].

Kinase Assays

For all enzymatic assays presented in the current study, EphA3, EphB3, EphA5 and EphA7 proteins were preincubated with 10 mM each ATP and MgCl₂ as described above in order to promote full autophosphorylation prior to assaying for enzymatic activity against peptide substrates. All proteins were purified using a HiTrapQ HP column (GE Healthcare) as described above in order to remove excess nucleotide from the reaction; all proteins were exchanged into identical reaction buffer by concentration and dilution. EphB2 and EphB4 were purchased from New England Biolabs in their active form and were not further modified before kinetic analysis. Enzymatic activity of all wild-type EPH RTKs and EphA3 mutants (N830A and K785N) were determined using the ADP-Quest Kit and following the protocol provided by DiscoverX (Fremont, CA) as described previously [20]. ADP production was followed by monitoring the increase in fluorescence (excitation at 530 nm and emission at 590 nm) using a fluorescence plate reader (Spectramax Gemini, Molecular Devices). All reactions were performed at room temperature in a final volume of 50 μl. Kinetic constants were determined by varying EPHOPT, OPTYF and OPTYK peptides concentrations from 1 to 4000 μM at 200 μM ATP. Protein concentrations of 10 nM to 5 μM were used in the assays. All experiments were performed in duplicate, and the values determined for kinase activity were corrected for background ADP production. K_m and V_{max} values were calculated using the Michaelis-Menten equation using Sigmaplot 9.0, and standard deviation was calculated from two independent experiments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The Structural Genomics Consortium is a registered charity (number 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research and the Wellcome Trust. S. Parker and B. Turk are supported by a grant from the U.S. National Institutes of Health (GM079498).

References

1. Himanen JP, Saha N, Nikolov DB. Cell-cell signaling via Eph receptors and ephrins. *Current Opinion in Cell Biology* 2007;19:534–542. [PubMed: 17928214]
2. Pasquale EB. Eph-ephrin bidirectional signaling in physiology and disease. *Cell* 2008;133:38–52. [PubMed: 18394988]
3. Surawska H, Ma PC, Salgia R. The role of ephrins and Eph receptors in cancer. *Cytokine & Growth Factor Reviews* 2004;15:419–433. [PubMed: 15561600]
4. Lackmann M, Boyd AW. Eph, a Protein Family Coming of Age: More Confusion, Insight, or Complexity? *Sci Signal* 2008;1 re2-
5. Noren NK, Pasquale EB. Eph receptor-ephrin bidirectional signals that target Ras and Rho proteins. *Cell Signal* 2004;16:655–666. [PubMed: 15093606]
6. Smith FM, Vearing C, Lackmann M, Treutlein H, Himanen J, Chen K, Saul A, Nikolov D, Boyd AW. Dissecting the EphA3/Ephrin-A5 interactions using a novel functional mutagenesis screen. *Journal of Biological Chemistry* 2004;279:9522–9531. [PubMed: 14660665]
7. Clifford N, Smith L, Powell J, Gattenlöhner S, Marx A, O'Connor R. The EphA3 receptor is expressed in a subset of rhabdomyosarcoma cell lines and suppresses cell adhesion and migration. *Journal of Cellular Biochemistry* 2008;105:1250–1259. [PubMed: 18814179]
8. Cheng H-J, Nakamoto M, Bergemann AD, Flanagan JG. Complementary gradients in expression and binding of ELF-1 and Mek4 in development of the topographic retinotectal projection map. *Cell* 1995;82:371–381. [PubMed: 7634327]
9. Feldheim DA, Nakamoto M, Osterfield M, Gale NW, DeChiara TM, Rohatgi R, Yancopoulos GD, Flanagan JG. Loss-of-function analysis of EphA receptors in retinotectal mapping. *Journal of Neuroscience* 2004;24:2542–2550. [PubMed: 15014130]
10. Lai C, Lemke G. An extended family of protein-tyrosine kinase genes differentially expressed in the vertebrate nervous system. *Neuron* 1991;6:691–704. [PubMed: 2025425]
11. Stephen LJ, Fawkes AL, Verhoeve A, Lemke G, Brown A. A critical role for the EphA3 receptor tyrosine kinase in heart development. *Developmental Biology* 2007;302:66–79. [PubMed: 17046737]
12. Bardelli A, Parsons DW, Silliman N, Ptak J, Szabo S, Saha S, Markowitz S, Willson JK, Parmigiani G, Kinzler KW, Vogelstein B, Velculescu VE. Mutational analysis of the tyrosine kinome in colorectal cancers. *Science* 2003;300:949. [PubMed: 12738854]
13. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, Sougnez C, Greulich H, Muzny DM, Morgan MB, Fulton L, Fulton RS, Zhang Q, Wendl MC, Lawrence MS, Larson DE, Chen K, Dooling DJ, Sabo A, Hawes AC, Shen H, Jhangiani SN, Lewis LR, Hall O, Zhu Y, Mathew T, Ren Y, Yao J, Scherer SE, Clerc K, Metcalf GA, Ng B, Milosavljevic A, Gonzalez-Garay ML, Osborne JR, Meyer R, Shi X, Tang Y, Koboldt DC, Lin L, Abbott R, Miner TL, Pohl C, Fewell G, Haipke C, Schmidt H, Dunford-Shore BH, Kraja A, Crosby SD, Sawyer CS, Vickery T, Sander S, Robinson J, Winckler W, Baldwin J, Chirieac LR, Dutt A, Fennell T, Hanna M, Johnson BE, Onofrio RC, Thomas RK, Tonon G, Weir BA, Zhao X, Ziaugra L, Zody MC, Giordano T, Orringer MB, Roth JA, Spitz MR, Wistuba, Ozenberger B, Good PJ, Chang AC, Beer DG, Watson MA, Ladanyi M, Broderick S, Yoshizawa A, Travis WD, Pao W, Province MA, Weinstock GM, Varmus HE, Gabriel SB, Lander ES, Gibbs RA, Meyerson M, Wilson RK. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 2008;455:1069–1075. [PubMed: 18948947]
14. Wood LD, Calhoun ES, Silliman N, Ptak J, Szabo S, Powell SM, Riggins GJ, Wang TL, Yan H, Gazdar A, Kern SE, Pennacchio L, Kinzler KW, Vogelstein B, Velculescu VE. Somatic mutations of GUCY2F, EPHA3, and NTRK3 in human cancers. *HumMutat* 2006;27:1060–1061.
15. Balakrishnan A, Bleeker FE, Lamba S, Rodolfo M, Daniotti M, Scarpa A, van Tilborg AA, Leenstra S, Zanon C, Bardelli A. Novel somatic and germline mutations in cancer candidate genes in glioblastoma, melanoma, and pancreatic carcinoma. *Cancer Research* 2007;67:3545–3550. [PubMed: 17440062]
16. Fox BP, Tabone CJ, Kandpal RP. Potential clinical relevance of Eph receptors and ephrin ligands expressed in prostate carcinoma cell lines. *Biochemical & Biophysical Research Communications* 2006;342:1263–1272. [PubMed: 16516143]

17. Day B, To C, Himanen JP, Smith FM, Nikolov DB, Boyd AW, Lackmann M. Three distinct molecular surfaces in ephrin-A5 are essential for a functional interaction with EphA3. *Journal of Biological Chemistry* 2005;280:26526–26532. [PubMed: 15901737]
18. Himanen JP, Chumley MJ, Lackmann M, Li C, Barton WA, Jeffrey PD, Vearing C, Geleick D, Feldheim DA, Boyd AW, Henkemeyer M, Nikolov DB. Repelling class discrimination: ephrin-A5 binds to and activates EphB2 receptor signaling. *Nature Neuroscience* 2004;7:501–509. [see comment].
19. Murai KK, Pasquale EB. Eph'ective signaling: forward, reverse and crosstalk. *Journal of Cell Science* 2003;116:2823–2832. [PubMed: 12808016]
20. Davis TL, Walker JR, Loppnau P, Butler-Cole C, Allali-Hassani A, Dhe-Paganon S. Autoregulation by the Juxtamembrane Region of the Human Ephrin Receptor Tyrosine Kinase A3 (EphA3). *Structure* 2008;16:873–884. [PubMed: 18547520]
21. Hubbard SR. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO Journal* 1997;16:5572–5581. [PubMed: 9312016]
22. Warner N, Wybenga-Groot LE, Pawson T. Analysis of EphA4 receptor tyrosine kinase substrate specificity using peptide-based arrays. *FEBS Journal* 2008;275:2561–2573. [PubMed: 18422655]
23. Minor W, Cymborowski M, Otwinowski Z. Automatic system for crystallographic data collection and analysis. *Acta Physica Polonica A* 2002;101:613–619.
24. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* 1997;276:307–326.
25. Leslie AG. The integration of macromolecular diffraction data. *Acta Crystallographica Section D-Biological Crystallography* 2006;62:48–57.
26. Evans P. Scaling and assessment of data quality. *Acta Crystallographica Section D-Biological Crystallography* 2006;62:72–82.
27. Read RJ. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta CrystallogrDBiolCrystallogr* 2001;57:1373–1382.
28. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta CrystallogrDBiolCrystallogr* 2004;60:2126–2132.
29. Potterton E, Briggs P, Turkenburg M, Dodson E. A graphical user interface to the CCP4 program suite. *Acta Crystallographica Section D-Biological Crystallography* 2003;59:1131–1137.
30. Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta CrystallogrDBiolCrystallogr* 2001;57:122–133.
31. Collaborative Computational Project n. The CCP4 suite: programs for protein crystallography. *Acta CrystallogrDBiolCrystallogr* 1994;50:760–763.
32. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. Procheck - A Program to Check the Stereochemical Quality of Protein Structures. *Journal of Applied Crystallography* 1993;26:283–291.
33. Davis IW, Murray LW, Richardson JS, Richardson DC. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* 2004;32:W615–W619. [PubMed: 15215462]
34. Hutti JE, Jarrell ET, Chang JD, Abbott DW, Storz P, Toker A, Cantley LC, Turk BE. A rapid method for determining protein kinase phosphorylation specificity. *NatMethods* 2004;1:27–29.
35. Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ. Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences* 1998;23:403–405. [PubMed: 9810230]
36. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997;25:4876–4882. [PubMed: 9396791]

Figure 1.

Phosphorylation motifs and optimal substrate design for EphA3. Biotinylated peptides bearing the indicated residue at the indicated position relative to a central tyrosine phosphoacceptor site were subjected to phosphorylation by EphA3 with radiolabeled ATP. Aliquots of each reaction were subsequently spotted onto a streptavidin membrane, which was washed, dried, and exposed to a phosphor screen. In the top panel is shown a representative array from three separate experiments. Quantified spot intensities representing the average of the three runs are provided in the lower panel; amino acids in bold show the highest significant difference for array positions from -2 to +4; numbers in parenthesis indicate the relative signal:to:noise ratio at each position. The optimized sequence derived from these results was used for all kinetic

and structural work; this sequence (named EPHOPT in the manuscript) is KQWDNYE-pY-IW.

Figure 2.

Views of the EphA3: EPHOPT complex structure. (A) The structure of EphA3 kinase in complex with the EPHOPT peptide. EphA3 is shown in ribbon representation and in teal; the substrate is shown in purple and in stick representation. The ATP analog AMP-PNP is shown in stick representation in orange. (B) Representative density, 1.3 σ . Shown is the backbone for four residues and phosphotyrosine of peptide and the Lys785 region of kinase. (C) Enlarged view of the EphA3: EPHOPT interface. Coloring is as in (A).

Figure 3.

Structural changes in EphA3 kinase upon binding substrates, and comparison of the EPHOPT and OPTYF complex structures. (A) A series of EphA3 structures without substrate bound [20] are shown superimposed upon the EphA3:EPHOPT complex structure. Coloring is as follows: mint green, EphA3:EPHOPT complex (PDB ID 3FXX); forest green, a low activity EphA3 conformation (2QOQ); in goldenrod and teal, two intermediate activity conformations (2QOB and 2GSF); pink, a higher activity conformation (2QO9); and in firebrick, a high activity conformation (2QOC). The EPHOPT substrate is not shown so that the differences in conformation in the region around AL residue Lys785 and α G residue Asn830 can be seen clearly. In order to assess the relative flexibility of these two regions, several other kinase

residues within 4Å of the EPHOPT substrate are also shown. These residues are clearly fixed in their orientation regardless of whether substrate is bound or what activity state the contributing structure represents. (B) A general view of the EphA3: EPHOPT interface. (C) The EphA3:OPTYF interface (PDB 3FY2). Compare the orientation of OPTYF residues Trp-4, Asn-3, Asp-2, and Glu+1 compared to those of EPHOPT in (B).

Figure 4.

Alignment of EPH kinase domains highlighting the region of substrate interaction. Alignment was performed using CLUSTALX [35,36], coloring is by chemical property. Specific residues discussed in the text are labeled and highlighted with boxes; residue numbers correspond to EphA3 numbering. The alignment corresponds to EphA3 residues 698–854. Secondary structural elements are indicated below the alignment.

Table 1
Crystallographic statistics

Atomic coordinates for the structures discussed in the text have been deposited into the RCSB and PDB codes are listed in the Experimental Procedures and in the table.

Dataset	EphA3:EPHOPT	EphA3:OPTYF
Spacegroup	P 1 21 1	P 1 21 1
Unit cell (Å)	53.46 38.20 76.65	53.82 38.26 76.37
Unit cell (°)	90.00 102.15 90.00	90.00 102.05 90.00
Data Collection		
Beamline	FRE-HR	FRE-HR
Wavelength	1.54178	1.54178
Resolution	26.5-1.7	26.73-1.8
Unique Reflections	33335	29459
Data Redundancy, -fold ^{**}	3.4 (2.4)	3.6 (3.5)
Completeness, %	98 (84)	100 (100)
I/sigI	28 (3)	21 (4)
Rsym ^{***}	0.044 (0.312)	0.049 (0.224)
Refinement		
Resolution	1.70	1.8
Number of Reflections	31644	26963
All Atoms (solvent)	5839(301)	5457 (233)
Rwork (Rfree) ^{****}	0.166(0.19)	0.179(0.21)
Rmsd bond length	0.01	0.011
Rmsd bond angle	1.28	1.26
Ramachandran Plot		
Most favoured, %	91.5	90.9
Additionally Allowed, %	8.1	8.7
Generously Allowed, %	0.4	0.4
Disallowed, %	0	0
PDB code	3FXX	3FY2
Modeled Residues		
	kinase: 606–774; 784–892; 895–904	607–773; 784–892; 896–904
	peptide: OWDNYE-pY-IW	WDNYE-F-IW

^{**} Highest resolution shell is shown in parenthesis.

^{***} $R_{sym} = 100 \times \sum(|I - \langle I \rangle|) / \sum(\langle I \rangle)$, where I is the observed intensity and $\langle I \rangle$ is the average intensity from multiple observations of symmetry related reflections.

Rfree value was calculated with 5% of the data.

Table 2**Kinetics**

Kinetic Data. Top panel, the effect of varying amino acids at the +2 position of substrate. The sequences of the peptides tested are as follows: EPHOPT = KQWDNYEpYIW; OPTYF = KQWDNYEFIW; OPTYK = KQWDNYEKIW. Bottom panel, the effect of mutating EphA3.

EPHA3 wild-type protein			
Peptide	Km (uM)	kcat(min-1)	kcat/Km (mM-1min-1)
EPHOPT	18 ±4.4	850 ±44	47222
OPTYF	30 ±5.7	421 ±30	14033
OPTYK	107±15.6	220 ±27	2056
EPHA3 N830A mutant			
Peptide	Km (um)	kcat (min-1)	kcat/Km (mM-1 min-1)
EPHOPT	39.5 ± 0.7	33 ± 1.4	835
OPTYF	152 ± 35	30 ± 1.4	197
EPHA3 K785E mutant			
Peptide	Km (uM)	kcat (min-1)	kcat/Km (mM-1 min-1)
EPHOPT	188 ± 64	34 ± 0.7	181
OPTYF	148 ± 10	23 ± 8	155