Structure of *Escherichia coli* tyrosine kinase Etk reveals a novel activation mechanism

Daniel C Lee¹, Jimin Zheng¹, Yi-Min She²
and Zongchao Jia^{1,*}

¹Department of Biochemistry, Queen's University, Kingston, Ontario, Canada and ²Department of Chemistry, Queen's University, Kingston, Ontario, Canada

While protein tyrosine (Tyr) kinases (PTKs) have been extensively characterized in eukaryotes, far less is known about their emerging counterparts in prokaryotes. The inner-membrane Wzc/Etk protein belongs to the bacterial PTK family, which has an important function in regulating the polymerization and transport of virulence-determining capsular polysaccharide (CPS). The kinase uses a unique two-step activation process involving intra-phosphorylation of a Tyr residue, although the molecular mechanism remains unknown. Herein, we report the first crystal structure of a bacterial PTK, the C-terminal kinase domain of *Escherichia coli* Tyr kinase (Etk) at 2.5-Å resolution. The fold of the Etk kinase domain differs markedly from that of eukaryotic PTKs. Based on the observed structure and supporting mass spectrometric evidence of Etk, a unique activation mechanism is proposed that involves the phosphorylated Tyr residue, Y574, at the active site and its specific interaction with a previously unidentified key Arg residue, R614, to unblock the active site. Both *in vitro* kinase activity and *in vivo* antibiotics resistance studies using structure-guided mutants further support the novel activation mechanism.

The EMBO Journal (2008) 27, 1758–1766. doi:10.1038/emboj.2008.97; Published online 22 May 2008

Subject Categories: signal transduction; structural biology

Keywords: activation; crystallography; Etk; protein tyrosine kinase; Wzc

Introduction

In eukaryotes, protein tyrosine (Tyr) kinases (PTKs) control various cellular pathways such as gene expression, metabolism, cell growth and apoptosis, as well as membrane transport (Cowan-Jacob, 2006). In bacteria, however, the physiological role of Tyr phosphorylation has just been revealed (Grangeasse *et al*, 2003). In fact, protein Tyr phosphorylation was initially thought to exist solely in eukaryotes (Levitzi and Gazit, 1995), and only since the last decade have researchers begun to investigate prokaryotic protein Tyr phosphorylation (Freestone *et al*, 1998; Grangeasse *et al*,

2007). So far, none of the PTK-related structures determined have a prokaryotic source. Among the ~20 PTKs characterized in Gram-negative and Gram-positive bacteria to date, all but three are homologues of Wzc/Etk (Cozzone *et al*, 2004; Grangeasse *et al*, 2007), the first and only extensively studied bacterial PTK family also known as bacterial Tyr (BY) kinases (Grangeasse *et al*, 2007).

Wzc/Etk PTKs are involved in synthesis and transportation of extracellular and capsular polysaccharides (EPS and CPS, respectively), which have important protective functions in promoting virulence in many bacterial strains, across the outer membrane (Whitfield, 2006). In *Escherichia coli*, four CPS polymerization and exportation mechanisms have been identified (Orskov *et al*, 1977; Whitfield and Roberts, 1999). Group-1 and Group-4 capsules, which are found in *E. coli* isolates causing intestinal infections, use a Wzy-dependent polymerization system (Whitfield, 2006). This novel, trans-envelope complex coordinates synthesis and translocation of EPS/CPS through (de)phosphorylation of PTK homologues (Whitfield and Roberts, 1999; Wugeditsch *et al*, 2001; Peleg *et al*, 2005; Grangeasse *et al*, 2007). In *E. coli*, there are three BY kinase homologues, Wzc_{CPS}, Etk (*Escherichia coli* Tyr kinase) and Wzc_{CA}, which are specifically involved in regulating translocation assemblies of Group-1 CPS, Group-4 CPS and colonic acid EPS (Ilan *et al*, 1999; Vincent *et al*, 1999; Paiment *et al*, 2002). Furthermore, studies have uncovered endogenous protein substrates of Wzc/Etk, namely a UDP-sugar dehydrogenase (Grangeasse *et al*, 2003; Mijakovic *et al*, 2003), a glycosyltransferase (Minic *et al*, 2007) and a heat-shock σ -factor (Klein *et al*, 2003), suggesting a more versatile function of the Wzc/Etk family in the bacterial cells. For more information, please see two recent reviews (Whitfield, 2006; Grangeasse *et al*, 2007).

Wzc/Etk proteins are inner-membrane PTKs in Gram-negative bacteria, with two trans-membrane α -helices separating the N-terminal periplasmic and C-terminal cytoplasmic domains. This arrangement somewhat resembles the eukaryotic receptor PTKs (Cowan-Jacob, 2006) despite the absence of any significant sequence homology between them. The ~270-residue cytoplasmic domains of Wzc/Etk contain the ATP-binding Walker-A and Walker-B motifs, and a Tyr-rich cluster (5–7 Tyr in a 12-residue stretch) at the C-terminus (Grangeasse *et al*, 2002). Wzc/Etk phosphorylation proceeds through a cooperative, two-step mechanism involving both intra- and inter-phosphorylation. First, a key upstream Tyr residue (Y569 in Wzc, Y574 in Etk) is autophosphorylated through an intra-molecular process, which results in protein kinase activation. Subsequently, the clustered C-terminal Tyr residues undergo inter-molecular phosphorylation (Grangeasse *et al*, 2002). Wzc has been shown to harbour ATPase activity, which is enhanced by intra-phosphorylation at this Tyr residue as well (Soulat *et al*, 2007). Moreover, recent studies have indicated a direct interaction between Wzc and Wza, the outer-membrane channel responsible for CPS/EPS transport (Nesper *et al*, 2003; Reid and Whitfield, 2005; Collins *et al*, 2007).

*Corresponding author. Department of Biochemistry, Queen's University, CDN-Kingston, Kingston, Ontario, Canada K7L3N6.
Tel.: +1 613 533 6277; Fax: +1 613 533 2497;
E-mail: jia@post.queensu.ca

Received: 12 November 2007; accepted: 25 April 2008; published online: 22 May 2008

Although not in direct contact with the rest of the CPS transport assembly, the Etk/Wzc kinase (cytoplasmic) domain has been shown to regulate CPS transport through phosphorylation of its C-terminal Tyr-cluster (Paiment *et al*, 2002). Mutational studies have shown that it is the degree of phosphorylation on the Wzc C-terminal Tyr-cluster, rather than the phosphorylation of specific Tyr residues, that regulates capsule transport (Paiment *et al*, 2002).

Several prominent structural studies have led to significant progress in characterizing the *E. coli* Group-1 assembly complex. The electron microscopic (EM) structure of Wzc was recently determined to 14-Å resolution, from which it can be discerned that Wzc adopts a native tetrameric conformation solely through its N-terminal, periplasmic domain (Collins *et al*, 2006). In addition, the high-resolution crystal structure of the cross-periplasmic-membrane Wza octamer channel showed that this assembly contains a large central cavity that is open to the extracellular environment, but closed to the periplasm (Dong *et al*, 2006). Furthermore, a recently determined EM structure of the Wzc tetramer complexed with the Wza octamer showed the Wza octamer channel in its open conformation, with four Wzc monomers interacting with the periplasmic end of the channel Wza. Interestingly, these Wzc monomers were not found to be associated with one another (Collins *et al*, 2006, 2007). Therefore, oligomerization of Wzc (and Etk), and their interaction with the respective channel protein (Wza and YccZ), seems to be critical to the export regulation of CPS. The regulatory machinery of the key Tyr residue phosphorylation, as well as the specific role of Wzc and Etk in these assemblies, however, remains unknown.

Results and discussion

The crystal structure of the Etk kinase domain

The crystal structure of the Etk kinase domain (residues 451–726, see crystallographic statistics in Supplementary Table 1) reveals a single-domain scaffold consisting of a central, eight-stranded, predominantly parallel β -sheet surrounded by 13 helices of varying lengths (Figure 1A). In the asymmetric unit, there are two molecules of Etk, which are essentially identical, including the bound ligands. The active site, which contains the well-conserved ATP-binding Walker motifs, is located at the end of the long E-helix and the loop regions that follow strands 3 and 6. We further determined the ADP-bound complex structure, which confirms the active site identification (Figure 2A and B). A structural homology search using Dali (Holm and Sander, 1995) has shown that Etk intriguingly shares a high z -score (15.6) with MinD (Figure 1B), an ATPase involved in inhibition of cell division (Lutkenhaus and Sundaramoorthy, 2003). Within the β -structure core, Etk and *Pyrococcus furiosus* MinD (Hayashi *et al*, 2001) overlap well with a root mean-square deviation value of 1.37 Å, despite having only ~18% overall sequence identity. The general folding of Etk also resembles that of Soj in the Gram-negative bacterium *Thermus thermophilus*, an ATPase sharing similar function and sequence with MinD (Leonard *et al*, 2005).

In contrast, the Etk kinase domain has a fold completely different from that of both eukaryotic cytoplasmic receptor Tyr kinases (RTKs) (Figure 1C) and other PTKs in general. Sequence alignment between the members of the Etk/Wzc

Figure 1 The Etk kinase domain structure. (A) The overall structure of the Etk kinase domain. The molecular switch of Etk activation, Y574, is coloured red. Helices and strands are labelled sequentially. ATP-binding Walker motifs and the C-terminal Tyr-rich cluster are also indicated. (B) Structure of *P. furiosus* cell division ATPase MinD (PDB code 1G3R; Hayashi *et al*, 2001). (C) Structure of the human insulin RTK domain (PDB code 1IRK; Hubbard *et al*, 1994).

PTK family in both Gram-negative and Gram-positive bacteria, together with MinD ATPases, reveals very clear sequence conservation over the central parallel strands, but not over the peripheral α -helices despite their similar spatial locations. In addition, the bacterial PTK-family members and MinD ATPases share the same Walker motifs (Figure 3). In MinD, the corresponding residues of Etk K545, T546 (for ATP binding), and D567, D569, D647 (for Mg^{2+} binding), as well as the ADP ligand, are all found at nearly identical positions in the active site (Hayashi *et al*, 2001; Figure 4A and B).

The structural similarity between Etk, a kinase, and MinD, an ATPase, motivated us to search for structural differences that might explain (i) why they differ in enzymatic functions and (ii) why Etk involves a two-step activation mechanism. Two significant differences in the active sites became evident. First, the left half of the MinD catalytic site (Figure 4B) is narrowed by an α -helix (P201–E209), while the Etk-binding pocket is not obstructed here, allowing greater access to the substrate (Figure 4A). However, on the other (right) side of the Etk-binding pocket, residue Y574, located at the beginning of helix-F, has its bulky side chain pointing directly into the active site, blocking the magnesium-binding aspartic acid trio (D567, D569 and D647, Walker motifs A' and B). In Wzc, phosphorylation of this Tyr residue is critical for Wzc kinase activity (Grangeasse *et al*, 2002). In contrast, in MinD this

Figure 2 The electron density of the active site in Etk and the Etk-ADP complex structures. (A) The $2F_o - F_c$ electron density at 1σ level in the complex. Dephosphorylated Y574 points directly to the active site, impairing substrate access. (B) A simulated annealing omit map ($F_o - F_c$) for ADP at 3σ density level. (C) The $2F_o - F_c$ electron density at 1σ level in Etk. A sulphate (or phosphate) group is located near R614. A modelled open conformation of Y574 is shown in green colour. (D) A simulated annealing omit map ($F_o - F_c$) for phosphate at 4σ density level.

		RK-cluster (479-498)													
Etk	<i>E. coli</i>	468	ATIPMSEWLDK	RTRLRKKNLFSNQQRHRTKN	IPFLAVDNP	507									
Wsc	<i>E. coli</i>	468	ASIFLSEWQK	ARDSVKTIK---	GIKRYKQSQL--	LAVGNP	502								
Wsc	<i>K. pneumoniae</i>	469	ASIPVSESNPK	NVIAKRLN---	KRDDSRPKVL--	LATENP	503								
CpsA	<i>E. pyrifoliae</i>	468	ASVPLSEWQR	KKDQEVQLKRKLAAKVDPHRL--	LALGNP	505									
AmsA	<i>E. Amylovora</i>	468	ASVPLSEWQR	KKDQETLLKRKLDAKDTPHNL--	LALGNP	505									
ORF	<i>S. Typhi</i>	467	ASIFLSEWQK	ARDSVKTIK---	GIKRYKQSQL--	LAVGNP	501								
		Walker-A										Gram-negative Tyrosine kinase			
Etk	<i>E. coli</i>	533	ILMITGATPDS	KTFVSSTL...	VLFI	ADLR	RGY	SHNLF	579						
Wsc	<i>E. coli</i>	528	VLMMTGVSP	SIKTFVCANL...	VLLI	CDMR	KGY	THELL	574						
Wsc	<i>K. pneumoniae</i>	525	LLMISGASPN	AKTFVSSNL...	VIFI	ADLR	KGY	THKLF	575						
CpsA	<i>E. pyrifoliae</i>	531	ILMVTGASPG	IKTFICANL...	VLFI	GDMR	RGY	THDLL	577						
AmsA	<i>E. Amylovora</i>	531	ILMITGATPDS	KTFVSSTL...	VLFI	ADLR	RGY	SHNLF	579						
ORF	<i>S. Typhi</i>	527	VLMITGVSP	SIKTFVCANL...	VLLI	CDMR	KGY	THELL	572						
CpsD	<i>S. agalactiae</i>	36	ILAITSFREGE	KTTSTSL...	TLII	ADTR	NSVM	SGTF	83						
Wsc	<i>S. pneumoniae</i>	37	VISVTSVNP	GEKTTTSVNI...	TLII	GDTR	NSVM	SGVF	83						
MinD	<i>E. coli</i>	4	IIVVTSKGG	GVKTTSSAAL...	TVVI	FDIG	LRNL	DLIM	50						
MinD	<i>P. furiosus</i>	4	IISIVSGKG	GGTKTTVTANL...	VLAV	EGLT	MANL	SLVL	50						
												Gram-positive Tyrosine kinase			
		Walker-A' Tyr574													
Etk	<i>E. coli</i>	533	ILMITGATPDS	KTFVSSTL...	VLFI	ADLR	RGY	SHNLF	579						
Wsc	<i>E. coli</i>	528	VLMMTGVSP	SIKTFVCANL...	VLLI	CDMR	KGY	THELL	574						
Wsc	<i>K. pneumoniae</i>	525	LLMISGASPN	AKTFVSSNL...	VIFI	ADLR	KGY	THKLF	575						
CpsA	<i>E. pyrifoliae</i>	531	ILMVTGASPG	IKTFICANL...	VLFI	GDMR	RGY	THDLL	577						
AmsA	<i>E. Amylovora</i>	531	ILMITGATPDS	KTFVSSTL...	VLFI	ADLR	RGY	SHNLF	579						
ORF	<i>S. Typhi</i>	527	VLMITGVSP	SIKTFVCANL...	VLLI	CDMR	KGY	THELL	572						
CpsD	<i>S. agalactiae</i>	36	ILAITSFREGE	KTTSTSL...	TLII	ADTR	NSVM	SGTF	83						
Wsc	<i>S. pneumoniae</i>	37	VISVTSVNP	GEKTTTSVNI...	TLII	GDTR	NSVM	SGVF	83						
MinD	<i>E. coli</i>	4	IIVVTSKGG	GVKTTSSAAL...	TVVI	FDIG	LRNL	DLIM	50						
MinD	<i>P. furiosus</i>	4	IISIVSGKG	GGTKTTVTANL...	VLAV	EGLT	MANL	SLVL	50						
												MinD ATPase			
		Walker-B													
Etk	<i>E. coli</i>	608	GFDVITRG	QVPPN...	LVIV	DTTP	650								
Wsc	<i>E. coli</i>	603	KFDLIPRG	QVPPN...	LVLI	DTTP	645								
Wsc	<i>K. pneumoniae</i>	604	GFDYISRG	SVPPN...	IVIL	DTTP	646								
CpsA	<i>E. pyrifoliae</i>	606	GFDFIARG	QVPPN...	LVLI	DTTP	648								
AmsA	<i>E. Amylovora</i>	608	GFDVITRG	QVPPN...	LVIV	DTTP	648								
ORF	<i>S. Typhi</i>	602	NFDLIPRG	QVPPN...	LVLI	DTTP	644								
CpsD	<i>S. agalactiae</i>	113	RLMVVPSG	QVPPN...	YIIV	DTTP	155								
Wsc	<i>S. pneumoniae</i>	113	NLFVVQSG	TVPEN...	YIIV	DTTP	156								
MinD	<i>E. coli</i>	82	NLYILPAS	QTR...	FIVC	CPA	123								
MinD	<i>P. furiosus</i>	80	NVIVLP	GAVDWE...	FILV	CPA	121								
												C-terminus Tyr-cluster			
Etk	<i>E. coli</i>	726	GAILNGVIK	RASTAYSYGYNYGY	SYSEKE	726									
Wsc	<i>E. coli</i>	720	GVILNSIF	RASATYQDGYGYEY	EYKSDAK	720									
Wsc	<i>K. pneumoniae</i>	722	GCILNGV	VKKATNKYGYGYND	YSYSDKK	722									
CpsA	<i>E. pyrifoliae</i>	726	GILNAVVR	KSANNYGYGYDYDY	YSYQGEKS	726									
AmsA	<i>E. Amylovora</i>	726	GILNAVVR	KSANNYGYGYDYDY	YSYQGEKS	726									
ORF	<i>S. Typhi</i>	719	GVILNSIF	RATGYQDGYGYEY	EYQSDSK	719									
CpsD	<i>S. agalactiae</i>	232	GILNKNV	NEVATYGDYGYGN	YGKRDKRK	232									
Wsc	<i>S. pneumoniae</i>	227	GVVLNKL	DISVDKYGVYGS	YGNYGKK	227									
MinD	<i>E. coli</i>	237	GVIPEDQ	SVLRASNQGEF	ILDINADAGK	237									
MinD	<i>P. furiosus</i>	195	GFVLNRY	GRSDRDIPPEAAED	VMEVPLLA	195									

Figure 3 Sequence alignment of prokaryotic Tyr kinases and MinD ATPases. The ATP-binding Walker motifs (green shades) are conserved among all sequences. The C-terminal Tyr-cluster is conserved among bacterial Tyr kinases but not ATPases. The two important residues for Etk activation, Y574 and R614, are only present in Gram-negative Tyr kinases.

Y574 position is occupied by N45, which does not cause narrowing of the MinD active site to the same extent as Y574 does in Etk. The amide nitrogen of this asparagine contributes to magnesium cation binding (Hayashi *et al*, 2001).

To our knowledge, no Ser/Thr kinase activity has been reported on Etk, although this possibility cannot be completely excluded. If Etk indeed is specific for Tyr, it is likely to be contributed by its deep pocket at the phosphorous-transferring site that is not readily accessible by shorter Ser/Thr side chain. To test this postulation, we modelled the Tyr residue of the peptide substrate at the active site with Ser, and cofactor ADP with ATP. The distance between Ser side-chain oxygen

and ATP γ -phosphorous is ~ 4.5 Å. This distance is too long for kinase activity, compared with that between the Tyr side-chain oxygen and ATP γ -phosphorous, which is ~ 2.5 Å. Therefore, it is likely that Etk is specific for Tyr. Indeed, from our mass spectrometry (MS) results it was clear that on Etk, all phosphorylation occurred on Tyr residues (Supplementary Figures 1 and 2), with seven on the C-terminal Y-cluster and one on Y574.

The regulatory role of Y574

In the Etk active site, due to steric hindrance of the Y574 side chain, the ATP cofactor would not be accessible to external

Figure 4 Etk activation through Y574 phosphorylation. (A) Dephosphorylated Y574 blocks substrate access to the active site. The residue, once phosphorylated, is likely to find an alternative conformation, such as binding to R614, therefore making the active site accessible to substrates. (B) The active site of *P. furiosus* MinD complexed with ADP (PDB code 1G3Q; Hayashi *et al*, 2001). (C) Modelled Etk C-terminal peptide (715–722, YNYYGYSY) is docked to the active site of the Etk–ADP complex structure, with the P-Y574 side chain interacting with R614. The magnesium cation, coloured green, is not present in the current Etk structure (due to the presence of EDTA in the crystallization buffer) and hence modelled from the same atom in MinD structure (PDB code 1G3R; Hayashi *et al*, 2001). Y718 of the peptide substrate is positioned to be phosphorylated. The Etk–peptide complex model is energy minimized. (D) The activation loop of insulin-like growth factor-1 receptor kinase (PDB code 1K3A; Favelyukis *et al*, 2001). The phosphorylation of Y1131, Y1135 and Y1136 stabilizes the activation loop far away from the active site (orange), which is now occupied by the ATP analogue (green) and the peptide substrate (magenta). When these three Tyr residues are dephosphorylated (PDB code 1IRK; Hubbard *et al*, 1994), the activation loop occupies the same location as the peptide substrate, inhibiting its kinase activity.

Table 1 Autokinase activity of the Etk kinase domain of wild type and mutants

	Wild-type activity (%)
Wild type	100 ± 2.4
Y574A	98.4 ± 8.0
Y574E	100 ± 12
Y574F	12.1 ± 1.0
Y574G	86.7 ± 2.4
Y574N	26.9 ± 1.9
R614A	14.6 ± 1.4
R614K	61.8 ± 6.0

peptide substrates. Indeed, mutation of this Tyr to Ala and Gly, which should remove this steric hindrance, were found to retain 98 and 87% of kinase activity, respectively (Table 1). As Y574 phosphorylation is known to activate Etk (Grangeasse *et al*, 2002), structures of both apo-Etk and Etk-ADP depict the unphosphorylated or inactive conformation of Etk because of the lack of a phospho-group in Y574. In the inactive structure, the side chain of Y574 points towards the α -phosphate of ADP, instead of the γ -phosphate. Nevertheless the Y574 side chain could readily rotate to approach γ -phosphate, becoming susceptible to phosphorylation.

In mass spectrometric and SDS–PAGE analyses of the intact protein masses, the initial protein samples showed a clear profile of phosphorylation with up to eight phosphate groups per molecule (Supplementary Figures 1 and 3). It is intriguing

that, however, in the crystal structure no phosphotyrosine was observed on the C-terminal Y-cluster and Y574. The high pH (9.5) of the buffer and long incubation time for crystallization (1–2 month) might have led to the loss of most phosphate groups through non-specific hydrolysis, whereas this slow hydrolysis finally ensured a soluble, homogenous, dephosphorylated Etk protein that favoured crystallization.

Were Y574 to be phosphorylated, the extra phosphate would clearly clash with the bound cofactor, ADP. It follows that, in order to vacate the active site for subsequent C-terminal Tyr-cluster phosphorylation, phosphorylated-Y574 (P-Y574) has to adopt an alternate conformation that unblocks the active site. Although P-Y574 was not found in the crystal structure, in both crystallographically independent kinase molecules, strong and unambiguous electron density for a sulphate moiety was observed at the back of Y574 side chain (Figure 2C and D). The sulphate group interacts with a number of conserved residues at this site, making a salt bridge with the side chain of R614 and hydrogen bonds with the side chains of Q616 and N583; hence, the sulphate group is tightly trapped. A 180° rotation of the Y574 side chain would readily place its side-chain hydroxyl group in almost perfect overlapping position with the sulphate oxygen atom. It is plausible that this alternative Y574 side-chain conformation, when coupled with the phosphate group, would represent P-Y574 in its open position. R614, along with Q616 and N583, might have an important function for attracting the phosphate group and the P-Y574 side chain away from the active site to open up the active site for

Figure 5 Effects of time course *in vitro* Tyr autophosphorylation on the C-terminus of Etk. Aliquots of the Etk protein solution (25 μ l) were mixed with 5 mM ATP and 10 mM MgCl₂ and incubated for various time intervals. The protein was then digested with 50 ng endoproteinase Lys-C and the resulting peptides were detected by MALDI QqTOF MS. The bar graph depicts the relative peak intensity of the multiple phosphorylated peptides containing the C-terminus of the protein. (A) Wild type Etk and (B) R614A mutant.

substrate entry. Perhaps not surprisingly then, R614 is well conserved in all Gram-negative PTKs as much as Y574 (Figure 3). It has been reported in two Wzc studies that mutation of the Tyr residue to Phe and Asn, the side chains of which are incapable of being phosphorylated, resulted in significant reduction of enzymatic activity (Grangeasse *et al*, 2002; Soulat *et al*, 2007). The same reduction in kinase activity was observed using Etk Y574F and Y574N mutants in our study (Table I). In addition, the P-Y574 aromatic ring would be stabilized further by interaction with the conserved H576, through π - π stacking (Figure 2C). The P-Y574 conformation, which is accompanied by a small main-chain shift, does not have any steric conflict with other residues. This small main-chain conformational change would be facilitated by the fully conserved G573, a flexible residue preceding Y574. Taken together, the evidence herein demonstrates the importance of Y574 side-chain steric hindrance and its removal for activation.

In order to experimentally verify the possible interaction of P-Y574 and R614, we further generated R614 to Ala and Lys mutants. If the proposed mechanism of phosphorylation-induced conformational switching is correct, the R614A mutant should significantly reduce the activity, whereas the R614K would retain the interaction with P-Y574. Indeed, the R614A mutant only exhibited 15% of wild-type kinase activity (Table I), and the reduced rate of autokinase activity is also demonstrated, on matrix-assisted laser desorption/ionization (MALDI), between the wild type and R614A mutant (Figure 5). The R614K mutant exhibits 62% of the wild-type activity, suggesting that R614 is actually able to interact with P-Y574, but to a lesser extent than R614 does with P-Y574. On the other hand, the Tyr574-to-Glu mutant, which from struc-

tural modelling shows a possible 3.2-Å salt bridge with R614, still retains all (100%) of the wild-type activity. The above results strongly suggest a conformation shift of the Y574 side chain, enabling the Y(E)574-R614 interaction that unblocks the active site, hence allowing subsequent substrates to enter the active site (modelled in Figure 4C).

Using *etk*-knockout *E. coli* cells, which exhibit compromised polymyxin-B resistance (Lacour *et al*, 2006), we performed *in vivo* complementation experiments using various full-length ETK mutants. As shown in Figure 6, when *etk* knockouts were rescued by full-length Etk and its mutants, we observed differential resistance of polymyxin-B. Mutants that resulted in the absence of Y574 steric hindrance (Y574A), always switched-on (Y574E) or wild-type-like (R614K) P-Y574-R614 interaction, showed higher polymyxin-B resistance than those with a compromised P-Y574-R614 switch (Y574F, Y574N, R614A). These *in vivo* results support the conclusion that Y574 blocks the active site and interaction of P-Y574 with R614 is required for activation.

To our knowledge, this model of a PTK activation mechanism—by the side-chain conformational change of a single Tyr residue in the active site—is the first among all other mechanisms of PTK activation. Y574 is, therefore, the molecular switch of Etk (and Gram-negative BY kinases) activation. Interestingly, Y574 is not a conserved residue in Gram-positive BY kinases (Figure 3), where a different kinase activation mechanism is speculated to involve a separate membrane-binding protein, which is a homologue of the N-terminal domain of the Gram-negative BY kinases (Mijakovic *et al*, 2003).

The novelty of this PTK activation mechanism is particularly evident when contrasted with the activation mechanism of eukaryotic PTKs. In eukaryotic RTKs, an entire activation loop must first be displaced by as much as 10 Å (through phosphorylation of specific Tyr residues) before kinase activity can ensue (Cowan-Jacob, 2006; Figure 4D). By comparison, the Etk conformational change following initial phosphorylation is much less profound. Two structural features that exert constraints on active-site conformational changes account for the much simpler PTK activation mechanism observed here. First, Y574 is located at the start of helix F. Any significant main-chain shifts would result in partial or even complete unwinding of helix-F, which we do not detect in the circular dichroism spectra (Supplementary Figure 4). Secondly, the main-chain loop prior to Y574 is rigid (containing hydrogen bonds to strand 3 and R614) and well defined in the electron density map, unlike the activation loop in mammalian kinases. The lack of the main-chain conformational change in Etk would not impede the access of substrate as long as the Y574 side chain is absent. As shown in Figure 4C, the substrate can readily approach the active site and cofactor from the protein surface, even though the active site is enclosed from the sides. The active-site rigidity in Etk provides additional support for our novel PTK activation model. It is the localized nature of the conformational change—the rearrangement of a single side chain—that sets the bacterial PTK activation model apart from existing eukaryotic PTK models.

Phosphorylation of the Y-cluster and its effect on CPS export

During preparation of Etk samples, several intriguing properties of the C-terminal Y-cluster phosphorylation were uncovered.

Figure 6 *E. coli* K12 W3110 *etk*-knockout cells complemented with full-length Etk mutants treated with polymyxin-B. RK→A represents a K478A, R479A, R481A triple mutant. **(A)** Various Etk mutants restored different levels of polymyxin resistance. **(B)** Western blot of the membrane fractions of various mutants prior to polymyxin-B exposure probed with anti-His tag antibody. The recombinant mutants showed similar expression levels and stability as the wild type.

First, when Etk samples were treated with alkaline phosphatase, the protein precipitated immediately and irreversibly (Supplementary Figure 3). This process could be monitored quantitatively to show time- and concentration-dependent aggregation of Etk (Supplementary Figure 5). Freshly purified, soluble Etk sample typically elutes in two fractions from a size-exclusion column. The first peak has a molecular weight of approximately 1 million Dalton, while the second peak has a molecular weight corresponding to a monomer. MALDI MS showed that these two fractions differ in their degree of phosphorylation, specifically on the C-terminal Y-cluster (Supplementary Figure 2). Therefore, lowering the phosphorylation state of the Y-cluster leads to Etk aggregation.

When fully dephosphorylated, the Etk kinase domain was found to aggregate *in vitro*, suggesting that absence of phosphorylation groups on the hydrophobic C-terminal Y-cluster exposes hydrophobic surfaces to the environment. We anticipate that full-length Etk would similarly experience strong hydrophobic interactions in a dephosphorylated state, thereby impairing its otherwise optimal association in the CPS export complex. When fully phosphorylated, the Etk kinase domain clearly favours the monomeric state *in vitro*, possibly because it lacks the N-terminal, periplasmic domain that provides the tetrameric contact previously shown (Collins *et al*, 2006).

We observed a positively-charged Arg- and Lys-rich flexible loop region (residues 478–497) that is in close proximity to the C-terminal Tyr-cluster (Figure 7) when placed in the Wzc tetramer EM structure (Collins *et al*, 2006). This arginine and lysine (RK)-cluster, which features a total of nine Arg and Lys residues in a stretch of 20 residues, is conserved across all Gram-negative Etk/Wzc Tyr kinases (Figure 3). Upon full phosphorylation, it is our conjecture that the resulting concentration of negative charges on the P-Tyr-cluster will provide strong attractive forces to the RK-cluster of an adjacent Etk molecule, affecting the tetrameric conformation of Etk in the corresponding kinase-channel complex, thereby affecting the CPS export machinery. Indeed, the mutant in which the

Figure 7 Structure match of Etk kinase domain to the EM structure (Collins *et al*, 2006). **(A)** The Etk kinase domain structure fits well into the EM tetramer. **(B)** From the bottom view, the RK-cluster is adjacent to the C-terminal Y-cluster of an adjacent molecule.

first three basic residues of the RK-cluster were replaced by Ala (K478A, R479A, R481A) could not restore polymyxin-B resistance in Δetk cells (Figure 6).

Our observation of the correlation between Etk C-terminal phosphorylation and aggregation, as well as potential interaction with the RK-cluster, is consistent with the published evidence, which suggests that an intermediate number of phosphorylation sites on the Tyr-cluster is required for CPS assembly (Drummlsmith and Whitfield, 1999; Wugeditsch *et al*, 2001; Paiment *et al*, 2002). A low degree of Tyr-cluster phosphorylation would result in strong hydrophobic attraction of the Tyr-clusters, thus interrupting oligomerization of the kinase and its subsequent interactions with other members (such as the channel) of the export pathway. On the other hand, upon high degree of Tyr-cluster phosphorylation, the positive-negative charge interaction between the RK-cluster and the Y-cluster would also affect the function of the Etk tetramer. Further studies are required to investigate such possibilities and uncover the CPS export regulatory mechanism.

At this stage, it is also unclear what regulates the phosphorylation-dephosphorylation or the opening and closure of the Y574 switch itself. While there is a conjugate phosphatase for Etk (Etp) and Wzc (Wzb) known to dephosphorylate the kinases, hence having an important function in CPS export regulation (Vincent *et al*, 1999), we observed P-Y574 dephosphorylation during crystallization and MS sample preparation as well, without the presence of Etp. It is possible that P-Y574 is rather unstable, which has the benefit of not letting the kinase to be turned on once and forever.

In conclusion, our structure has provided the first glimpse of a prokaryotic protein Tyr kinase, which contains a completely different fold as compared with its mammalian counterparts. The activation mechanism is also novel, involving only the movement of a single Tyr side chain following phosphorylation, instead of the large conformational change found in eukaryotic PTK activation. Further structural observation and functional studies have provided new insights into how the degree of phosphorylation on the C-terminal Y-cluster phosphorylation would affect the export of CPS.

Materials and methods

Expression and purification of the Etk kinase domain

Etk C-terminal kinase domain (451–726) and Y574A, R614A mutants were subcloned into pET expression vectors, yielding fusion proteins with an N-terminal His₁₀ tag. One-litre cultures of BL21 (DE3) cells carrying plasmid for the respective recombinant protein were grown at 37°C in Terrific Broth (Bioshop, Burlington, Canada) with 100 µg/ml ampicillin. Protein expression was induced with 0.1 mM isopropyl β-D-thiogalactoside at room temperature overnight. Etk substituted with selenomethionine was expressed in the *metA*[−] *E. coli* strain DL41 in LeMaster medium (Hendrickson *et al*, 1990). The lysate was purified by nickel nitrilotriacetate agarose affinity chromatography in 50 mM phosphate buffer (pH 8.5) and 300 mM NaCl. Protein samples were eluted using 150 mM imidazole. After being dialysed into Tris buffer (pH 9.5), the samples were passed through a size-exclusion column (HiLoad 26/60, Superdex 200; GE) on a ΔTKA Explorer fast protein liquid chromatography (FPLC). The low-molecular-weight (monomeric) fraction was further dialysed into 100 mM Tris buffer (pH 9.5) and 300 mM NaCl, and concentrated to ~20 mg/ml for crystallization trials.

Crystallization of the Etk kinase domain

The Etk kinase domain was crystallized with 20 mg/ml of protein in 100 mM Tris, pH 9.5, and 300 mM NaCl, in 140 mM ammonium sulphate (Sigma-Aldrich), 17% PEG 8000 (Fluka), 14% glycerol

(Bioshop), 14 mM ethylenediamine tetraacetic acid (EDTA) (Bioshop) pH 8, and 86 mM Bis-Tris (Bioshop), pH 6.0 or 6.5, with addition of 1 M sodium bromide (Fisher) or 40% acetonitrile (Hampton Research). Very thin plate-shaped crystals grew from precipitation over 4 weeks–2 months, and were hardly reproducible. In fact, only a few diffraction-quality crystals would emerge from thousands of crystallization drops. The crystallization solution sufficed as the cryoprotectant. The ADP-bound form was produced by in-drop soaking with identical buffer solution plus 20 µg/ml of ADP over 48 h. A single-wavelength anomalous dispersion data set up to 2.6-Å resolution at selenium K-absorption edge was collected at the National Synchrotron Light Source, Brookhaven National Laboratory (Upton, NY, USA). Two data sets at 2.5 (native) and 3.0-Å (ADP bound) resolution were collected at the Cornell High Energy Synchrotron Source (Ithaca, NY, USA).

Data collection and structure determination

X-ray diffraction data were processed with HKL2000 (Otwinowski and Minor, 1997). The positions of the heavy atom were determined with SOLVE (Terwilliger and Berendzen, 1999) and phase-calculated by RESOLVE (Terwilliger, 2000). The initial model was traced manually with COOT (Emsley and Cowtan, 2004), and subsequently refined using CNS (Brunger *et al*, 1998), XtalView/Xfit (McRee, 1999) and Phenix (Adams *et al*, 2002).

Kinase activity assay

Kinase activity assay was performed with Etk kinase domain wild-type, Y574A, Y574E, Y574F, Y574G, Y574N, R614A and R614K (20 µM, in 50 mM Tris, pH 8.5) being incubated with 200 µM of ATP and 200 µM of MgCl₂ at room temperature for 5 min. The ATP solution contained 10 µCi of [γ-³²P]. The proteins were precipitated from 10 µl of the reaction mixture with 20% tetrachloric acid (TCA) and washed twice with 500 µl of TCA. The resulting pellet was resuspended with 1 ml of standard Glycyl SDS buffer and mixed with 5.3 ml of ScintiVerse (Fisher). Scintillation readings were taken from both the protein sample and 10 µl of reaction mixture. Specific activity for each reaction was determined as the count ratio between the protein sample and the reaction mixture.

Mass spectrometry

The proteins of native Etk and the R614A mutant were treated by cold acetone (99.9%) precipitation at −20°C to remove the salt buffer solution (300 mM NaCl, 100 mM Tris-HCl). Following immediate wash, the resulting pellets were dissolved in a mixture of the solvents of formic acid/water/methanol (v/v/v, 1:1:2). Intact protein mass determination was performed with an Applied Biosystems/MDS Sciex QStar XL quadrupole time-of-flight (QqTOF) mass spectrometer equipped with a nanospray source designed by MDS Sciex (Concord, Ontario, Canada). The instrument was operated with the ionspray voltage of 3500 V and the declustering potential of 80 V, and data acquisition was performed using Analyst QS 1.1 software.

To further identify the protein phosphorylation sites, an in-solution digestion with endoprotease Lys-C was essentially used to produce large proteolytic peptides. The protein was incubated with 50 ng Lys-C (Roche Diagnostic Corp., Indianapolis, IN, USA) in 25 mM ammonium bicarbonate (pH 7.6) solution at 37°C for 4 h. The digested peptides were subsequently deposited on a MALDI target by mixing with an equal volume of 0.5 µl 2,5-dihydroxybenzoic acid matrix (100 mg/ml in 50% acetonitrile (ACN); Sigma), and the solvent was allowed to evaporate. When dry, the target was loaded into the QStar QqTOF instrument with an oMALDI™ II source and a nitrogen laser operating at 337 nm. After MALDI mass MS peptide mapping, the selected peptides were sequenced by MS/MS using argon as the collision gas to confirm the identity.

Polymyxin-B resistance assay

Polymyxin-B resistance assay was performed on the basis of the microdilution method (Amsterdam, 1991). *E. coli* K-12 W3110 and *etk*-knockout strains and the plasmid of the *etk* full-length construct were kind gifts from the Genome Analysis Project in Japan (<http://ecoli.aist-nara.ac.jp>). Plasmids containing wild type, full-length *etk* and various mutants were transformed into Δ*etk* cells. Overnight Luria-Bertani (LB) broth cultures were diluted to an optical density at 600 nm (OD₆₀₀) of 0.10 with 10 µM isopropyl-β-thiogalactoside (where recombinant Etk expression level is similar to that in

wild-type cells) and 1.2 µg/ml polymyxin-B (Bioshop), for a total volume of 200 µl. The cultures were incubated at 37°C for 16 h, followed by OD₆₀₀ measurement. Western blotting was performed with anti-His (Santa Cruz Biotechnology) and subsequently goat anti-rabbit (Bio-Rad) antibody on the membrane fraction of the overnight LB broth cells prior to polymyxin-B treatment. The membrane fraction was prepared by solubilizing lysis pellets in 2 mg/ml n-dodecyl-β-D-maltoside (Bioshop) at 4°C overnight.

Accession numbers

The structure of Etk kinase domain has been deposited into the Protein Data Bank.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

References

- Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D* **58**: 1948–1954
- Amsterdam D (1991) Susceptibility testing of antimicrobials in liquid media. In: *Antibiotics in Laboratory Medicine*, Lorian V (ed), pp 72–78. Baltimore, MD: Williams and Wilkins
- Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D* **54**: 905–921
- Collins RF, Beis K, Clarke BR, Ford RC, Hulley M, Naismith JH, Whitfield C (2006) Periplasmic protein–protein contacts in the inner membrane protein Wzc form a tetrameric complex required for the assembly of *Escherichia coli* group 1 capsules. *J Biol Chem* **281**: 2144–2150
- Collins RF, Beis K, Dong C, Botting CH, McDonnell C, Ford RC, Clarke BR, Whitfield C, Naismith JH (2007) The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *Proc Natl Acad Sci USA* **104**: 2390–2395
- Cowan-Jacob SW (2006) Structural biology of protein tyrosine kinases. *Cell Mol Life Sci* **63**: 2608–2625
- Cozzzone AJ, Grangeasse C, Doublet P, Duclos B (2004) Protein phosphorylation on tyrosine in bacteria. *Arch Microbiol* **181**: 171–181
- Dong C, Beis K, Nesper J, Brunkan-Lamontagne AL, Clarke BR, Whitfield C, Naismith JH (2006) Wza the translocon for *E. coli* capsular polysaccharides defines a new class of membrane protein. *Nature* **444**: 226–229
- Drummlersmith J, Whitfield C (1999) Gene products required for surface expression of the capsular form of the group 1 K antigen in *Escherichia coli* (O9a:K30). *Mol Microbiol* **31**: 1321–1332
- Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**: 2126–2132
- Favelyukis S, Till JH, Hubbard SR, Miller WT (2001) Structure and autoregulation of the insulin-like growth factor 1 receptor kinase. *Nat Struct Biol* **8**: 1058–1063
- Freestone P, Trinei M, Clarke SC, Nystrom T, Norris V (1998) Tyrosine phosphorylation in *Escherichia coli*. *J Mol Biol* **279**: 1045–1051
- Grangeasse C, Cozzzone AJ, Deutscher J, Mijakovic I (2007) Tyrosine phosphorylation: an emerging regulatory device of bacterial physiology. *Trends Biochem Sci* **32**: 86–94
- Grangeasse C, Doublet P, Cozzzone AJ (2002) Tyrosine phosphorylation of protein kinase Wzc from *Escherichia coli* K12 occurs through a two-step process. *J Biol Chem* **277**: 7127–7135
- Grangeasse C, Obadia B, Mijakovic I, Deutscher J, Cozzzone AJ, Doublet P (2003) Autophosphorylation of the *Escherichia coli* protein kinase Wzc regulates tyrosine phosphorylation of Ugd, a UDP-glucose dehydrogenase. *J Biol Chem* **278**: 39323–39329
- Hayashi I, Oyama T, Morikawa K (2001) Structural and functional studies of MinD ATPase: implications for the molecular recognition of the bacterial cell division apparatus. *EMBO J* **20**: 1819–1828
- Hendrickson WA, Horton JR, LeMaster DM (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J* **9**: 1665–1672
- Holm L, Sander C (1995) Dali: a network tool for protein structure comparison. *Trends Biochem Sci* **20**: 478–480
- Hubbard SR, Wei L, Ellis L, Hendrickson WA (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* **372**: 746–754
- Ilan O, Bloch Y, Frankel G, Ullrich H, Geider K, Rosenshine I (1999) Protein tyrosine kinases in bacterial pathogens are associated with virulence and production of exopolysaccharide. *EMBO J* **18**: 3241–3248
- Klein G, Dartigalongue C, Raina S (2003) Phosphorylation-mediated regulation of heat shock response in *Escherichia coli*. *Mol Microbiol* **48**: 269–285
- Lacour S, Doublet P, Obadia B, Cozzzone AJ, Grangeasse C (2006) A novel role for protein-tyrosine kinase Etk from *Escherichia coli* K-12 related to polymyxin resistance. *Res Microbiol* **157**: 637–641
- Leonard TA, Butler PJ, Lowe J (2005) Bacterial chromosome segregation: structure and DNA binding of the Soj dimer—a conserved biological switch. *EMBO J* **24**: 270–282
- Levitzi A, Gazit A (1995) Tyrosine kinase inhibition: an approach to drug development. *Science* **267**: 1782–1788
- Lutkenhaus J, Sundaramoorthy M (2003) MinD and role of the deviant Walker A motif, dimerization and membrane binding in oscillation. *Mol Microbiol* **48**: 295–303
- McRee DE (1999) XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* **125**: 156–165
- Mijakovic I, Poncet S, Boel G, Maze A, Gillet S, Jamet E, Decottignies P, Grangeasse C, Doublet P, Le MP, Deutscher J (2003) Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases. *EMBO J* **22**: 4709–4718
- Minic Z, Marie C, Delorme C, Faurie JM, Mercier G, Ehrlich D, Renault P (2007) Control of EpsE, the phosphoglycosyltransferase initiating exopolysaccharide synthesis in *Streptococcus thermophilus*, by EpsD tyrosine kinase. *J Bacteriol* **189**: 1351–1357
- Nesper J, Hill CM, Paiment A, Harauz G, Beis K, Naismith JH, Whitfield C (2003) Translocation of group 1 capsular polysaccharide in *Escherichia coli* serotype K30. Structural and functional analysis of the outer membrane lipoprotein Wza. *J Biol Chem* **278**: 49763–49772
- Orskov I, Orskov F, Jann B, Jann K (1977) Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. *Bacteriol Rev* **41**: 667–710
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**: 307–326
- Paiment A, Hocking J, Whitfield C (2002) Impact of phosphorylation of specific residues in the tyrosine autokinase, Wzc, on its activity in assembly of group 1 capsules in *Escherichia coli*. *J Bacteriol* **184**: 6437–6447

- Peleg A, Shifrin Y, Ilan O, Nadler-Yona C, Nov S, Koby S, Baruch K, Altuvia S, Elgrably-Weiss M, Abe CM, Knutton S, Saper MA, Rosenshine I (2005) Identification of an *Escherichia coli* operon required for formation of the O-antigen capsule. *J Bacteriol* **187**: 5259–5266
- Reid AN, Whitfield C (2005) Functional analysis of conserved gene products involved in assembly of *Escherichia coli* capsules and exopolysaccharides: evidence for molecular recognition between Wza and Wzc for colanic acid biosynthesis. *J Bacteriol* **187**: 5470–5481
- Soulat D, Jault JM, Geourjon C, Gouet P, Cozzzone AJ, Grangeasse C (2007) Tyrosine-kinase Wzc from *Escherichia coli* possesses an ATPase activity regulated by autophosphorylation. *FEMS Microbiol Lett* **274**: 252–259
- Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D* **56**: 965–972
- Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D* **55**: 849–861
- Vincent C, Doublet P, Grangeasse C, Vaganay E, Cozzzone AJ, Duclos B (1999) Cells of *Escherichia coli* contain a protein-tyrosine kinase, Wzc, and a phosphotyrosine-protein phosphatase, Wzb. *J Bacteriol* **181**: 3472–3477
- Whitfield C (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem* **75**: 39–68
- Whitfield C, Roberts IS (1999) Structure, assembly and regulation of expression of capsules in *Escherichia coli*. *Mol Microbiol* **31**: 1307–1319
- Wugeditsch T, Paiment A, Hocking J, Drummelsmith J, Forrester C, Whitfield C (2001) Phosphorylation of Wzc, a tyrosine autokinase, is essential for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *J Biol Chem* **276**: 2361–2371