Application of Mistic to improving the expression and membrane integration of histidine kinase receptors from *Escherichia coli*

Georgia Kefala · Witek Kwiatkowski ·
Luis Esquivies · Innokentiy Maslennikov ·
Senyon Choe

Abstract Integral membrane proteins have become the focus of interest of many laboratories and structural genomics consortia, but their study is hampered by bottlenecks in production, solubilization, purification and crystallization. In our laboratory we have addressed the problem of high-level protein expression in the membrane of *Escherichia coli* by use of Mistic, a novel *Bacillus subtilis* protein, as a fusion partner. In this study we examine the effect of Mistic on protein expression and membrane integration levels of members of the *E. coli* histidine kinase receptor family. We find that Mistic fusion invariably increases the overall yield by targeting the cargo proteins more efficiently to the membrane and may even replace the signal sequence. Mistic fusion methods will likely be instrumental for high-level expression of other integral membrane proteins.

Keywords Membrane protein · *Escherichia coli* Protein expression · Mistic · Histidine kinase receptor

Abbreviations

IMP	Integral membrane protein
SP	Signal peptide
HKR	Histidine kinase receptor
PCR	Polymerase chain reaction
PBS	Phosphate buffer saline
M	Mistic fused protein, misticated
NM	Non-mistic fused, non-misticated

G. Kefala · W. Kwiatkowski · L. Esquivies · I. Maslennikov ·
S. Choe (✉)
Structural Biology Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Rd, La Jolla, CA 92037, USA
e-mail: choe@salk.edu

FL	Full length
T	Truncated
MAD	Multi-wavelength anomalous diffraction

Introduction

Integral membrane proteins (IMPs) account for 25–30% of both prokaryotic and eukaryotic proteomes [1, 2] and are involved in numerous cellular processes, such as transport, lipid biosynthesis, cell adhesion and signal transduction. They attract increasing interest from biochemists and structural biologists, in an effort to better understand their functions and potentially use the information for rationalized drug design. Nevertheless, due to bottlenecks associated with recombinant protein expression, extraction from the membrane by means of detergents, purification and crystallization, less than 200 unique structures of membrane proteins or peptides are currently (July 2007) known (Membrane Protein Data Bank, [3]).

For the purpose of functional and structural analysis large amounts of folded, active membrane proteins are typically required. Mistic a novel *B. subtilis* protein discovered recently, [4] can be used to produce large quantities of otherwise low or non-expressing membrane proteins in *Escherichia coli* (*E. coli*). *E. coli* is the organism of choice for expression of recombinant proteins because it is inexpensive, fast and readily scalable to high-throughput production needs. In *E. coli*, the newly synthesized IMPs utilize the *E. coli* translocation machinery to get inserted into the cell membrane and obtain proper folding. However, in the case of recombinant IMPs, the protein translation machinery can easily be overloaded, leading to toxicity.

Mistic, a small 13 kDa protein, associates tightly with the *E. coli* membrane, whether expressed alone or in fusion with IMPs. Without any recognizable signal sequence (also referred as signal peptide, SP), it likely bypasses direct interaction with the *E. coli* translocon complex. Mistic, when fused to the N-terminus of IMPs, enables the cargo proteins to fold into their native conformations in the membrane, thus yielding high-level expression of IMPs [4].

In this study, we have focused on the *E. coli* family of histidine kinase receptors (HKR). HKRs, together with their cytoplasmic cognate response regulators, form the so-called Two Component Systems (TCS), which are the predominant signal transduction mechanisms in prokaryotic organisms [5], also present in fungi and plants [6]. Most HKRs have two predicted transmembrane helices and are comprised of a periplasmic sensor domain and a large cytoplasmic C-terminal portion, divided into at least two subdomains: a dimerization domain containing the conserved histidine, and a catalytic ATP binding domain. In brief, environmental signals trigger the dimeric HKR to catalyze its autophosphorylation on the conserved histidine residue. Subsequently, the HKR catalyzes the transfer of the phosphoryl group from the histidine to a conserved aspartate residue on the cognate response regulator, which undergoes conformational change. This conformational change enables inter- or intra-molecular interactions, through which the response to the environmental signal is achieved [7]. Most response regulators are transcription factors, regulating the transcription of specific genes. Herein, we discuss how Mistic has been used as N-terminal fusion with proteins of the *E. coli* HKR family, to enhance their expression and integration in the bacterial membrane.

Materials and methods

All chemicals are purchased from Sigma, unless otherwise noted. NuPAGE gels are from Invitrogen, protein standards from Invitrogen and BIO-RAD.

Target selection

Our focus is on *E. coli* membrane kinase receptors. Our target selection was initially based on a publication reporting the global topology analysis of the *E. coli* inner membrane proteome [8], and subsequently combined search results from the public *E. coli* databases EchoBASE [9] and Colibri [10]. We identified 25 *E. coli* membrane kinases, 23 of which are transmembrane HKRs of TCS, and two are tyrosine kinase receptors. In this study, we have focussed on seven HKRs. They are BasS, CpxA, EvgS, KdpD, PhoQ, RstB, and ZraS.

Cloning and expression

Two different constructs were prepared for each kinase, one as fusion in frame with Mistic within the Gateway-adapted Mistic-containing octahistidine-tagged vector (pMIS3.0E, [4]), which we refer to as 'Misticated' targets, and one non-Misticated in a Gateway-adapted nonylhistidine-tagged vector (pET28). Both vectors add to the expressed proteins an N-terminal histidine tag, followed by a thrombin cleavage site. The program SignalP v3.0 [11] was used to assess the possibility of the presence of a signal peptide (SP).

The target HKR genes were amplified and Gateway (Invitrogen)—adapted by performing two successive rounds of Polymerase Chain Reaction (PCR). In the first round, the gene of interest was amplified by colony PCR, using a single cell *E. coli* K-12 MG1655 colony as template, and gene-specific primers. After PCR purification (Qiagen), the product of the first PCR served as the template for a second PCR, using a pair of generic primers to introduce the attB sites required for Gateway cloning and an N-terminal thrombin cleavage site.

The final PCR product was purified by gel purification (Qiagen), and used to set up a BP reaction with the pENTR207 vector, following the manufacturer's instructions (Invitrogen). The recombinant plasmids were used to transform *E. coli* BL21 (DE3) cells (Invitrogen). Cells from an overnight 5 ml culture were inoculated in 150 ml Terrific Broth (TB) medium and grown at 37°C to an OD 600 nm of 1. The temperature was then lowered to 18°C and protein production was induced by addition of 0.5 mM β-d-thiogalactopyranoside (IPTG), and overnight incubation.

Membrane isolation

The cells were collected by centrifugation, washed in Phosphate Buffered Saline (PBS), (137 mM NaCl, 10 mM phosphate, 2.7 mM KCl, pH 7.4), resuspended in 30 ml of lysis buffer (20 mM Tris pH 8.0, 10 mM EDTA, 5 mM β-mercaptoethanol (BME), 5 mM PMSF and 0.5 mg/ml lysozyme), and lysed by sonication three times for 1 min each, using 1 s pulses, on ice.

The lysed cells were centrifuged at 100,000*g* for 2 h. The pellet consisting of non-lysed cells, inclusion bodies and membrane fractions was resuspended in 30 ml lysis buffer without lysozyme and centrifuged at 10,000*g* for 20 min, for inclusion body pelleting. Membranes were collected by subsequent centrifugation of the supernatant at 100,000*g* for 2 h. The pelleted membrane fraction was well resuspended in 30 ml of ice cold salt wash buffer (20 mM Tris pH   8.0,   500 mM NaCl,   5 mM BME,

5 mM PMSF 10 mM EDTA) and stirred overnight at 4°C. The next morning, the membrane fractions were pelleted by centrifugation at 100,000g for 2 h, the pellet was weighted and then resuspended in 5 ml ice cold storage buffer consisting of 20 mM Tris pH 8, 0.1 M NaCl, 5 mM PMSF, 5 mM BME and 20% v/v glycerol, aliquoted and stored at −80°C. The amount of whole cell lysate and membrane fraction to be analysed by SDS–PAGE was normalized based on the wet weight of the pelleted cells after overnight induction, and final membrane fraction, respectively.

Total protein concentration in the membrane fraction for the kinase assay is estimated as follows: 100 μl of membrane fraction in storage buffer are centrifuged at 21,000g for 10 min and the pellet is well resuspended in 2% SDS to 100 μl final volume. The sample is then heated at 95°C for 10 min, centrifuged 21,000g for 3 min and the OD 280 nm is measured against 2% SDS as blank. Extinction coefficient and theoretical pI were calculated for each protein using the ExPASy Protparam tool [12].

Western blot

The samples were separated by SDS–PAGE and transferred to nitrocellulose membranes by semi-dry electroblotting. The unoccupied sites on the nitrocellulose membrane were blocked with 5% non-fat milk powder in PBS for 2 h at room temperature, followed by washing 2× with PBS plus 0.02% Tween 20 (PBST) and 1× with PBS. The nitrocellulose membrane was then incubated with Penta-His Antibody, BSA-free, Mouse monoclonal (0.2 mg/ml, Qiagen) at 1:3000 dilution in blocking solution, overnight at 4°C. The next morning, the membrane was washed 2× with PBST and 1× with PBS and then incubated with Goat Anti-Mouse Horse Radish Peroxidase Conjugate (0.8 mg/ml, BIO-RAD) at 1:3000 dilution in PBS for 1 h at room temperature. The membrane was then washed as above and incubated in freshly prepared substrate solution (HRP conjugate kit, BIO-RAD) for 5–10 min, for color development.

Phosphorylation assay

Autophosphorylation assays of the membrane fractions containing the overexpressed kinases were performed in the presence of radioactive ATP. About 1 μg of protein (as estimated by the SDS method described above) was diluted to 1 μl volume with storage buffer and mixed with 7 μl of ultra pure H_2O and 10 μl of 2× buffer: 100 mM Tris pH 8.0, 100 mM KCl, 10 mM MgCl$_2$, 10% glycerol. 2 μl ATP mixture was then added, containing 10 μCi [γ^{32}P]

ATP and 20 μM cold ATP. The reaction was carried out at 25°C for 15 min and was terminated by adding SDS–PAGE sample buffer. The samples were analyzed on a 10% SDS polyacrylamide gel by electrophoresis, the gel was dried and exposed on KODAK BioMax XAR film at −70°C, using Lightning Plus intensifying screens (Dupont).

Results and discussion

Effects of Mistic fusion on membrane expression levels

Each of the seven HKRs was expressed both as a fusion with Mistic (M, Misticated) and without Mistic (NM, non-Misticated) in order to assess the effect of Mistic on the level of protein expression, and the efficiency of integration into the *E. coli* membrane. In all cases the membrane fraction was isolated and analyzed on SDS–PAGE. Care was taken to normalize the amount of sample loaded on the gel. In all seven cases, fusion to Mistic improved the amount of membrane integrated recombinant protein (Fig. 1). In particular, two NM proteins, BasS and EvgS, had no detectable level by Coomassie stain, in contrast to the very high expression of the same proteins as Mistic fusions. CpxA, PhoQ, KdpD and RstB NM are expressed, but at lower yields compared to their Mistic fused counterparts. In the case of ZraS, a very low amount of Mistic fused protein is detected on the Coomassie stained gel, whereas the NM ZraS is barely visible. ZraS is one of the four targets included in this study that has a predicted SP (Table 1), therefore in order to investigate

Fig. 1 Membrane expression of full length Misticated vs. non-Misticated proteins. NuPAGE (4–12%) Bis-Tris SDS gel of the final membrane fractions in storage buffer. In order to normalize the amount of protein, the weight of the membrane fraction was taken into account for calculating the volume to be loaded. Lanes 1 and 12: See Blue plus 2 protein standard (Invitrogen). Lane 2: BasS M, Lane 3: BasS NM, Lane 4: CpxA M, Lane 5: CpxA NM, Lane 6: EvgS M, Lane 7: EvgS NM, Lane 8: PhoQ M, Lane 9: PhoQ NM, Lane 10: ZraS M, Lane 11: ZraS NM, Lane 13: KdpD M, Lane 14: KdpD NM, Lane 15: RstB M, Lane 16: RstB NM

Table 1 Summary of full length target information and results

Protein name	Protein size (kDa) M/NM	Signal peptide prediction	pI (theoritical) M/NM	Expression in the membrane[a] M/NM	Functional assay M/NM
BasS	59.57/44.52	No	6.10/6.60	+++/−	Not done
CpxA	70.30/55.25	No	5.97/6.45	+++/+	+/+
EvgS	153.42/138.37	Yes	5.90/6.03	+++/−	Not done
KdpD	117.40/102.34	No	6.16/6.49	+++/+	Not done
PhoQ	73.97/58.92	Yes	5.88/6.19	+++/++	+/+
RstB	67.83/52.78	Yes	6.00/6.33	++++/+++	Not done
ZraS	69.70/54.65	Yes	6.79/9.18	+/−	+/+

[a] Approximate total protein yield in the membrane fraction of 150 ml *E. coli* culture. Ranges of quantities as calculated by measuring OD 280 as described in "Materials and methods"

++++: 15 mg

+++: 8–12 mg

++: 6 mg

+: 1.5–4 mg

−: 0.5–0.8 mg

possible interference of the putative SP with Mistic for membrane integration, we prepared additional constructs lacking the SP (see next paragraph).

Effects of the SignalP-predicted signal peptide on expression levels

When we fused full length (FL) HKRs in frame to the C terminus of Mistic, ZraS was the only HRK that showed low levels of expression (Fig. 1). In order to investigate possible interference of the predicted SP with Mistic's capability to target the fused protein to the *E. coli* membrane, we removed the predicted SP, producing what we call a Truncated (T) version of the protein. Two such targets we selected among four HKRs with a Signal P-predicted SP (Table 1) are ZraS, which expresses at low levels as FL, and PhoQ, which shows high levels of expression as FL protein, to compare. We thus prepared two truncated constructs, with and without Mistic fusion: T–M and T–NM. We evaluated the expression levels in the isolated membrane fraction by SDS–PAGE and western blotting for all four constructs of ZraS and PhoQ (Fig. 2a, b).

Misticated ZraS FL expressed at low levels (also see Fig. 1), whereas the NM version showed barely detectable expression, even after western blotting with anti-His antibodies. However, when the predicted SP sequence was removed, ZraS expressed at high levels only in fusion with Mistic (Fig. 2). Removal of the predicted SP of PhoQ, which already expresses at high levels as FL–M protein, resulted in equally good expression without the SP sequence in the membrane fraction. Therefore in cases where a SP is predicted, it is advisable to also test the T–M

protein expression if the initial expression test of the full length protein is unsuccessful. It is important to note however, that one must proceed cautiously, since predictions by SignalP could be misleading and instead of a SP, a transmembrane helix might be removed.

Comparison of whole cell lysates and membrane fractions of non-Misticated proteins

In order to investigate whether the lower expression levels of FL NM targets in the *E. coli* membrane are the result of low overall protein expression or of unsuccessful targeting of the protein to the membrane, we compared whole cell lysates (containing all the expressed protein, both in the membrane and in inclusion bodies), and the final membrane fraction. With Misticated proteins, there is insignificant difference between whole cell lysate and membrane fractions in total quantity. However, in all cases of non-Misticated targets, the levels of recombinant protein are significantly higher in the whole cell lysate than in the membrane fraction as detected by Coomassie stain (Fig. 3). RstB and PhoQ are the only two among seven HKRs we tested that expressed at significant level without the help of Mistic (Fig. 1). Low amounts of the non-Misticated proteins in the membrane fraction may be partly due to overloading of the *E. coli* translocation machinery. It has been shown that during overexpression of membrane proteins in *E. coli*, exposed hydrophobic targeting signals can overload the Signal Recognition Particle and could saturate the system. As a result of saturation, membrane integration of further membrane proteins is inhibited, potentially leading to cytotoxicity and cell death [13]. Under the conditions used in this study (described in materials and
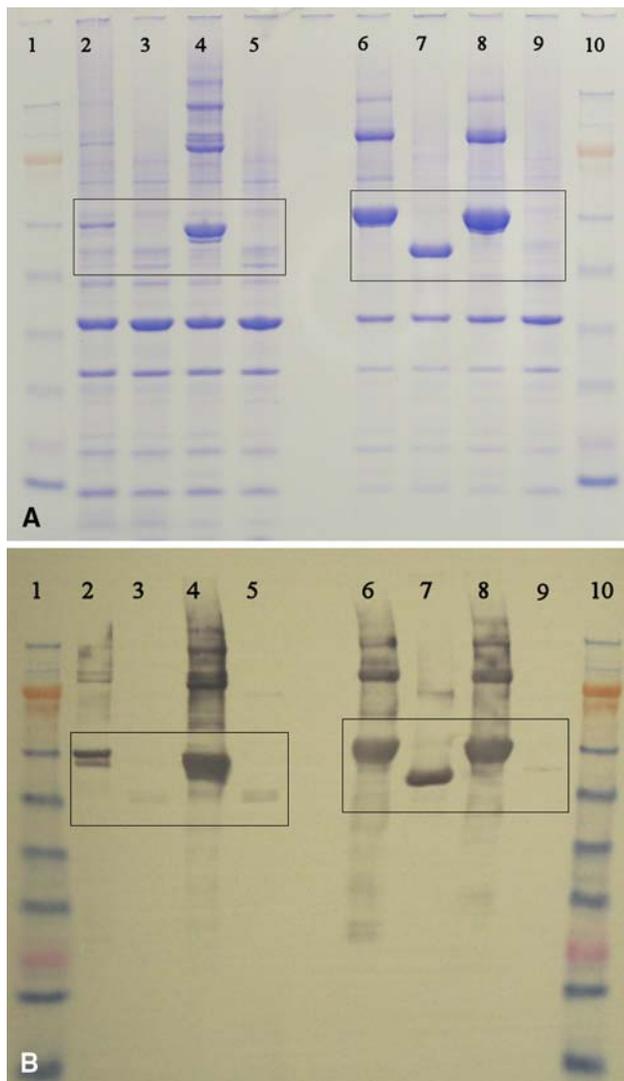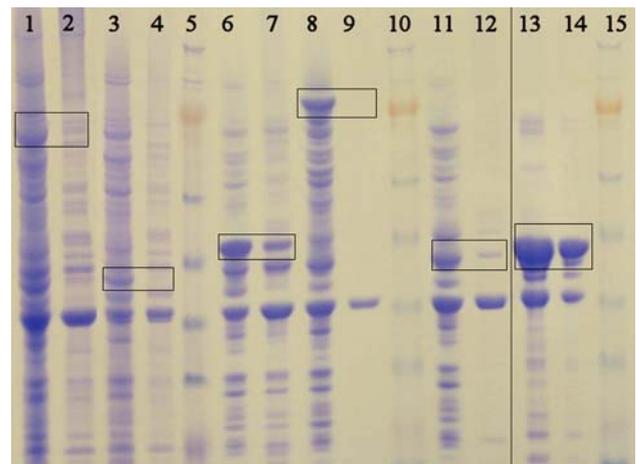
Fig. 3 Comparison of total expression vs. membrane integration for full length (FL) non-Misticated kinases. NuPAGE (4–12%) Bis-Tris SDS gel of the whole cell lysates (wcl) and final membrane fractions (m) in storage buffer for the full length, Non Mistic Fused (NM) proteins. Lanes 5, 10 and 15: See Blue plus 2 protein standard (Invitrogen). Lane 1: KdpD wcl, Lane 2: KdpD m, Lane 3: BasS wcl, Lane 4: BasS m, Lane 6: CpxA wcl, Lane 7: CpxA m, Lane 8: EvgS wcl, Lane 9: EvgS m, Lane 11: ZraS wcl, Lane 12: ZraS m, Lane 13: RstB wcl, Lane 14: RstB m

Fig. 2 Effect of the signal sequence on the expression and membrane integration levels. (**a**) NuPAGE (4–12%) Bis-Tris SDS gel and (**b**) western blot analysis of the final membrane fractions in storage buffer for the full length (FL) and truncated (T) proteins Zras and PhoQ, either as Mistic fusions (M), or without Mistic (NM). Lanes 1 and 10: See Blue plus 2 protein standard (Invitrogen). Lane 2: ZraS FL–M, Lane 3: ZraS FL–NM, Lane 4: ZraS T–M, Lane 5: ZraS T–NM, Lane 6: PhoQ FL–M, Lane 7: PhoQ FL–NM, Lane 8: PhoQ T–M, Lane 9: PhoQ T–NM

methods), we observed no cytotoxicity for either the Misticated targets, or for the non-Misticated ones. All cultures were induced at OD 600 of 1 and at cell harvesting the OD 600 was always in the range of 1.1–1.4. Hydrophobicity variation in combination with protein size may partially account for the differential membrane integration ratio. However a simple comparison of the ratio of hydrophobic (V, I, L, M, F, W and C) versus non-hydrophobic amino acids in the entire sequences of the seven proteins does not show any obvious correlation to the ratios of membrane integration versus inclusion body.

Kinase activity assay

In order to better evaluate the potential of Mistic in overexpressing IMPs, we investigated whether the fusion with Mistic affects the functionality of the protein. The case study group presented here is comprised of HKR, so the assay of choice was autophosphorylation in the presence of radioactive ATP. Whole membrane fractions were used and care was taken so that equal amounts of proteins were tested (see Materials and methods). The assay was performed on a subset of the FL kinase receptors: CpxA, PhoQ and ZraS. Both the Misticated and non-Misticated proteins embedded in the membrane show autophosphorylating activity, at comparable levels between them (Fig. 4). We surmise that fusion with Mistic does not affect the functional and structural properties of the cytoplasmic catalytic domain of the kinase receptors. Functional integrity of transmembrane and extracellular domains of the proteins has not been examined in the present study.

Conclusions

Given the increased interest in structural and functional studies of membrane proteins in the recent years, Mistic can provide a powerful tool for expressing large amounts of otherwise difficult to obtain targets. In this study we examined the effects of Mistic fusion to improve the yields of the family of *E. coli* HKR. We estimate that typical yields from our scaled-up cultures average at 8 mg of pure
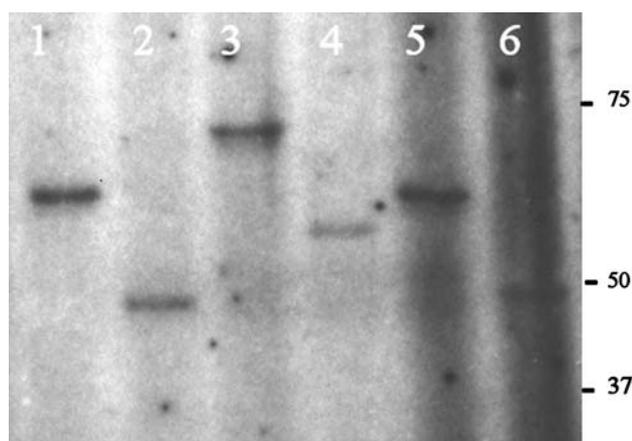
Fig. 4 Kinase activity assay. The membrane fractions containing the overexpressed kinases were analysed on a 10% gel and autoradiography was performed. Lane1: CpxA M, Lane 2: CpxA NM, Lane 3: PhoQ M, Lane 4: PhoQ NM, Lane 5: ZraS M, Lane 6: ZraS NM. The protein standard used is the Precision Plus Protein Standard (BIO-RAD)

protein per liter of culture, with yields ranging from 3 to 35 mg/l of culture in some cases. Under the conditions described herein, we have not encountered cytotoxicity due to the protein overexpression. Further improvements in the use of Mistic to express IMPs in the *E. coli* membrane are currently focused on the study of homologues from other *Bacillus species*, and on the improvement of the existing expression vector. Initial results with Mistic homologues that have recently been identified, indicate integrating ability comparable to the *B. subtilis* Mistic [14]. Comparison between the different Mistic homologues might help to identify the important regions of the sequences for membrane integration and lead to an engineered Mistic sequence optimized for membrane integration capabilities. Modifications of the expression constructs such as different protease cleavage sites or length of the linker between Mistic and the protein of interest are also important variables. Variations of these alternative expression vectors can

be designed to optimize the yields. Mistic offers the advantages of inexpensive and fast high yield production of IMPs in *E. coli* cells. For structural studies by X-ray crystallography, selenomethionine incorporation for MAD can be easily exploited, without requiring major modifications of the protein production procedure.

References

1. Frishman D, Mewes HW (1997) Nat Struct Biol 4:626–628
2. Kihara D, Kanehisa M (2000) Genome Res 10:731–743
3. Raman P, Cherezov V, Caffrey M (2006) Cell Mol Life Sci 63:36–51
4. Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R, Choe S (2005) Science 307:1317–1321
5. Wolanin PM, Thomason PA, Stock JB (2002) *Genome Biol* 3, REVIEWS3013
6. Loomis WF, Shaulsky G, Wang N (1997) J Cell Sci 110:1141–1145
7. Stock AM, Robinson VL, Goudreau PN (2000) Annu Rev Biochem 69:183–215
8. Daley DO, Rapp M, Granseth E, Melen K, Drew D, von Heijne G (2005) Science 308:1321–1323
9. Misra RV, Horler RS, Reindl W, Goryanin II, Thomas GH (2005) Nucleic Acids Res 33:D329–D333
10. Médigue C, Viari A, Hénaut A, Danchin A (1993) Microbiol Rev 57:623–654
11. Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) J Mol Biol 340:783–795
12. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) Nucleic Acids Res 31:3784–3788
13. Valent QA, de Gier JW, von Heijne G, Kendall DA, ten Hagen-Jongman CM, Oudega B, Luirink J (1997) Mol Microbiol 25:53–64
14. Roosild TP, Vega M, Castronovo S, Choe S (2006) BMC Struct Biol 6:10