Original Research Paper Biotechnology

Protein-Protein Interaction (PPI) interface study of bacterial Universal Stress Protein (UspC) and Chromosome Partitioning Protein (mukB): Computational Approaches

Alok Jha

Assistant Professor, Department of Biotechnology and Life Sciences, Institute of Bio Medical Education & Research. Mangalayatan University, Beswan, Aligarh, UP, India. Pin: 202145.

ABSTRACT Computational approaches have attracted tremendous attention for the prediction of protein-protein interaction networks and also because of the ability of such methods to validate the results. In this study, we report protein-protein interaction (PPI) interfaces of bacterial universal stress protein, UspC and chromosome partitioning protein, mukB by using in silico methods. Universal stress protein, UspC has a functional role in providing resistance to DNA damaging agents and the predicted partner chromosome partitioning protein, mukB is involved in chromosome segregation and condensation and associated signal transduction. The study reveals the hot spot residues in the protein-protein interaction interface and also the solvation free energy gain upon formation of the assembly that indicates the areas to be targeted with small molecules. The geometry of PPI interfaces also indicates the possibility of finding new therapeutic interventions.

KEYWORDS : Universal stress protein, UspC, Protein-protein interaction, PPI interface, Hot spots, Chromosome partitioning protein, mukB

Introduction:

The universal stress protein appears to belong to all E.*coli* stress and starvation stimulons under the global control of gene expression.[16] In different words, the production of the protein was found to be stimulated by variety of conditions including stationary phase, starvation for carbon, nitrogen, phosphate, sulphate and amino acids, and exposure to heat, oxidants, metals, uncouplers, polymixin, cycloserine, ethanol, antibiotics and other stimulants.[16] Moreover, recent bioinformatics data suggests about the regulation and possible roles of usp paralogues in E.*coli* as well. It is known that E.*coli* has six usp family genes, uspA, uspC, uspD, uspE, uspF and uspG and it appears that most usp genes are monocistronically expressed with some exceptions such as uspD.[17,18] The corresponding Usp proteins can be divided into two subfamilies on the basis of sequence similarities in the Usp domain.[18] Three of the Usp proteins, UspA, UspC (yecG), and UspD (yiiT), belong to the same related family, whereas, UspF (ynaF) and UspG (ybdQ) belong to the second sub family. The UspE that is a tandem type protein seems to have evolved from a duplication event. The Usp domain 2 of UspE is more related to that of the UspFG sub family, whereas, the sequence of domain 1 appears to be closer to that of the UspA sub family. Usp genes including uspA, uspC, uspD and uspE respond to stress conditions causing growth arrest and under extreme conditions production of Usp proteins are repressed.[10,17,18] A search for sequence similarity in several databases suggests that Usp proteins or conserved Usp domains are present in Eukaryotes; therefore, appear to be of importance in these organisms as well as in bacteria and archae. The bacterial usp genes usually encode either a small Usp protein (around 14-15 kD) that has one Usp domain, or a larger version (around 30 kD) consists of two Usp domains in tandem. The relative number and distribution of these variants differ depending on the organism. Moreover, some proteins of the Usp family in Archae, cyanobacteria and plants are large proteins in which the Usp domain is present together with other functional domains e.g. Na^+/H^+ antiporter domains, Cl- voltage channels, amino acid permeases and protein kinase domains. etc.[16] An E.coli temperature sensitive mutant was isolated which produces spontaneously normal size anucleate cells at low temperature.[19] Another study suggested that mukB mutants of E.coli are defective in the correct partitioning of replicated chromosomes.[20] This results in the appearance of normal sized anucleate (chromosome less) cells during cell proliferation. Based on the nucleotide sequence of the mukB gene, the MukB protein of 177kD was predicted that is filamentous protein with globular domains at the ends, and also having DNA and nucleotide binding abilities. MukB forms a homodimer with a rod-and-hinge structure having a pair of large, C-terminal globular domains at one end and a pair of small, N terminal globular domains at the opposite end; it tends to bend at a middle hinge site of the rod section. MukB was reported to bind DNA and also to ATP and GTP in the presence of Zn^{2+}.[20] The product of mukB gene has features of a myosin-like motor protein and probably has a role in chromosome condensation and/or movement.[20] Other genes mukE and mukF which are in an operon with mukB, are also required for efficient partitioning and their products presumably interact with MukB.[22] Of the ~12 bacterial genomes that have been sequenced, mukB, mukE, and mukF are found only in E. coli and Haemophilus influenzae.[21] It was further confirmed that Ca^{2+} or Mg^{2+} may be required for the formation of complex consisting of mukB, mukE and mukF, and thus may participate in a particular step during chromosome partitioning in E.coli [21,22] Moreover, the substantial analysis of proteins reveals that proteins involved in the same cellular processes are interacting with each other. The study of PPIs is also important to infer the protein function within the cell. The detailed study of PPIs has shown the way to model the functional pathways to exemplify the molecular mechanisms of cellular processes.[1,2] There are several ways to establish the result of two or more proteins interacting with a definite functional objective. Other studies marked the significant properties of PPIs such as PPIs can modify the kinetic properties of enzymes and change the specificity of a protein for its substrate through interaction with different binding partners.[8,9] PPIs can also act as a mechanism to allow for substrate channeling; construct a new binding site for small effector molecules, inactivate or suppress a protein. Uncovering protein-protein interaction information helps in the identification of drug targets.[1-3]

Materials and Methods:

Sequence based prediction approaches:

Predictions of PPIs have been carried out by integrating evidence of known interactions with information regarding sequential homology. We performed sequence based prediction of PPIs within bacteria using BIPS-BIANA Interolog prediction server that is based on the known interactions of the orthologous genes of other organisms (interologs).[4]

Sequence similarity measures: Sequence similarity between proteins relies on basic local alignment search tool (BLAST) alignments.[14] The query protein universal stress protein UspC and chromosome partitioning protein, mukB from E. *coli* were aligned against all sequences with known interactions stored in the BIANA MySQL data base.[4] The threshold of 90% of template coverage has been used to ensure that the prediction is not inferred from local regions of the template interaction. Also, the geometric mean of

individual identities (joint identities) and the geometric mean of individual BLAST E values (joint E values) are considered, i.e. E value ≤ 10^{-5}, Similarity ≥ 30%, and alignment coverage ≥ 60%. [4]

Domain interactions: Based on the hypothesis that universal stress proteins UspC and chromosome partitioning protein, mukB (E.*coli)* have interacting domains. Also, these domains could be interacting domains in the iPfam or the 3DID databases. The BIPS-BIANA server measured the similarity of the target sequences with Pfam domains and HMMER program as a function of the E-value cut off of 10^{-5} in the pFam A database. [4]

Further, we used STRING server version 10.0 as the STRING database aims to provide a critical assessment and integration of protein-protein interactions, including direct (physical) as well as indirect (functional) associations. [5] It has scalable algorithms for transferring interaction information between organisms. For this purpose, hierarchical and self consistent orthology annotations have been introduced for universal stress proteins and other interacting proteins, grouping the proteins into families at various levels of phylogenetic resolution. The STRING server version 10.0 [5] includes a completely redesigned prediction pipeline for inferring protein–protein associations from co-expression data, an API interface for the *R* computing environment and thus statistical analysis for Universal stress proteins, UspC and chromosomal partitioning protein, mukB from *E.coli* were performed.

Phylogenetic Tree:
We generated Phylogenetic tree for UspC and chromosome partitioning protein, mukB *(*E.*coli)* by using Clustal Omega. [15] The underlying principle behind this method is that the coevolution between the interacting proteins can be reflected from the degree of similarity from the distance matrices of corresponding phylogenetic trees of the interacting proteins. The set of organisms common to the two proteins are selected from the multiple sequence alignments (MSA) and the results are used to construct the corresponding distance matrix for each protein. The BLAST score [14] could also be used to fill the matrices. The linear correlation is calculated among these distance matrices.

Model Generation:
We retrieved the sequence of chromosome partitioning protein, mukB from Uniprot (sequence id UniProtKB-P22523; MUKB_ECOLI) and generated model using the server Phyre2 Intensive method for model building. 91% of residues were modeled at >90% confidence.[11]

PPI Identification:
We used KFC2 server to identify the protein-protein interaction interface.[6] The KFC model is a machine learning approach for predicting binding hot spots within protein-protein interactions. The KFC model is comprised of two decision tree based classifiers: K-FADE (based on shape specificity features calculated by the Fast Atomic Density Evaluator or FADE) or K-CON (based on biochemical contact features). Each decision tree that provides a set of hierarchical rules for hot spot classification is trained by a supervised learning process to recognize the local structural environments that are indicative of hot spots. [7] Those residues were classified as hot spots if their mutation to alanine resulted in a change of binding energy (ΔΔG) greater than 2 kcal/mol.K-FADE predicts hot spots using the size of the residue and the radial distribution of shape specificity and interface points. K-CON predicts hot spots in terms of a residue's intermolecular atomic contacts, hydrogen bonds, interface points and chemical type. [6]

PPI analyses:
We analyzed Universal stress proteins, UspC and chromosome partitioning protein, mukB (E. *coli*) complexes for protein-protein interaction (PPI) interface by using PDBePISA [12] server. The oligomeric state and symmetry/space group are calculated. Each assembly was assigned a complexation significance score (CSS) that gives information about the importance of the interface in complex formation.

Results and discussion:
The protein-protein interactions for universal stress protein, UspC (E.*coli*) and chromosome partitioning protein, mukB were established by a sequence based approach and a detection method called two hybrid pooling approach with assumption that homologous proteins would have similar behavior. The approach uses sequence similarity between proteins based on the sequence alignment. The alignment of UspC with predicted partner protein mukB sequences is based on 60% identity and 70% of the total length of the target protein and 90% of the template, where, template coverage is fixed to 90%. In a second approach the similarity of the target sequence with Pfam domains as a function of e-value was calculated. This results in the assignation of one or several Pfam domains to the query and target sequences. Then the Interolog prediction server BIPS-BIANA based on interolog information compared the iPfam and 3DiD databases for domain-domain interactions. [4] The homology conditions were maintained with a joint e-value of 1e-10 and joint identity of 80%. The predictions were filtered assuming the traditional definition of interologs that both proteins suppose to interact if they are orthologous with proteins that interact. Both proteins were clustered by comparing with cluster of orthologous genes (COG) database and selected using GO annotations as they share the largest number of similar GO terms.[4] The UspC is involved in providing resistance to DNA damaging agents, whereas, mukB is involved in chromosome segregation, chromosome condensation, signal transduction, and share some molecular functions including protein or transcription factor binding.

The possible interaction of UspC and mukB was predicted from PPI network of UspC (bitscore 285 and e-value 1e-101) on the basis of active interaction sources e.g. co-expression, co-occurrence, gene fusion, neighborhood, experiments and databases by using a deterministic network drawing algorithm (STRING Server) that uses a spring model to generate the network images.[5] Nodes are modeled as masses and edges as springs; the final position of the nodes in image is computed by minimizing the 'energy' of the system. The network nodes are proteins and the edges represent the predicted functional associations. The high confidence edges are given a higher 'spring strength' so that they will reach an optimal position before lower confidence edges. The predicted functional partners include chromosome partitioning proteins MukBEF (mukB, mukE and mukF with scores 0.773, 0.771 and 0.758 respectively), stress induced protein uspE (score 0.770), DNA binding response regulator in two component regulatory system kdpE(score 0.725) and fused sensory histidine kinase kdpD (score 0.715), septal ring assembly factor yiiU (score 0.764) that stimulates cell division. Based on these interactions we selected chromosome partitioning protein mukB for further analysis.

The UspC interaction shows 11 nodes, 25 edges, with average node degree 4.55, and clustering coefficient of 0.876.(Fig.1.) The PPI enrichment p-value is 0.000292 and the functional enrichments in the network shows assigned GO annotations for different biological processes with GO accession id GO:0030261, chromosome condensation (3 genes) with a false discovery rate of 0.000391; GO:0007059; chromosome segregation (3 genes) with a false discovery rate of 0.00586; GO:0007049, cell cycle (4 genes) with a false discovery rate of 0.0142; GO:0051301, cell division (4 genes) with a false discovery rate of 0.0142. The cellular component has been assigned a GO accession GO: 0009295, nucleoid (3 genes) with a false discovery rate of 0.00216. In confidence mode the thickness of the line indicate the degree of confidence prediction of the interaction. Action mode shows additional information about the binding, activation etc. [5]

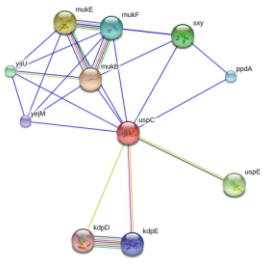

Fig 1. Protein-protein interaction network for UspC (E.*coli*) In evidence mode the edge lines predict the associations with different colored lines that are Red line- indicates the presence of fusion evidence, Green line-neighborhood evidence, Blue line-cooccurrence evidence, Purple line-experimental evidence, Yellow line-textmining evidence, Light blue line- database evidence, Black line-coexpression evidence.

Further, phylogenetic tree was generated on the basis of multiple sequence alignment of universal stress proteins UspC and Chromosome partitioning protein mukB from E.coli using the Clustal W and Clustal Omega parameters. [15] The phylogenetic tree suggested a common evolutionary significance of both the protein with a distance of 0.69718. (Fig.2) The phylogenic tree generation method includes tree format (distance matrix) and clustering method (Neighbor Joining) with distance correlation for more divergent sequences. The percent identity matrix was used with exclusion of gaps that is meant to forcing the alignment to use only positions where information can be included from all sequences.

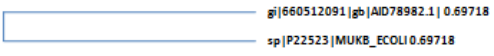

Fig. 2. Phylogenetic tree of universal stress protein UspC and chromosome partitioning protein mukB from E. *coli*

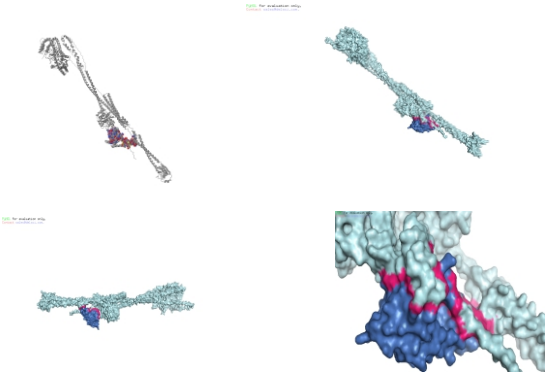

Fig.3. (a) Hot Spots for UspC-mukB complex (E.*coli*) (b,c) UspC-mukB complex with FADE (d) PPI Interface for UspC-mukB complex

We identified the PPI interface and hot spots residues that participate in the interaction.[6,7] Predicted hot spot residues based on K-FADE (Fast Atomic Density Evaluator) and K-CON (biochemical contact features) scores include Glu621 (K-FADE=0.03; K-CON=-0.51), Ala1413 (K-FADE=0.29; K-CON=-0.67), Ile1438 (K-FADE=0.81; K-CON=-0.19), Lys1451 (K-FADE=0.67; K-CON=-0.02), Pro1469 (K-FADE= 1.31; K-CON= -0.52), Leu1471 (K-FADE= 1.22; K-CON= -0.14), Pro1472 (K-FADE= 0.71; K-CON= -0.53), Glu1473 (K-FADE= 0.98; K-CON= -0.36), Leu1475 (K-FADE= 0.35; K-CON=-0.52) in chain A. The Hot spot residues in chain B include Met45 (K-FADE= 0.34; K-CON= -0.15), Ala50 (K-FADE= 0.15; K-CON= -0.63), Leu57 (K-FADE= 0.25; K-CON= -0.07), Ser91 (K-FADE= 0.32; K-CON=-0.42), Asn110 (K-FADE=0.11; K-CON=-0.37), Phe115 (K-FADE= 0.08; K-CON= -0.52), Ala119 (K-FADE=0.73; K-CON= -0.47), Cys121 (K-FADE= 0.33; K-CON= -0.45), Ser122 (K-FADE= 1.00; K-CON= -0.36), Lys124 (K-FADE= 0.41; K-CON= -0.21), Arg125 (K-FADE= 0.67; K-CON= -0.01), Ile127 (K-FADE= 1.32; K-CON=0.10) and Thr128 (K-FADE= 0.74; K-CON= -0.54).[6,7] The K-FADE and K-CON Conf or confidence values indicate for the worst (0) and the best (1) hotspot residues. Therefore, some best hotspot residues are Pro1469, Leu1471 and Ser122, Ile127. The hot spot residues and the PPI interfaces for UspC-mukB are depicted in Fig. 3. The PPI analyses from PDBePISA server [12] suggested that the UspC and chromosome partitioning protein mukB complex has assemblies with multimeric state of 2 with formula AB, composition AB and the dissociation pattern is A+B. The surface area of the complex that indicates the total solvent-accessible surface area of the assembly in $Å^2$ is 95893.0 and the buried area that indicates the total solvent accessible area of the assembly, buried upon formation of all assembly's interfaces in $Å^2$ is 3046.2. ΔG^{int} that indicates the solvation free energy gain upon formation of the assembly in kcal/mol. is -20.0 kcal/mol. The value of ΔG^{diss}, which indicates the free energy of assembly dissociation in kcal/mol is 13.1 kcal/mol.[12] The free energy of dissociation corresponds to the free energy difference between dissociated and associated states. Positive values of ΔG^{diss} indicate that an external driving force should be applied in order to dissociate the assembly, therefore, the assemblies with $\Delta G^{diss} > 0$ are thermodynamically stable. The rigid body entropy change at dissociation $T\Delta S^{diss}$ in kcal/mol is 13.1 kcal/mol. The symmetry number that indicates the number of different but equivalent orientations of the assembly, which can be obtained by rotation, its value is 1 in the UspC-mukB complex. The engaged interfaces correspond to interfacing structures A+B with N_{occ} 1 that indicates the total number of the corresponding interfaces in the assembly. Also, the buried area is 1523.1 $Å^2$ (50%) that indicates the interface area calculated as difference in total accessible surface areas of isolated and interfacing structures divided by two. The solvation free energy gain upon formation of interface Δ^iG is -20.0 kcal/mol. The value is calculated as difference in total salvation energies of isolated and interfacing structures. Negative Δ^iG corresponds to hydrophobic interfaces, or positive protein affinity. This value does not include the effect of satisfied hydrogen bonds and salt bridges across the interface. The surface of the UspC-mukB complex contains 7603 atoms (63.6%) and 645 atoms (61.3%) respectively and the interface contains 163 atoms (1.4%) and 161 atoms (15.3%) respectively. The complex surface contains 1402 residues (94.3%) and 126 residues (93.3%) respectively and the interface contains 43 residues (2.9%) and 45 residues (33.3%) respectively. (12) The interface solvent accessible area in Å are 1576.4 (1.7%) and 1469.8 (18.7%) respectively. The solvation energies of the isolated structures in the UspC-mukB complex in kcal/mol are -968.6 and -126.2 respectively. The gain on complex formation in kcal/mol are -9.8 (1.0%) and -10.2 (8.1%) respectively and the average gain accounts for -4.5 kcal/mol (0.5%) and -5.5 kcal/mol (4.3%) respectively. The P-values for the UspC-mukB complex are assigned as 0.109 and 0.140 respectively. The interface residues involved in hydrogen bond formation in the UspC-mukB complex include Asn612 with Tyr46 (3.48 Å), Ser615 and Glu620 with Ala50 (3.68 and 3.47 Å), Glu620 with Ala51 (3.42 Å), Glu1436 with Val134 (3.87 Å), Ser1439 with Ile127 (3.41 Å) and Glu1441 with Thr128, Ser129 and Arg125 (3.47, 3.25 and 2.21 Å), Val1452 with His111 (2.38 Å), Gln1454 with Glu15 (3.07 Å), Gln1470 with Cys121 (3.63 and 3.18 Å).The interface residues involved in salt bridge formation include Glu1441 with Arg125 (3.24 and 2.21 Å). The interface in complex Formation Significance Score (CSS) assigned is zero. CSS ranges from 0 to 1 as interface relevance to complex formation increases. However, achieved CSS implies that the interface does not play any role in complex formation and seems to be a result of crystal packing only.[12] The energetic signatures in the isolated proteins are retained in the bound forms that can help determining the binding orientation of proteins on complex formation. The analysis of such energetic motifs could help identify binding sites and their orientations from the monomeric, isolated partners for a diverse ensemble of protein-protein assemblies. The

change in solvent accessible surface area (SASA) upon binding and along with an estimate of its contribution to the binding free energy, the identification of various residues in PPIs is useful not only to provide insights into mechanism of protein-protein recognition but also to indicate the areas to be targeted with small molecules.[13] The ability of small molecules to modulate specific PPIs can be pursued for designing small molecules that bind with high affinity to relatively large and flat protein-protein interfaces. However, there are other sites that contrary to hot spots have an explicit concave/convex geometry appealing for therapeutic intervention.

Conclusion:

Protein-protein interaction plays a central role in cellular function, improving the understanding of complex formation has various implications, including the rational design of new therapeutic agents and the mechanisms governing signal transduction networks. Usually, large, flat and relatively featureless binding sites of protein complexes pose many challenges for drug design. An integrated approach using molecular docking and coevolutionary analysis can face such challenges. This approach can accurately predict and characterize the binding sites for protein-protein interactions as well as provide clusters of bound small molecules on the druggable regions of the predicted binding site. These bound small molecules, peptides or peptidomimetics can be chemically combined to create candidate drugs.

References:

1. Rao, V. S., Srinivas, K., Sujini, G. N., Sunand Kumar, G. N. (2014) Protein-Protein Interaction Detection: Methods and Analysis. International Journal of Proteomics, 1-12.
2. Zhang, M., Su, S., Bhatnagar, R.K., Hassett, D. J., Lu. L. J. (2012) Prediction and Analysis of the Protein Interactome in Pseudomonas aeruginosa to Enable Network-Based Drug Target Selection. 7, 7: 1-13.
3. Tuncbag, N., Kar, G., Keskin, O., Gursoy, A., Nussinov, R. (2009) A survey of available tools and web servers for analysis of protein-protein interactions and interfaces. Briefings in Bioinformatics. 1-16.
4. Garcia, J.G., Schleker, S., Seetharaman, J. K., Oliva, B. (2012) BIPS: BIANA Interolog Prediction Server. A tool for protein–protein interaction inference. Nucleic Acids Res. 40, 147-151.
5. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D. Cepas, J.H., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., Kuhn, M., Bork, P., Jensen, L.J., Mering, C.V. (2015) STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 43:447-452.
6. Darnell, S.J., LeGault, L. Mitchell, J.C. (2008) KFC Server: interactive forecasting of protein interaction hot spots. Nucleic Acids Research. 36: 265-269.
7. Kortemme, T., Baker, D. (2002) A simple physical model for binding energy hot spots in protein–protein complexes. Proceedings of the National Academy of Science. 99,22:14116-14121.
8. Valencia, A., Pazos. F. (2002) Computational methods for the prediction of protein interactions. Current Opinion in Structural Biology.12:368–373.
9. Tkaczuk K.L., Shumilin I.A., Chruszcz M., Evdokimova E., Savchenko A., Minor W. (2013) Structural and functional insight into the universal stress protein family. Evol Appl. 6, 3: 434–49.
10. Siegele D.A. (2005) Universal stress proteins in Escherichia coli. J Bacteriol. 187, 18:6253–4.
11. Kelly LA et al. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols. 10, 845-858.
12. Krissinel, E., Henrick, K. (2007). 'Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774--797.
13. Laraia, L., McKenzie, G., Spring, D R., Venkitaraman, A R., Huggins, D J. (2015) Overcoming Chemical, Biological, and Computational Challenges in the Development of Inhibitors Targeting Protein-Protein Interactions. Chem Biol. 18; 22, 6: 689–703.
14. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402.
15. The EMBL-EBI bioinformatics web and programmatic tools framework. (2015) Nucleic acids research. 4:580-4.
16. Kvint, K., Nachin, L., Diez, A., Nystro¨m, T. (2003) The bacterial universal stress protein: function and regulation. Current Opinion in Microbiology. 6:140–145.
17. Nystro¨m T, Neidhardt FC. (1992) Cloning, mapping and nucleotide sequencing of a gene encoding a universal stress protein in Escherichia coli. Mol Microbiol. 6:3187-3198.
18. Gustavsson N, Diez A, Nystro¨m T. (2002) The universal stress protein paralogues of Escherichia coli are co-ordinately regulated and co-operate in the defence against DNA damage. Mol Microbiol. 43:107-117.
19. Niki, H., Jaffe, A., Imamura, R., Ogura, T., Hiraga, S. (1991) The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E.coli. The EMBO Journal vol. 10, 1: 183 – 193.
20. Niki, H., Imamura, R., Kitaokal, M., Yamanaka, K., Ogura, T., Hiraga,S. (1992) E.coli MukB protein involved in chromosome partition forms a homodimer with a rod-and-hinge structure having DNA binding and ATP/GTP binding activities. The EMBO Journal. 11, 13: 5101 - 5109.
21. Yamazoe, M., Onogi, T., Sunako, Y., Niki, H., Yamanaka, K., Ichimura, T., Hiraga, S. (1999) Complex formation of MukB, MukE and MukF proteins involved in chromosome partitioning in Escherichia coli. EMBO J.21:5873-84.
22. Yamanaka, K., Ogura, T., Niki, H., Hiraga, S. (1996). Identification of two new genes, mukE and mukF, involved in chromosome partitioning in Escherichia coli. Mol. Gen. Genet. 250, 241-251.