ADAPTIVE RESPONSE OF *ESCHERICHIA COLI* TO ALKYLATING AGENTS: MOLECULAR ASPECTS AND BIOTECHNOLOGICAL APPLICATIONS IN THE BIOREMEDIATION FIELD

Valentina Rippa

Dottorato in Scienze Biotecnologiche – XXII ciclo
Indirizzo Biotecnologie Industriali
Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche – XXII ciclo
Indirizzo Biotecnologie Industriali
Università di Napoli Federico II

ADAPTIVE RESPONSE OF *ESCHERICHIA COLI* TO ALKYLATING AGENTS: MOLECULAR ASPECTS AND BIOTECHNOLOGICAL APPLICATIONS IN THE BIOREMEDIATION FIELD

Valentina Rippa

Dottoranda:	Valentina Rippa
Relatore:	Dott.ssa Angela Duilio
Coordinatore:	Prof. Ettore Benedetti

Learn from yesterday,
live for today,
hope for tomorrow.
The important thing is not to stop questioning.

Albert Einstein

INDEX

SUMMARY

The increasingly stringent environmental regulations on hazardous wastes has encouraged the search for innovative solutions for the remediation of contaminated wastes. In this field, bioremediation is seen as an attractive solution due to its reputation as a low cost, environmentally friendly and publicly acceptable treatment technology. The aim of this research project was to explore new potential candidates for the bio-treatment of wastes and environments contaminated by alkylating agents. The study has been specifically focused on AidB, an enigmatic component of the response to alkylation stress in bacterium *Escherichia coli*. First, AidB protein was functionally characterized: it was showed to bind with high affinity DNA regions containing an upstream element and to have transcriptional activity. At this regard, it was intriguing to speculate that AidB might stimulate the transcription of genes whose products are responsible for alkylation resistance. Successively, given that the knowledge of the domain architecture is necessary for understanding the multifunctional properties of a protein, structural and functional characterization of domains present in AidB was performed. Specifically, its N-terminal region was shown to be exhibit acyl-CoA dehydrogenase activity while the short C-terminal domain was shown to be responsible for the DNA binding activity and for regulatory function. The study was then aimed at investigate the mechanism by which AidB directly protects *E. coli* cells against alkylating compounds. It was demonstrated that this protein prevents alkylation damage and it does so by protecting DNA and, presumably, by inactivating alkylators before they are able to react with their target. Interestingly, a recent report on the three dimensional structure of AidB bound to double strand DNA supported this model, revealing that the protein is well equipped to sterically occlude DNA from attack by damaging agents. Importantly, the unique chemical environment of FAD active site provided a rationale for a possible role of AidB in deactivation of nitrosoguanidines such as *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) and *N*-ethyl-*N'*-nitro-*N*-nitrosoguanidine (ENNG). Coupled with structural analysis, the results obtained in this work supported the hypothesis that AidB might act as a detoxification enzyme to destroy nitrosoguanidines: indeed, it was demonstrated that *aidB* mutant cells display decreased resistance to MNNG and ENNG and no change in sensitivity to other classes of alkylators; besides, AidB was showed to allow more efficient gene transcription in *E. coli* cells exposed to nitrosoguanidines rather than to other mutagens. Therefore, AidB represents a promising tool for the bio-treatment of sites contaminated by certain alkylating agents. On the basis of data described above, this experimental work was ultimately targeted at identify as well as at characterize *E. coli* AidB homologues in bacteria used for bioremediation applications. Specifically, the acyl-CoA dehydrogenase coded by the PP4780 gene from *Pseudomonas putida* KT2440 was the object of this investigation. The PP4780 gene was expressed in Δ*aidB* and wild type *E. coli* strains and its involvement in the protection against alkylating agents was tested. Interestingly, the complementation of the Δ*aidB* mutation by PP4780 restored the resistance phenotype to lethal and mutagenic effects of MNNG and ENNG; besides, recombinant cells that overexpress PP4780 were shown to possess increased resistance to nitrosoguanidines as compared with wild type and *aidB*-overexpressing cells. On the basis of these observations, the acyl-CoA dehydrogenase from *P. putida* has been demonstrated to be involved in the response to alkylation stress, presumably functioning as a detoxification enzyme. In conclusion, the data obtained strongly support the possibility of developing new successful strategies for the bioremediation of sites contaminated by alkylating compounds.

RIASSUNTO
L'impiego intensivo di composti alchilanti in diversi settori dell'attività umana (industria farmaceutica, chimica, agricola, alimentare) pone seri problemi connessi all'impatto che dette sostanze, altamente inquinanti e dotate di un elevato potere mutagenico e citotossico, hanno sull'ambiente e sulla salute degli esseri viventi. Di qui nasce l'interesse verso lo sviluppo di processi di biorisanamento economicamente sostenibili e a ridotto impatto ambientale da applicare in alternativa ai convenzionali processi di smaltimento basati sul trattamento chimico-fisico. Il biorisanamento, nello specifico, sfrutta sistemi ossidativi -microbici ed enzimatici- per la rottura e successiva trasformazione (ed eventuale completa mineralizzazione) delle molecole inquinanti.
In risposta all'esposizione a composti alchilanti, numerosi microrganismi, in particolare i batteri, hanno sviluppato sistemi deputati a riparare i danni da alchilazione del DNA nonché attività enzimatiche capaci di catabolizzare, completamente o in parte, questi substrati. Tra i suddetti meccanismi, il sistema Ada-dipendente del batterio Gram negativo *Escherichia coli* riveste un interesse particolare. Nello specifico, la cellula batterica risponde alla condizione di stress ambientale inducendo la sintesi e l'accumulo di 4 proteine: Ada, AlkB, AlkA, capaci di riparare il DNA alchilato, e AidB, una flavoproteina non presente in batteri strettamente correlati ad *E. coli* (*Vibrio*, *Shewanella*, *Klebsiella*), omologa ad acil-CoA deidrogenasi (ACADs) e capace di legare il DNA senza specificità di sequenza (1,2,3). La funzione del fattore proteico AidB non è nota sebbene diversi studi abbiano evidenziato il suo coinvolgimento nella risposta ai danni da alchilazione. Cellule di *E. coli* sovraproducenti AidB risultano, infatti, meno sensibili all'azione di molecole alchilanti quali N-metil-N'-nitrosoguanidina (MNNG), ma il meccanismo responsabile dell'insorgenza di questa resistenza è tuttora sconosciuto (4). Recentemente, la risoluzione della struttura tridimensionale di AidB (5) ha fornito informazioni utili a delucidarne la funzione biologica. Tali studi strutturali hanno messo in luce che AidB è un tetramero di subunità identiche, ognuna formata da 3 domini che nel loro insieme presentano il *fold* tridimensionale tipico delle ACADs, e da un dominio C-terminale non riscontrato tra i membri di questa famiglia enzimatica. Nel tetramero le regioni C-terminali, ricche di amminoacidi carichi positivamente e quindi potenzialmente in grado di interagire con lo scheletro del DNA carico negativamente, risultano situate in prossimità della superficie della molecola e quindi lontane dal sito catalitico legante il coenzima FAD e le molecole di substrato. Queste osservazioni strutturali supportano l'ipotesi che l'interazione di AidB con il DNA sia finalizzata alla protezione dell'acido nucleico dall'azione dannosa di agenti mutageni. Gli studi cristallografici hanno inoltre dimostrato che il sito catalitico della proteina presenta caratteristiche strutturali non riscontrate negli omologhi ACADs. Sulla base di queste osservazioni è stato ipotizzato che substrati quali acil-CoA tioesteri a catena ramificata non sono in grado di accedere alla tasca di legame del FAD, (giustificando la debole attività acil-CoA deidrogenasica di AidB) e che la proteina possa essere capace di legare alcune classi di molecole alchilanti e catalizzarne la degradazione. Al di là dei suddetti ruoli biologici, esperimenti di proteomica funzionale condotti nel corso di questo progetto di tesi, hanno indotto ad ipotizzare che AidB abbia proprietà di regolazione dell'espressione genica. Nello specifico, AidB è stata identificata come proteina di *E. coli* risultante in grado di legare elementi *upstream* (UP): sequenze nucleotidiche, poste a monte della *box* -35 di varie regioni promotrici, riconosciute e legate da fattori proteici il cui ruolo è quello di attivare o reprimere la trascrizione (6). AidB, in qualità di regolatore trascrizionale, potrebbe

intervenire nel meccanismo di difesa della cellula batterica, inducendo la produzione di attività enzimatiche con proprietà detossificanti. Le potenzialità esibite da AidB risultano interessanti alla luce della prospettiva di realizzare organismi ingegnerizzati capaci di biodegradare sostanze alchilanti.
Questo progetto di tesi si è proposto di definire la funzione biologica della proteina AidB, attraverso le seguenti tappe:

- Caratterizzazione funzionale e strutturale dei domini presenti in AidB
- Caratterizzazione dell'attività trascrizionale
- Analisi funzionale di AidB nella risposta a molecole alchilanti
- Ricerca di proteine omologhe ad *Ec*AidB in batteri impiegati nel campo del biorisanamento

Caratterizzazione funzionale e strutturale dei domini presenti in AidB.
Le informazioni ricavate dagli allineamenti multipli di sequenze e dalla risoluzione della struttura cristallografica hanno suggerito che AidB è caratterizzata da 2 domini strutturalmente e funzionalmente indipendenti: un dominio di interazione col DNA corrispondente alla breve regione carbossi-terminale della proteina e una regione avente attività deidrogenasica corrispondente alla porzione polipeptidica contenente i 3 domini che nel loro insieme presentano il *fold* tipico delle ACADs. Al fine di caratterizzare i suddetti domini, è stata effettuata la produzione ricombinante della proteina intera e di due mutanti di delezione: AidBCt (il frammento carbossi-terminale) e AidBΔCt (la porzione proteica deleta della regione C-terminale).
Caretterizzazione del dominio di legame al DNA.
I due mutanti e la proteina AidB, espressi in fusione con una coda di istidine, sono stati purificati e saggiati *in vitro* per la capacità di legare il DNA attraverso esperimenti di ritardo della mobilità elettroforetica (EMSA). E' stata confermata la presenza di un dominio di interazione col DNA nella regione C-terminale della proteina. Esperimenti di *gel filtration* hanno inoltre evidenziato la struttura monomerica di tale regione.
Caratterizzazione del dominio catalitico.
L'attività ossido-reduttasica delle proteine AidBΔCt e AidBCt purificate è stata misurata e confrontata con quella di AidB (campione di riferimento), registrando spettrofotometricamente l'ossidazione del substrato isovaleril-CoA. Solo il mutante AidBΔCt è risultato in grado di catalizzare la deidrogenazione dell'isovaleril-CoA mostrando valori di attività specifica analoghi a quelli della proteina intera. L'attività catalitica risiede, quindi, nella regione N-terminale di AidB. Mediante *gel filtration* è stato dimostrato che AidBΔCt è il dominio deputato alla tetramerizzazione.

Caratterizzazione dell'attività trascrizionale.
Allo scopo di indagare l'attività trascrizionale di AidB, si è proceduto all'identificazione di regioni promotrici legate dalla proteina in maniera sequenza-specifica. Tra queste, mediante esperimenti EMSA, è stato identificato il promotore del gene *aidB* (P*aidB*) caratterizzato dalla presenza di un elemento *upstream*. Il ruolo di AidB nella regolazione trascrizionale del proprio gene è stato esaminato mediante esperimenti di fusione trascrizionale. Nello specifico, il ceppo di *E. coli wild type* MV1161 ed il ceppo Δ*aidB* MV5924 sono stati trasformati con un plasmide contenente il gene *reporter lacZ* sotto il controllo di P*aidB*. Sono state allestite colture cellulari dei due ceppi ricombinanti ed è stata analizzata la quantità di β-galattosidasi prodotta. Negli estratti del ceppo mutante è stato registrato un incremento della produzione di β-galattosidasi di ~10 volte rispetto allo *strain* selvatico. Questi risultati hanno suggerito

che, nella cellula batterica non esposta all'azione di molecole alchilanti, la proteina AidB reprime la trascrizione del proprio gene. Al fine di verificare che l'incrementata espressione del gene *lacZ* registrata nel ceppo mutante, fosse effettivamente imputabile all'assenza della proteina AidB, è stato effettuato un esperimento di complementazione: lo *strain* MV5924 è stato trasformato con un plasmide di espressione recante il gene *aidB wild type.* Nel ceppo complementato, i valori di attività β-galattosidasica sono risultati equiparabili a quelli registrati per le cellule *wild type*, dimostrando il coinvolgimento diretto di AidB nella repressione trascrizionale del suo stesso gene. Il ceppo mutante recante il costrutto *reporter*, è stato anche trasformato o con il plasmide esprimente la sola regione ammino-terminale di AidB o con quello contenente la porzione genica relativa al frammento carbossi-terminale. Nel ceppo complementato con il dominio C-terminale, sono stati registrati valori di attività β-galattosidasica analoghi a quelli determinati per le cellule *wild type*, dimostrando che questo dominio possiede attività di regolazione trascrizionale.
Il coinvolgimento di AidB nella repressione trascrizionale del proprio gene è stato anche confermato mediante esperimenti di trascrizione *in vitro*.

Analisi funzionale di AidB nella risposta a molecole alchilanti.
Lo studio è stato poi incentrato sulla caratterizzazione del ruolo di AidB nel meccanismo di resistenza della cellula all'azione di substrati alchilanti. Il primo *step* ha previsto l'identificazione di composti mutageni a cui il ceppo Δ*aidB* MV5924 risulta sensibile. A tale scopo, sono state allestite crescite delle cellule *wild type* e Δ*aidB* in presenza di diversi agenti alchilanti: MMS (metilmetansulfonato), MNU (metilnitrosourea), MNNG (N-metil-N'-nitro-N-nitrosoguanidina) ed ENNG (N-etil-N'-nitro-N-nitrosoguanidina). Il ceppo mutante è risultato sensibile alle molecole MNNG ed ENNG evidenziando il coinvolgimento della proteina in un meccanismo di difesa dall'azione delle due nitrosoammine.
AidB, in risposta a condizioni di stress ambientale, influenza il processo trascrizionale.
Esperimenti di proteomica funzionale hanno dimostrato che AidB è parte di un complesso proteico che si assembla in maniera specifica sulla regione *upstream* del promotore ribosomiale *rrnB* P1. Da studi EMSA è risultato che questa proteina lega con elevata affinità regioni di DNA contenenti elementi UP. Al fine di attribuire un significato a queste evidenze, sono stati condotti esperimenti di fusione trascrizionale. Nello specifico, sono stati preparati costrutti *reporter* contenenti il gene *lacZ* sotto il controllo di diverse regioni promotrici:

- Il promotore ribosomiale *rrnB* P1 contenente e non il suo elemento UP
- Il promotore *leuA*, esempio di sequenza promotrice non ribosomiale priva di un elemento UP
- Il promotore *ompF*, esempio di sequenza promotrice non ribosomiale contenente un elemento UP

Colture cellulari dei ceppi *wild type* e Δ*aidB* trasformati con i costrutti *reporter* sono state allestite in assenza e in presenza di composti alchilanti (MMS, MNU, MNNG, ENNG) ed è stata analizzata la quantità di β-galattosidasi presente nei corrispondenti estratti cellulari. Durante la normale crescita batterica, AidB non ha sortito alcun effetto sul processo trascrizionale; d'altro canto, sotto condizioni di stress ambientale, nello *strain* selvatico è stato registrato un incremento dell'espressione del gene *lacZ* rispetto al ceppo mutante. In particolare, è stato dimostrato che la presenza di AidB rende più efficiente la trascrizione a partire da regioni promotrici contenenti un elemento *upstream*. E' stata quindi messa in luce l'attiva partecipazione della proteina nel

meccanismo di resistenza a molecole alchilanti. Altra osservazione interessante scaturita da tale studio è che l'espressione del gene *reporter* risulta maggiore quando il ceppo selvatico è esposto all'azione di MNNG o ENNG piuttosto che agli altri mutageni utilizzati (MMS e MNU), avvalorando l'ipotesi che AidB sia coinvolta nel processo di degradazione delle nitrosammine.

AidB protegge il DNA dall'attacco di agenti alchilanti.

Le informazioni ricavate dagli studi strutturali unitamente alla capacità di AidB di legare il DNA, hanno indotto ad ipotizzare un suo coinvolgimento in meccanismi di protezione dell'acido nucleico dall'attacco di molecole alchilanti. Il progetto si è dunque proposto di indagare, *in vivo* e *in vitro*, tale funzione. Nell'ambito degli esperimenti *in vivo*, crescite del ceppo di *E. coli wild type* e del ceppo mutante Δ*aidB*, trasformati con un vettore plasmidico, sono state allestite in assenza e in presenza di molecole alchilanti (MMS, MNU, MNNG, ENNG). Il DNA plasmidico è stato isolato dalle suddette colture e trattato con due enzimi, la DNA glicosilasi AlkA e la endonucleasi IV da *E. coli*. L'enzima AlkA riconosce e catalizza la scissione di purine e pirimidine alchilate (7), la endonucleasi IV genera rotture di singolo filamento accanto al sito abasico (8). Il trattamento con i due enzimi converte, dunque, il plasmide super-avvolto (*supercoiled*), contenente basi alchilate in una forma rilassata. Tali forme topologiche mostrano una differente velocità di migrazione se sottoposte a separazione elettroforetica, consentendo quindi di determinare l'esistenza e l'entità del danno subito dalle molecole di DNA in presenza e in assenza di AidB. Le analisi condotte sul DNA isolato dalle colture esposte ad agenti mutageni hanno rivelato che, nelle cellule *wild type*, il plasmide è presente principalmente nella forma *supercoiled*; in assenza di AidB, invece, è stato osservato un incremento della forma rilassata, indice dell'avvenuta alchilazione del DNA.

Relativamente agli esperimenti *in vitro*, due tipi di reazioni sono state allestite:

1) la proteina AidB è stata incubata con il DNA plasmidico; il complesso proteina-DNA è stato alchilato con l'agente chimico MMS e poi trattato con gli enzimi AlkA ed endonucleasi IV. Dalle analisi condotte su gel di agarosio, il plasmide è risultato presente nella forma *supercoiled.*
2) Il DNA plasmidico è stato dapprima alchilato con MMS e poi incubato con AidB; la miscela di reazione è stata trattata con gli enzimi AlkA ed endonucleasi IV e sottoposta ad elettroforesi su gel di agarosio. E' stato rivelato l'incremento dell'intensità del segnale corrispondente alla forma rilassata e la conseguente scomparsa della forma *supercoiled*.

Complessivamente, questi esperimenti hanno dimostrato che l'interazione di AidB col DNA è finalizzata alla protezione dell'acido nucleico dall'attacco delle molecole alchilanti e non al riparo dei danni da alchilazione.

I risultati conseguiti dagli studi EMSA e dagli esperimenti di fusione trascrizionale hanno indotto ad ipotizzare che la proteina AidB possa proteggere in maniera preferenziale regioni geniche contenenti un elemento UP. Al fine di investigare questa ipotesi, l'analisi di protezione è stata ristretta ad una specifica porzione: il frammento *lacZ* posto sotto il controllo di sequenze promotrici con e senza elemento UP. Colture dei ceppi *wild type* e Δ*aidB*, trasformati con i costrutti *reporter* precedentemente descritti, sono state allestite in presenza di substrati alchilanti. Il DNA plasmidico è stato isolato dalle suddette colture ed idrolizzato al fine di rilasciare il gene *lacZ*. Tale frammento è stato dapprima trattato con gli enzimi AlkA ed endonucleasi IV e successivamente sottoposto ad elettroforesi su gel di agarosio in condizioni denaturanti, separando i due filamenti di DNA. Questo trattamento ha permesso di

evidenziare l'avvenuta alchilazione del gene *lacZ* correlandola alla comparsa di frammenti tronchi a singolo filamento. Le analisi condotte sui campioni isolati dalle cellule mutanti hanno rivelato la completa degradazione del gene *lacZ*; nei campioni isolati dalle cellule *wild type*, invece, è stata osservata la presenza di frammenti di DNA *full lenght* nel caso in cui l'espressione del gene *lacZ* è guidata da un promotore contenente una regione UP e la presenza di alcuni frammenti tronchi nel caso in cui il gene in esame è sotto il controllo di un promotore senza elemento UP.
I risultati ottenuti hanno dimostrato che l'azione protettiva di AidB si esplica in generale sul DNA e preferenzialmente sulle regioni contenenti elementi *upstream*.

Ricerca di proteine omologhe ad *Ec*AidB in batteri impiegati nel campo del biorisanamento.

Gli studi cristallografici supportano l'ipotesi che la proteina AidB sia capace di legare specifiche classi di molecole alchilanti (le alchil-nitrosammine) ed esserne coinvolta nel processo di degradazione. Tale ipotesi è stata avvalorata dai risultati conseguiti nel corso di questo progetto di tesi: 1) la dimostrazione che il ceppo di *E. coli* Δ*aidB* è particolarmente sensibile alle alchil-nitrosammine MNNG ed ENNG; 2) l'osservazione che nelle cellule trattate con agenti alchilanti, in particolare con alchil-nitrosammine, la presenza di AidB esercita un'influenza positiva sul processo trascrizionale. La caratterizzazione delle proprietà detossificanti di AidB risulta particolarmente suggestiva alla luce della prospettiva di realizzare microrganismi ingegnerizzati capaci di biodegradare molecole alchilanti. In virtù di questa interessante potenzialità, tale progetto si è proposto di ricercare proteine omologhe ad *Ec*AidB in sistemi microbici impiegati nel campo del biorisanamento e caratterizzarne il coinvolgimento nella risposta della cellula a substrati alchilanti. Nello specifico, da studi di allineamento di sequenza è emerso che *Ec*AidB presenta un'elevata similarità con una putativa acil-CoA deidrogenasi (codificata dal gene PP4780) da *Pseudomonas putida* KT2440, un batterio Gram negativo ampiamente utilizzato nel biorisanamento (9). Alla luce di queste premesse, si è proceduto all'espressione ricombinante di PP4780 in *E. coli* e alla successiva analisi del coinvolgimento del corrispondente prodotto proteico nel meccanismo di resistenza ad agenti alchilanti. A tale scopo, sono state allestite colture del ceppo selvatico e delle cellule ricombinanti esprimenti PP4780 o *aidB*, in assenza e in presenza di MNNG o ENNG. Dall'analisi comparativa dei profili di crescita è emerso che le cellule esprimenti PP4780 risultano più resistenti, rispetto agli altri ceppi esaminati, alla condizione di stress ambientale. Ulteriore evidenza del ruolo esercito dal prodotto del gene PP4780 nella risposta a substrati alchilanti è stata ottenuta attraverso la complementazione del ceppo Δ*aidB* con questa molecola proteica. Dall'analisi comparativa dei profili di crescita è emerso che la putativa acil-CoA deidrogenasi da *P. putida*, analogamente ad AidB, è in grado di ripristinare la resistenza del batterio alla condizione di stress in atto. Tali evidenze sperimentali hanno suggerito l'attiva partecipazione di questo fattore proteico nel meccanismo di difesa della cellula batterica dall'azione di molecole alchilanti.
Questo progetto, quindi, attraverso la caratterizzazione del ruolo biologico della proteina AidB e attraverso le prime indagini condotte su PP4780, ha contribuito ad acquisire informazioni utili per la realizzazione di un sistema di biorisanamento ambientale.

BIBLIOGRAFIA

1. Landini, P., and M. R. Volkert. 1995. Transcriptional activation of the *Escherichia coli* adaptive response gene *aidB* is mediated by binding of methylated Ada protein. *J. Biol. Chem.* **270**:8285-8289.
2. Landini, P., and M. R. Volkert. 2000. Regulatory responses of the adaptive response to alkylation damage: a simple regulon with complex regulatory features. *J. Bacteriol.* **182**:6543-6549.
3. Rohankhedkar, M. S., S. B. Mulrooney, W. J. Wedemeyer, and R. P. Hausinger. 2006. The AidB component of the *Escherichia coli* adaptive response to alkylating agents Is a flavin-containing, DNA-binding protein. *J. Bacteriol.* **188**:223-230.
4. Landini, P., L. I. Hajec, and M. R. Volkert. 1994. Structure and transcriptional regulation of the *Escherichia coli* adaptive response gene *aidB*. *J. Bacteriol.* **176**:6583-6589.
5. Bowles, T., A. H. Metz, J. O'Quin, Z. Wawrzak, and B. F. Eichman. 2008. Structure and DNA binding of alkylation response protein AidB. *Proc. Natl. Acad. Sci. USA* **105**:15299-15304.
6. Hirvonen, C. A., W. Ross, C. E. Wozniak, E. Marasco, J. R. Anthony, S. E. Aiyar, V. H. Newburn, and R. L. Gourse. 2001. Contributions of UP elements and the transcription factor FIS to expression from the seven *rrn* P1 promoters in *Escherichia coli*. *J. Bacteriol.* **183**:6305-6314.
7. O'Brien, P. J., and T. Ellenberger. 2004. The *Escherichia coli* 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site. *J. Biol. Chem.* **279**:26876-26884.
8. Levin, J. D., A. W. Johnson, and B. Demple. 1988. Homogeneous *Escherichia coli* endonuclease IV. Characterization of an enzyme that recognizes oxidative damage in DNA. *J. Biol. Chem.* **263**:8066-8071.
9. Reva, O. N., C. Weinel, M. Weinel, K. Böhm, D. Stjepandic, J. D. Hoheisel, and B. Tümmler. 2006. Functional genomics of stress response in *Pseudomonas putida* KT2440. *J. Bacteriol.* **188**:4079-4092.

Introduction

1. Biotechnology and bioremediation: a new challenge for pollution management

Environmental pollutants are compounds that are released into the environment at high concentrations, usually as a consequence of human activities; they cause instability, disorder, harm to the ecosystem. Contaminants are either compounds of industrial origin that present chemical structures alien to the biosphere (xenobiotics), e.g. polychlorobiphenyls (PCBs), polichlorodioxins, trinitrotoluene (TNT) and azo dyes, or natural compounds that have been mobilized to a bioavailable form toxic to living organisms, e.g. hydrocarbons present in fossil fuels and heavy metals present in minerals (Fig.1). Major sources of pollution are: (i) chemical and pharmaceutical industries that produce a wide array of xenobiotics and synthetic polymers; (ii) pulp and paper bleaching, which are the main sources of chlorinated organic compounds in the environment; (iii) mining, which releases heavy metals into biogeochemical cycles; (iv) fossil fuels (coal and petroleum), which may be accidentally released in large amounts into the ecosystem (oil spills) and whose combustion increases significantly CO_2 atmospheric levels (green-house effect) and causes deposition of nitric and sulfuric acids (acid rain and smog); and (v) intensive agriculture, which releases massive amounts of fertilizers, pesticides, and herbicides (Dua, M. *et al.*, 2002; Rieger, P.G. *et al.*, 2002).

The removal of pollutants from the environment via natural physico-chemical and biological processes (natural attenuation) is, in general, a slow and unpredictable way of counteracting anthropogenic pollution and irreversible damage to the biosphere. The urgent need of rehabilitating areas fouled by pollutants has encouraged the search for innovative solutions. In this field, bioremediation is seen as an attractive strategy due to its reputation as a low cost, environmentally friendly and publicly acceptable treatment technology. The term “bioremediation” has been used to describe the process of using and manipulating detoxification abilities of living organisms to transform hazardous organic contaminants into harmless metabolites or mineralize the pollutants into carbon dioxide and water (Fig.1) (Lovley, D.R. 2003; Wackett, L.P. 2003; Wackett, L.P. *et al*., 2000). Although most organisms are endowed with detoxification abilities, microorganisms, particularly bacteria, have been the most well-studied and the most frequently used for bioremediation strategies. Bacteria, which evolved more than three billion years ago, have developed strategies to obtain energy from virtually every compound, playing a crucial role in sustainable development of the biosphere and in biogeochemical cycles. The abundance of microorganisms, together with their great ability for horizontal gene transfer and their high growth rates, allows them to evolve quickly and to adapt to environmentally changing conditions, even to extreme environments that do not allow proliferation of other living organisms. The huge genetic diversity of microorganisms accounts for their great metabolic versatility (De Lorenzo, V. 2001; Lovley, D.R. 2003; Timmis, K.N. *et al*., 1999).

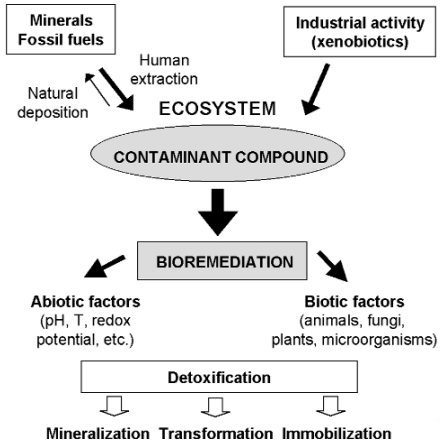


Fig.1: Main sources of pollution in the ecosystem and the factors that influence bioremediation processes.

A large number of microorganisms have been isolated and applied to both *in situ* and *ex-situ* bioremediation processes in recent years, and the identification of new microbes with novel metabolic potential offers an attractive route to solve environmental problems (Dua, M. *et al.*, 2002). On the other hand, advances in genetic and protein engineering techniques have opened up new avenues towards the design of genetically engineered microorganisms (GEMs) and enzymes with the desired biodegradation properties (Debarati, P. *et al.*, 2005). The relative cheapness of the processes, the wide degradation potential offered by the different types of microorganisms which can be employed, and the new frontiers opened by genetic engineering, render bioremediation one of the most promising alternatives for efficient cleanup of pollutants.

2. Alkylating agents as environmental contaminants

Alkylating agents comprise a broad class of highly reactive chemical compounds that introduce alkyl groups into biologically active molecules and prevent normal functioning. Living organisms are continuously exposed to alkylating molecules released into the ecosystem at high concentrations, usually as a consequence of human activities (Vaughan, P. *et al.*, 1991; Taverna, P. and Sedgwick, B. 1996).

Alkylating compounds such as 1,2-dichloroethane, 1-chloro-2,3-epoxypropane and methyl isothiocyanate have a widespread use in the industry as solvents, intermediates in chemical synthesis and fumigant pesticides. Food (cured meats and different fish products), beverages and tobacco smoke can be a source of exposure to alkylating *N*-nitroso compounds (Jagerstad, M. and Skog, K. 2005), a class of agents highly reactive. Alkylating compounds found in the environment are also produced by microorganisms or may be formed by chemical reactions. Some Streptomyces sp. release alkylating antibiotics, such as streptozotocin and azaserine, into the soil creating an urgent need for an adaptive response in other microorganisms. Furthermore, certain algae and fungi, growing in saline environments generate methyl chloride (MeCl) as a product of chloride detoxification (Sedgwick, B. and Vaughan, P. 1991). Alkylating agents may be chemically formed by nitrosations, in slightly acidic conditions, of amides, amines, amino acids and peptides (Sedgwick, B. 1997; Sedgwick, B. and Vaughan, P. 1991). These reactions could occur in decaying matter, in acidic soils or in putrid water.

Numerous alkylating agents found in the environment are known to be extremely cytotoxic. These chemicals react deleteriously with cellular macromolecules (DNA, RNA and proteins) either directly or following metabolic activation (Taverna, P. and Sedgwick, B. 1996; Vaughan, P. *et al.*, 1991). Arguably, the most important cellular target is the DNA molecule. The alkylating agents can introduce methyl or larger alkyl groups into all the available nitrogen and oxygen atoms in DNA bases (Fig.2) in addition at the anionic oxygen of the phosphodiester backbone (Sedgwick, B. and Lindahl, T. 2002; Sedgwick, B. 2004). Alkyl base lesions can arrest replication, interrupt transcription, or signal the activation of cell-cycle checkpoints or apoptosis. In mammals they could be involved in carcinogenesis, neurodegenerative disease and aging.

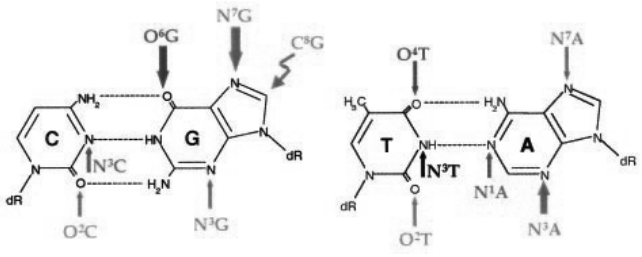


Fig.2: Sites of alkylation on the DNA bases. Thick arrows indicate sites alkylated by most of the agents. The curly arrow indicates an additional site alkylated by alkyl radicals.

The majority of evidence indicates that among the 11 identified base modifications, 3-methyladenine and O^6-methylguanine are mainly responsible for the biological effects of alkylation agents (Singer, B. 1976). The alkylating agent-DNA interaction can also result in cross-linking or strand-breaking reactions. Depending on their processing by the cell, these lesions can give rise either to mutation or to cell death. The relative proportions of the different base lesions depend on the nature of the alkylating agent, its reaction mechanism and the secondary structure of the DNA target (Table 1).

Table 1: Relative proportions of reaction at each base position by common alkylating agents. nd= not detected.

Position of alkylation	S_N2		S_N1	
	ssDNA / RNA	dsDNA	ssDNA / RNA	dsDNA
Adenine				
N1	18	3.8	2.8	1.3
N3	1.4	10.4	2.6	9
N7	3.8	1.8	1.8	1.7
Guanine				
N3	~1	0.6	0.4	0.8
O^6	-	0.3	3	6.3
N7	68	83	69	67
Uracil/Thymine				
O^2	nd	nd	nd	0.11
N3	nd	nd	nd	0.3
O^4	nd	nd	nd	0.4
Cytosine				
O^2	nd	nd	nd	0.1
N3	10	<1	2.3	0.6
Diester	2	0.8	~10	17

Based on the reaction mechanism used, alkylating agents fall into two chemical categories: S_N1 reagents, which typically react through a monomolecular mechanism (meaning spontaneous departure of the leaving group), and S_N2 agents, that act through a bimolecular reaction (leaving group departure occurs only upon reaction with another species). The S_N1 type agents such as *N*-methylnitrosourea (MNU), and *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine (MNNG) introduce alkyl adducts at ring nitrogenes, exocyclic oxygens of the bases, and at oxygens of the sugar-phosphate backbone; S_N2 agents, such as methyl methanesulfonate (MMS) and methyl iodide (MeI) target mainly ring nitrogenes. These lesions are less readily formed in duplex DNA because the modification sites are involved in base pairing, and are therefore shielded from alkylation. Additionally, alkylated lesions can arise through the reaction of DNA with alkyl radicals. These modified base lesions are both toxic and mutagenic (Delaney, J.C. and Essigmann, J.M. 2004; Delaney, J.C. and Essigmann, J.M. 1999) and their repair facilitates cellular survival. Environmental mutagens such as 1,2-dimethylhydrazine, tert-butylhydroperoxide and diazo-quinones generate methyl radicals that react with guanine residues in DNA to form miscoding 8-methylguanine adducts (Hix *et al*., 1995).

One implication of the mutagenic property of alkylating agents is their carcinogenic potential. Cytotoxic alkylating agents are commonly used as drugs in cancer chemotherapy (Hurley, L.H. 2002; Chaney, S.G. and Sancar, A. 1996). DNA damage seems to be a critical component of the efficacy of these drugs, and increased DNA repair is a common cause of developed resistance (Gerson, S.L. and Willson, J.K. 1995). Both mono-and bi-functional alkylators are used as chemotherapeutic agents and these drugs function by creating a range of adducts from alkyl lesions to interstrand cross-links (Margison, G.P. *et al*., 2002; Ludlum, D.B. 1990).

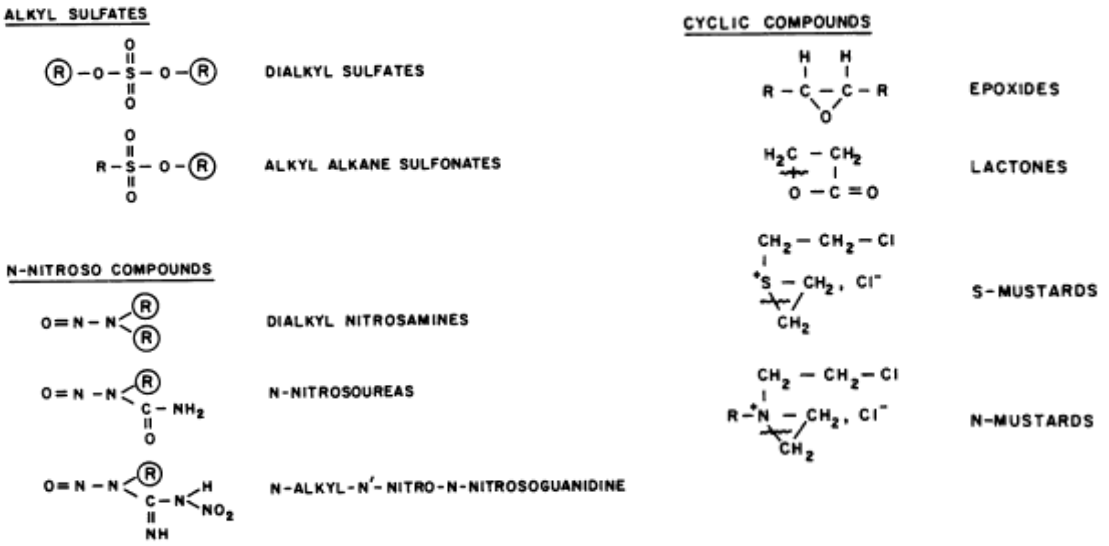


Fig.3: Structural formulas of some alkylating agents.

Useful chemotherapeutic alkylating agents include alkyl sulfonates, nitrosoureas, nitrogen mustards, ethylenimines derivatives, and triazines (Fig.3).
Therefore, the release of alkylating molecules in the ecosystem is a remarkable source of pollution and perturbation of life since many of these compounds are toxic, relatively new, stable, recalcitrant and less easy to degrade than many other organic compounds. These concerns have led to new and/or stricter regulations regarding alkylating wastewater discharges. Various methods of alkylating agents removal from a contaminated site have been proposed and used but bioremediation mediated by microorganisms has emerged as friendly and cost-competitive alternative.

3. Microbial utilization of alkylating compounds

Microorganisms excel at using organic substances, natural or synthetic, as sources of nutrients and energy. The explanation for their remarkable range of degradative abilities is that, by the time human beings came on the scene, microorganisms had already coexisted for billions of years with an immense variety of organic compounds. The vast diversity of potential substrates for growth led to the evolution of enzymes capable of transforming many unrelated natural organic compounds by many different catalytic mechanisms. The resulting giant “library” of microbial enzymes serves as raw material for further evolution whenever a new chemical becomes available (Butler, C.S. and Mason, J.R. 1997; Ellis, B.M.L. 2000).
Microorganisms can use a wide array of pollutants through aerobic or anaerobic processes. In aerobic environments, oxygen is the most common final electron acceptor. However, many polluted environments are often anoxic, e.g., aquifers, aquatic sediments and submerged soils. In such environments, biodegradation is carried out by either strict anaerobes or facultative microorganisms using alternative electron acceptors, such as nitrate (denitrifying organisms), sulphate (sulfate reducers), Fe(III) (ferric-ion reducers), CO_2 (methanogens), or other acceptors (chlorate, Mn, Cr, U, etc.) (Gibson, J. *et al.*, 2002; Lovley, D.R. 2003; Widdel, F. *et al.*, 2001).
Structurally diverse alkylating molecules such as the fumigant pesticides methyl bromide (MeBr), trichloronitromethane, methyl isothiocyanate (MITC) and the alkylating N-nitroso compounds are among the most prevalent pollutants released

into the ecosystem (Zhang, Y. *et al*., 2005)**.** Bioremediation is seen as an attractive route to reduce the concentration and toxicity of these contaminants.
A wide diversity of microorganisms with the capacity to biotransform alkylating compounds have been identified. Methanotrophs, pseudomonads and nitrifiers are the attractive models for such bioremediation processes. In the case of MeBr, methanotrophic and nitrifying bacteria are able of co-oxidizing MeBr to CO_2 during the oxidation of methane and ammonia, respectively (Oremland, R.S. *et al*., 1994; Rasche, M.E. *et al*., 1990). The degradation of MeBr is catalyzed by methane and ammonia monooxygenase enzyme, respectively. Trichloronitromethane can be dehalogenated by *Pseudomonas* species, with the major metabolic pathway occurring through three successive reductive dehalogenations to nitromethane (Castro, C.E. *et al*., 1983). Microorganisms responsible for degradation of MITC specifically target the isothiocyanate functional group. Recently, several bacteria expressing monooxygenase enzymes have been reported to degrade *N*-nitrosodimethylamine (NDMA), a potent alkylating agent that has been detected in discharges of industries as rubber manufacturing, leather tanning and, in addition, sewage treatment plant effluent (Fournier, D. *et al*., 2009)**.** NDMA is a potent carcinogen, so the metabolism of NDMA and other nitrosamines by mammals has been widely studied (Mitch, W.A. and Sedlack, D.L. 2004). Metabolic conversion of NDMA is initiated by the cytochrome P450-dependent mixed function oxidase system (P-450) (Haggerty, H.G. and Holsapple, M.P. 1990; Lee, V.M. *et al*., 1996) and follows either the α-hydroxylation or a denitrosation pathway of the nitrosamine. The α-hydroxylation pathway results in the formation of the strongly methylating methyldiazonium ion ($CH_3N^+{\equiv}N$), which alkylates biological macromolecules such as DNA, RNA, and proteins. In this manner NDMA can exert its genotoxic effects (Sohn, S.O. *et al*., 2001). Alternately, NDMA can be oxidized via a denitrosation route, which leads to the formation of methylamine (CH_3NH_2) and formaldehyde as metabolites. Interestingly, various bacteria expressing broad-specificity monooxygenase enzymes have been observed to degrade NDMA, including *Pseudomonas mendocina* KR1 (toluene-4-monooxygenase; Sharp, J.O. *et al*., 2005), *Rhodococcus ruber* ENV425 (propane monooxygenase; Fournier, D. *et al*., 2009) and *Methylosinus trichosporium* OB3b (soluble methane monooxygenase; Yoshinari, T. and Shafer, D. 1990). However, few data exist concerning the pathways of NDMA and other nitrosamines metabolism in bacteria. Studying the biochemistry and genetics of the microbial pathways is crucial to develop efficient bioremediation strategies. Indeed, although microorganisms have acquired the ability to use pollutants as carbon and energy sources, their efficiency at removing such molecules might not be optimal for cleaning up present-day pollution. Microorganisms have evolved towards ecological fitness rather than biotechnological efficiency; thus, it would take a long time for bacteria capable of cleaning up anthropogenic pollution to evolve by natural selection. Hence, developing an understanding of microbial communities and their response to the natural environment and pollutants, becomes crucial to recreate and accelerate natural processes in the test tube as well as to accomplish their rational manipulation to design more efficient biocatalysts for different biotechnological applications. These include: (i) bioremediation of polluted sites, (ii) biotransformation of toxic compounds into fine chemicals and other high added-value products (green chemistry), and (iii) development of *in situ* biomonitoring devices and biosensors to monitor pollutant bioavailability (De Lorenzo, V. 2001; Timmis, K.N. *et al*, 1999).

4. The *Escherichia coli* adaptive response to alkylating agents

Alkylating agents potentially cytotoxic and mutagenic occur in the environment and also in the living cells, as byproducts of normal metabolism. Since these molecules are ubiquitous and hence unavoidable, all organisms (eubacteria, archaebacteria, and eukaryotes) possess several defense systems to overcome their effects. The repair of alkylation damage to DNA involves at least four different mechanisms: (i) direct repair mediated by methyltransferases or oxidative demethylases; (ii) base excision repair initiated by DNA glycosylases; (iii) mismatch repair system; and (iv) nucleotide excision repair (Pegg, A.E. 2000). Among the systems protecting cells against the action of alkylating agents, there is the adaptive response to DNA damage. To defend against fluctuating levels of environmental alkylating compounds, many bacteria (*Bacillus subtilis*, *Gloetrichia ghosei*, *Micrococcus luteus*, *Pseudomonas aeruginosa*, *Shigella sonnei*, and *Xanthomonas maltophilia*) mount an inducible response that enhances cellular resistance to alkylation damage. The adaptive response to alkylating agents has been most extensively studied in model organism *Escherichia coli*. Decades ago, it was discovered that the relationship between mutation frequency in chemically-treated *E. coli* and the length of exposure to the alkylating agent was not linear, but rather increased initially and then reached a plateau (Cerda-Olmedo, E. and Hanawalt, P. C. 1968; Neale, S. 1972). Interestingly, the height of that plateau was related to the dose of alkylating molecule to which the bacteria were exposed (Jimenez-Sanchez, A. and Cerda-Olmedo, E. 1975). These observations, combined with the subsequent discovery that exposure of *E. coli* cells to sub-lethal doses of alkylating agents renders them resistant to higher concentrations of drug, led to the discovery of the so-called "adaptive response" (Jeggo, P. *et al*., 1977; Samson, L. and Cairns, J. 1977). The inducible protective effect required active protein synthesis (Samson, L. and Cairns, J. 1977), suggesting that it was due to the production of proteins that could defend against the toxic and mutagenic effects of alkylating agents. Later studies demonstrated that the phenomenon involved increasing expression of four genes, (*ada, alkA, alkB,* and *aidB*) (Schendel, P. F. and Robins, P. E. 1978; Evensen, G. and Seeberg, E. 1982), three of which produce proteins of established function. The first protein product of the adaptive response to be identified was Ada. The characterization of this factor revealed its bifunctional nature: Ada repairs alkylated bases and also regulates the adaptive response (Fig.4).

<u>Ada as a multi-substrate repair protein</u>.

Ada is the key enzyme of the adaptive response; this protein is composed of two major domains, a 19 kDa C-terminal (C-Ada19) and a 20 kDa N-terminal (N-Ada20), linked by a hinge region. These domains are capable of carrying out two types of DNA repair reaction: C-Ada19 directly dealkylates the mutagenic bases O^6-methylguanine and O^4-methylthymine, and transfers the methyl groups on to its Cys-321 residue (Demple, B. *et al*., 1985). N-Ada20 demethylates S_p-diastereo-isomers of the apparently innocuous methylphosphotriester lesion by methyl transfer on to the Cys-38 residue (Kondo, H. *et al.*, 1986; Lindahl, T. *et al.*, 1988). This methylation reaction alleviates the normal repulsion that exists between the negatively charged cysteine center and the DNA backbone (He, C. *et al*., 2005), increasing the affinity of Ada for DNA and converting the protein into a transcription factor regulated by post-translational modification. Cys-38 of Ada can be directly methylated by S_N2 agents, such as methyl iodide (MeI), which may be an alternative of Ada activation as a positive gene regulator (Takahashi, K. *et al*., 1988). Ethylating agents, despite

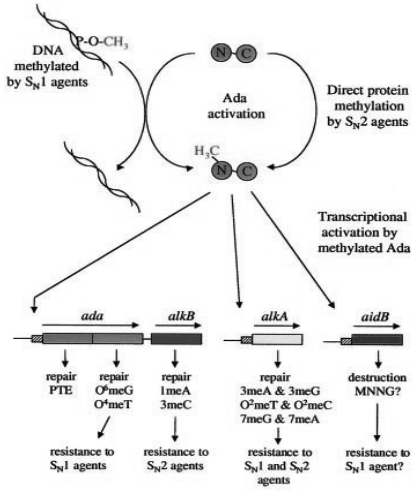


Fig.4: Representation of the *E. coli* Ada response. Ada is activated as a positive regulator by methylation of its Cys-38 in the N-terminal domain. This activation occurs by repair of methylphosphotriesters (PTE) in DNA or by direct protein methylation. Activated Ada induces expression of several genes.

produce more triesters in the DNA backbone than their corresponding methylating counterparts (Singer, B. and Kusmierek, J.T. 1982), are weaker at inducing the adaptive response. This effect is not due to an inability of the protein to remove ethyl groups, however, as the ethylphosphotriesters are repaired by Ada (Margison, G. P. *et al*., 1985) indicating that ethylation of the Cys-38 residue is less efficient at activating transcriptional activity than methylation.
Both DNA repair actions of Ada are suicidal, resulting in loss of the repair function of the Ada protein in a stoichiometric manner, and requiring the activity to be replenished through new protein synthesis. Interestingly, when repair occurs via the Ada pathway, the restoration of a single alkylated base consumes an entire protein molecule. Although this repair method seems energetically expensive, it allows reversal of damage in a single step, without compromising the integrity of the DNA backbone or the coding content of the DNA strand. For these reasons, direct reversal of base damage, such as that exhibited by the suicidal Ada protein, can be advantageous. The question concerning the mechanism of the Ada response termination remains open. It has been postulated that the activated Ada protein is simply diluted by cell division after withdrawal of alkylating agents (Lindahl, T. *et al.*, 1988). A second hypothesis is that unmethylated Ada, when all repairable methylphosphotriesters have been repaired, shuts off the adaptive response by accumulation in the cell and competition with the methylated protein for binding to the regulated genes (Saget, B.M. and Walker, G.C. 1994). This negative feedback loop to some extent mitigates the energetic burden of a suicidal protein repair system.

Ada as a transcriptional regulator.
The Ada protein is activated as a transcriptional factor by methylation of its Cys-38 residue in the amino-terminal half of the protein. N-methylated Ada acquires the ability to bind to specific DNA sequences in the promoter region of the *ada-alkB* operon and the *alkA* and *aidB* genes (Landini, P. and Volkert, M.R. 2000) (Fig.4). Binding of N-methylated Ada makes the promoters accessible to RNA polymerase and therefore activates their transcription (Nakabeppu, Y. and Sekiguchi, M. 1986). The mechanisms by which Ada interacts with the promoters of the Ada regulon and RNA polymerase have been reviewed in detail (Landini, P. and Volkert, M.R. 2000). The activation of *ada*, *alkB, alkA* and *aidB* genes confers increased cellular resistance to the mutagenic and cytotoxic effects of alkylating agents. Three of the four genes are known to be involved in the repair of DNA lesions and each employs a different mechanism.

AlkB, a DNA dioxygenase.
AlkB was shown to be an α-ketoglutarate-Fe(II)-dependent dioxygenase that hydroxylates the aberrant methyl or larger alkyl group of a range of DNA adducts (1-methyladenine, 3-methylcytosine, 3-ethylcytosine, 1-ethyladenine, hydroxyethyl,

propyl, hydroxypropyl lesions) creating an unstable intermediate that will decompose spontaneously to release formaldehyde and the repaired base (Trewick, S.C. *et al*., 2002; Choi, J. Y. *et al*., 2006).

AlkA, a DNA glycosylase with broad substrate specificity.

The *E. coli alkA* gene encodes a 3-methyladenine-DNA glycosylase, which is responsible for excising a broad spectrum of lethal adducts from the genome, initiating the process of base excision repair (BER). AlkA removes the abundant lesion 3-methyladenine and it also can excise the minor products 3-methylguanine, 7-methylguanine, O^2-alkylpyrimidines (Thomas, L., *et al*., 1982; Lindahl, T. 1976; Choi, J. Y. and Guengerich, F.P. 2004).

E. coli has a second 3-methyladenine-DNA glycosylase (Tag) and a second O^6-methylguanine-DNA methyltransferase (Ogt) that are expressed constitutively, and will repair some 3-methyladenine and O^6-methylguanine lesions in DNA during the vulnerable period in which the adaptive response is being induced (Samson, L. 1992). Ingeniously, methylphosphotriesters, a relatively innocuous type of DNA damage, that are not repaired by Ogt or any other known constitutive activity, are used as the inducing signal for the adaptive response. Repair of one of the two stereo-isomers of methylphosphotriesters therefore serves solely as a sensor for changing levels of DNA alkylation damage in bacteria (Lindahl, T. *et al*., 1988).

AidB, the most enigmatic protein involved in the Ada-response.

Despite years of effort, the function of the fourth adaptive response protein, AidB, is still unclear. In an attempt to elucidate the possible function of this protein, the deduced amino sequence of the 60.5 kDa AidB was used in homology searches in Swiss Protein and GenBank-EMBL databases. While no homology was found with any known DNA repair protein, significant homology was detected with several mammalian acyl-coenzyme A dehydrogenases (ACADs), a family of enzymes which use a flavin adenine dinucleotide (FAD) to catalyze the α,β,-dehydrogenation of acyl-coenzyme A (acyl-CoA) conjugates (Landini, P. *et al*., 1994; Rohankhedkar, M.S. *et al*., 2006). Multiple sequence alignments have showed that numerous residues of ACADs are conserved in AidB sequence; in particular, the conservation of Glu425, that corresponds to the catalytic glutamate in most of the short- and medium-chain acyl-CoA dehydrogenases, suggests that a dehydrogenase or oxidase activity is essential to the physiological function of the protein.

It has been determined that AidB shows 24.6% identity with human isovaleryl-CoA dehydrogenase (IVD) precursor (Landini, P. *et al*., 1994), an enzyme involved in leucine metabolism in mammalian cells. AidB has been shown recently to bind stoichiometric amounts of redox active FAD, and IVD activity has been detected from both crude cell extracts overexpressing AidB and purified preparations (Landini, P. *et al*., 1994; Rohankhedkar, M.S. *et al*., 2006). However, the level of IVD activity observed in AidB is quite low compared to other ACADs. Probably the IVD activity in AidB is a side reaction that is distinct from its functional role (Rohankhedkar, M.S. *et al*., 2006).

The Ada-dependent upregulation of AidB *in vivo*, in response to the presence of alkylating agents is one of two ways in which synthesis of the protein can be triggered. Expression of the *aidB* gene is also induced by anaerobiosis or by addition of sodium acetate to growth medium, at a slightly acidic pH (ranging 6.0 to 6.8), in an Ada-independent fashion (Volkert, M. R. *et al*., 1989). This second pathway of expression is mediated by *rpoS* (Volkert, M. R. *et al*., 1994), a gene which encodes

an alternative sigma factor of RNA polymerase, mainly active in late-logarithmic and stationary phases of *E. coli* growth. Discovery of the IVD activity of AidB led to the hypothesis that the protein, like other acyl-CoA dehydrogenases plays a role in energy production during fermentation and anaerobic metabolism (Brockman, H. L, and Wood, W.A. 1975). This hypothesis could explain the observed induction of AidB in response to anaerobiosis, but does not provide a explanation for its Ada-dependent induction.

In an attempt to define the function of AidB in the adaptive response, the effect of its overexpression in *E. coli* cells exposed to methylating agent MNNG, was tested. Overexpression of AidB reduced the mutagenic effect of MNNG (Landini, P. *et al*., 1994) prompting the hypothesis that the protein may act as a detoxification enzyme to destroy nitrosoguanidines or their reactive intermediates. Paradoxically, strains with insertionally inactivated *aidB* showed two different phenotypes: those with alkylation sensitivity identical to wild-type and those that display increased resistance to both lethal and mutagenic effects of MNNG, but no alteration in sensitivity to other alkylating compounds (Volkert, M. R. *et al*., 1986). The finding that the same phenotype can be produced both by insertional mutation and by overexpression of the *aidB* gene is apparently paradoxical. It has been observed that insertion mutants studied all affected only the terminal third of the gene. It is therefore possible that these mutants produce a functional fragment of AidB protein. The observed phenotypes could be explained if expression of the first two-thirds of the gene produced a functional protein that is either more stable or more active in the resistant strains and similar to wild-type AidB in the non-resistant strains. Since sensitivity to only MNNG is affected in the insertional mutants, it has been suggested that AidB may be involved either in repair of MNNG specific DNA lesions or in an MNNG detoxification pathway. MNNG, in fact, must be metabolically activated to exert its mutagenic and lethal activity (Lawley, P. D. 1974). In *E. coli* such a pathway involves glutathione; interestingly, it has been detected that a nonprotein thiol: glutathione-deficient mutants are generally more resistant to MNNG (Sedgwick, B. and Robins, P. 1980). AidB might be involved in this or in an alternative detoxification pathway, either by acting directly on MNNG or by synthesis or modification of some MNNG-reactive compound. It is noteworthy that when *E. coli* is exposed to sodium acetate in a medium with a final pH of 6.5 (inducing conditions for *aidB*), it becomes more resistant to the mutagenic effect of MNNG (Oktyabrsky, O. N. *et al*., 1993).

More recent studies have shown that the AidB protein is able to bind to dsDNA (Rohankhedkar, M.S. *et al*., 2006). AidB was suspected to have this ability on the basis of its coregulation with DNA repair enzymes and because a short region of its C terminus is homologous to a DNA-binding domain of human topoisomerase I. No sequence specificity was observed, but further studies are warranted to characterize this in greater detail. The presence of an enzymatically active flavin and the protein's demonstrated DNA-binding capability led to the suggestion that AidB might catalyze the repair of alkylated DNA by a

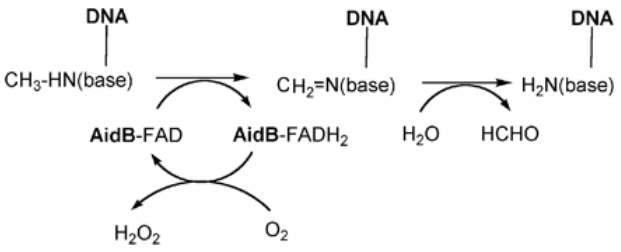


Fig.5: Hypothesis that AidB is a methylated-base dehydrogenase. The Schiff's base product would hydrolyze to release the free base plus formaldehyde, and the reduced flavin could either react with oxygen to produce hydrogen peroxide or transfer electrons to another suitable electron acceptor.

dehydrogenase mechanism, such as that shown in Fig.5 (Rohankhedkar, M.S. *et al*., 2006).

Whether AidB acts to reduce mutagenicity by a DNA repair or a detoxification mechanism is currently a matter of speculation. However, It has been observed that, whereas *ada* and *alkB* genes form a transcriptional unit, *alkA* and *aidB* are well separated from this operon and from each other. These observations raise the possibility that AidB may be of secondary importance for repair of alkylation damage to DNA. To help resolve this issue, recently the crystal structure of the AidB protein has been determined (Bowles, T. *et al*., 2008). The structural analysis revealed that AidB is well equipped to sterically occlude dsDNA from chemical attack. Importantly, the structure is not consistent with a DNA repair function. Alternatively, the unique chemical environment of AidB's putative FAD active site provides a rationale for a possible role in deactivation of alkylating agents.

5. Structural properties and potential cellular function of AidB

As predicted (Landini, P. *et al*., 1994; Rohankhedkar, M.S. *et al*., 2006), the structure of *E. coli* AidB is representative of the ACAD family of flavoproteins. ACADs form tightly bound homodimers or homotetramers and typically have three domains: an N-terminal α-helical domain, a middle β-sheet domain, and a C-terminal helical bundle (Kim, J.J. and Miura, R. 2004). However, AidB contains unique features that distinguish it functionally from the ACAD enzymes. One AidB subunit (Fig.6A) consists of an N-terminal α-helical domain (domain I, residues 1–179), a seven-stranded β-barrel (domain II, residues 180–285), a central α-helical region (domain III, residues 286–444), and an α-helical domain at its C terminus (domain IV, residues 445–540). Domains I–III collectively constitute the ACAD core fold. Domain IV, however, is not present in the ACAD family (Fig.7). In the crystal structure, AidB forms a tetramer (Fig.6B). Sedimentation velocity ultracentrifugation showed that AidB (60,590 Da per subunit) sediments as a 234-kDa protein, and gel filtration analysis was consistent with a tetramer in solutions. The AidB tetramer is a dimer-of-dimers. As in ACAD structures, AB and CD dimers are each formed from extensive contacts between domains II and III. Two FAD molecules (one per subunit) are

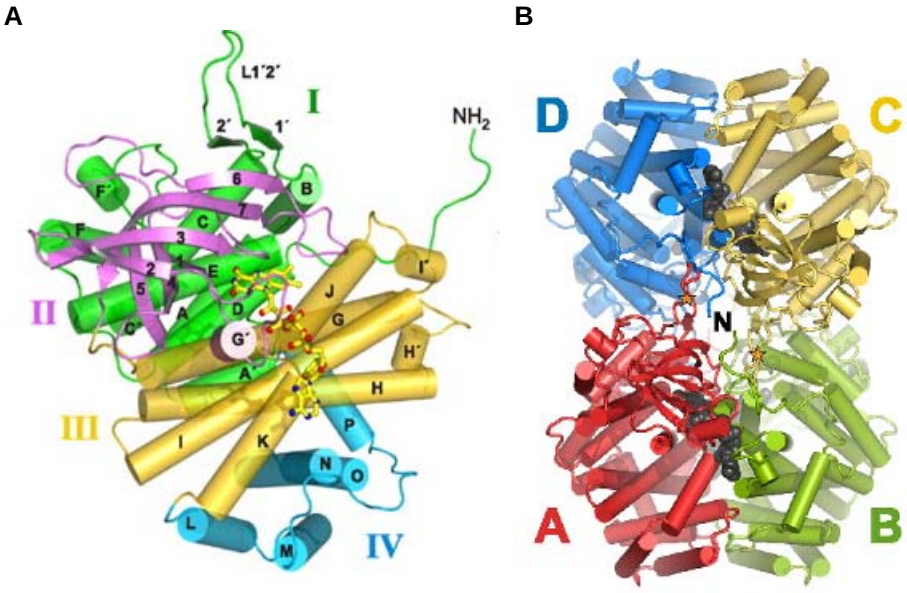


Fig.6: The structure of *E. coli* AidB. (A) Stereo image of one subunit colored according to domains. The FAD is shown as ball and stick. (B) Representation of the AidB tetramer. Four subunits A, B, C, and D associate with dihedral symmetry as a dimer of dimers. The FAD cofactors at each of the A/B and C/D dimer interfaces are shown as black CPK spheres. N-terminal extensions and L1'2' loops that form the AB/CD interface unique to AidB are labeled with the letter N and orange stars, respectively.

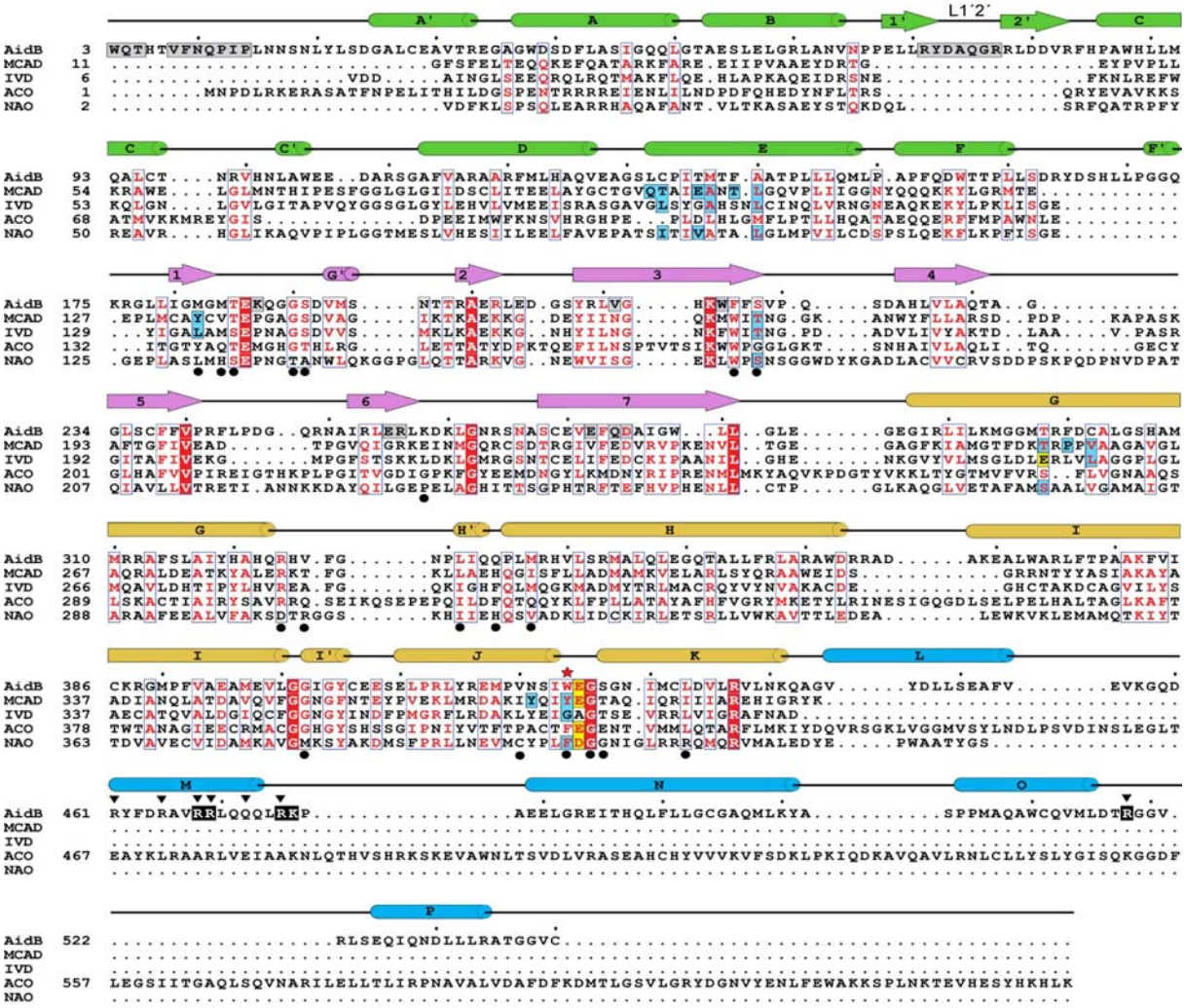


Fig.7: Structure-based sequence alignment of AidB, medium-chain acyl-CoA dehydrogenase (MCAD, PDB ID code 3MDE) (Kim, J.J. *et al*., 1993), isovaleryl acyl-CoA dehydrogenase (IVD, PDB ID code 1IVH) (Tiffany, K.A. *et al*. 1997), acyl-CoA oxidase II (ACO, PDB ID code 1IS2) (Nakajima, Y. *et al*., 2002), and nitroalkane oxidase (NAO, PDB ID code 2C0U) (Nagpal, A. *et al*., 2006). AidB residues important for tetramerization and DNA binding are highlighted gray and black, respectively, and those predicted to contact DNA are marked with black triangles. ACAD substrate binding and catalytic residues are highlighted blue and yellow, respectively. Black circles denote FAD binding residues, and AidB Trp-424 is marked with a red star.

bound at each of these A/B and C/D dimer interfaces (Fig.6B).

Unlike the ACADs, however, AidB has a unique quaternary architecture formed by an extended random coil at the extreme N terminus and a β-hairpin loop (L1'2') inserted between helices αB and αC that project outward to interlock AB and CD dimers together (Fig.6B and Fig.7). Therefore AidB, as result of the novel N-terminal interface and the additional C-terminal domain IV, adopts an homotetrameric architecture different from the ACAD tetramer. In addition, several unique features distinguish AidB's FAD cavity from the ACAD active sites despite the conservation of their general properties.

The FAD active-site cavity as a sink for reactive alkylating agents.

Proteins within the ACAD superfamily share conserved features inside their FAD active sites (Kim, J.J. and Miura, R. 2004). The FAD binds in an extended conformation at the A/B subunit interface, with the isoalloxazine ring buried within a central cavity that contains catalytic residues. Immediately adjacent to the FAD binding site is a substrate access channel that leads from the protein exterior into the flavin active site. A conserved phenylalanine/tyrosine residue (e.g., MCAD Tyr-375) helps position the substrate for catalysis (Fig.8). The α,β-dehydrogenation/oxidation reactions in ACAD enzymes are typically catalyzed by a conserved glutamate carboxylate (e.g., MCAD Glu-376) and the isoalloxazine N5 nitrogen, which are positioned adjacent to the acyl chain Cα and Cβ carbons, respectively (Ghisla, S. and Thorpe, C. 2004). The FAD–protein interactions within domains II and III are largely conserved among AidB and ACADs (Fig.7). In addition, the substrate access channel in AidB is similar in size and shape to that of ACADs. However, the unique features of the FAD pocket provide a structural rationale for AidB's weak isovaleryl dehydrogenase activity and for the absence of flavin spectral perturbation in the presence of isovaleryl-CoA (Landini, P. *et al*., 1994; Rohankhedkar, M.S. *et al*., 2006). Specifically, AidB differs significantly from the ACAD enzymes in the residues important for positioning the substrates in close proximity to the flavin. The AidB channel is blocked at the flavin ring by Trp-424, which is in the same position in the primary sequence as MCAD Tyr-375 (Figs.7 and 8). Thus, the bulky Trp-424 side chain sterically blocks access to the back half of the channel and, more importantly, to the isoalloxazine N5 nitrogen involved in the α,β-dehydrogenation reaction, and significantly restructures the environment of the AidB putative substrate cavity. It has been proposed that the glycine in this position in IVD allows for binding branched acyl chains (Kim, J.J. and Miura, R. 2004, Tiffany, K.A. *et al*. 1997). In addition to the Trp-424 steric block, the 5-Å displacement of the Glu-425 carboxylate from its normal position in ACADs would render the carboxylate ineffective at abstracting the Cα-H hydrogen from the acyl-CoA substrate. Although the FAD cavity is spacious enough for Trp-424 and Glu-425 side chains to shift position, such a large conformational change to accommodate an extended acyl chain is not likely to occur in light of the small active-site changes observed upon acyl-CoA binding in MCAD and isobutyryl-CoA dehydrogenase (Kim, J.J. *et al*., 1993, Battaile, K.P. *et al*., 2004). Thus, it appears that the elongated FAD cavity in AidB is an evolutionary

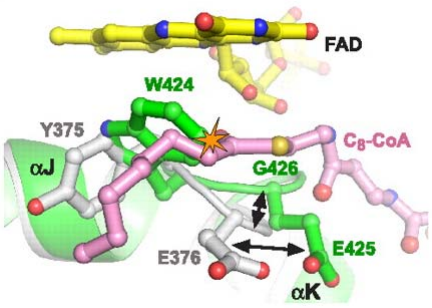


Fig.8: Superposition of AidB (green) and MCAD (gray) structures. The AidB flavin is yellow, and the MCAD octanoyl-CoA substrate is pink. The black arrows highlight the shift in Glu-425 position, and the steric clash between Trp-424 and a fatty acyl substrate is shown as an orange starburst.

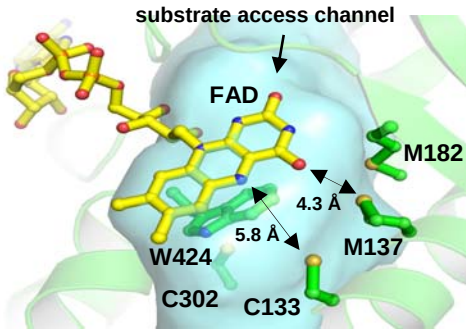


Fig.9: The sulfur-rich FAD binding pocket of AidB. The solvent-accessible surface of the substrate access channel is shown in transparent blue and is lined with Cys-133, Met-137, Met-182, Cys-302, and the re-face of the FAD isoalloxazine ring (yellow carbons). Trp-424 is imbedded within this channel.

remnant and not a bona fide active site for fatty or amino acid acyl-CoA thioesters. The volume of the FAD binding pocket (~290 $Å^2$) is large enough to accommodate a small molecule. AidB has been proposed to detoxify MNNG or one of its reactive metabolites (Landini, P. *et al*., 1994). Unlike MMS or MNU, MNNG is activated by thiols, including glutathione, to produce highly reactive methylation agents (e.g., methyldiazonium ion) (Lawley, P. D. 1974). Interestingly, AidB's FAD cavity is rich in thiol and methylsulfide groups (Fig.9), a characteristic distinctly different from other structural homologues. In fact, four of 13 solvent-accessible residues lining the AidB pocket are cysteine or methionine, whereas there are no such residues in contact with substrates in MCAD, IVD, ACO, or NAO (Fig.7). It is therefore intriguing to speculate that Cys-133, Met-137, Met-182, and Cys-302 may alter the redox potential of AidB's pocket relative to those of other, functionally unrelated enzymes. In particular, Cys-133 and Met-137 sulfur atoms are positioned 5.8 and 4.3 Å from the flavin, respectively, and are oriented toward the mouth of the substrate channel (Fig.9). Thus, it appears that AidB's putative active site may act as a sink for reactive MNNG derivatives.

Structural studies also suggested that the C-terminal domain of AidB, highly electropositive, is responsible for DNA binding. Therefore AidB does not engage DNA at the acyl-CoA substrate binding cavity, the only entry point into the FAD active site. These observations argue against the notion that AidB repairs DNA by an FAD-dependent mechanism. AidB's tetrameric assembly exposes DNA binding surfaces at each end of the tetramer. Nonspecific DNA binding at the ends of the tetramer suggests that the protein might function to protect naked DNA from attack by alkylating agents. A similar role has been observed in the Dps protein, which protects DNA in starved *E. coli* against oxidative damage (Almiron, M. *et al*., 1992; Martinez, A. and Kolter, R. 1997). Both AidB and Dps are up-regulated during stationary phase and are *rpoS* dependent (Volkert, M.R. *et al*., 1994; Altuvia, S. *et al*., 1994). Interestingly, endogenous methylating agents such as nitrosamines are formed as by-products of stationary-phase metabolism (Taverna, P. and Sedgwick, B. 1996). This leads to an accumulation of alkylation damage to DNA during stationary phase, as demonstrated by an increase in spontaneous mutation in methyltransferase (*ada ogt*)-deficient *E. coli* in nondividing cells (Rebeck, G.W. and Samson, L. 1991). Induction of AidB expression, therefore, could serve to prevent endogenous stationary-phase alkylation damage in a manner similar to Dps protection of oxidative damage. Dps protection of DNA is believed to occur by binding DNA duplexes within the pores of hexagonally packed Dps dodecamers (Grant, R.A. *et. al*, 1998). Interestingly, crystal form of AidB displays that the putative DNA binding faces are clustered symmetrically around 25-Å pore that is perfectly sized to accommodate DNA. Thus, it appears that AidB tetramers cluster around DNA to restrict access by damaging agents. Coupled with unique features of FAD site, these observations suggest that AidB may be the cell's line of defense to prevent alkylation damage by protecting DNA and by detoxifying the alkylating compounds before they are able to react with DNA.

6. Potential role for AidB as a detoxification enzyme: a new challenge for pollution cleanup

Despite detailed understanding of Ada, AlkA, and AlkB, the mechanism by which AidB protects against DNA damage in the adaptive response is less well understood. Although the specific function remains to be determined, the structural features of AidB's unique DNA binding domain, subunit organization, and FAD chemical

environment help to support the involvement of the protein in a detoxification pathway of certain alkylators.
The possibility that AidB might degrade alkylating molecules, among the most prevalent pollutants released into the environment, makes this protein a promising tool for bioremediation applications. Besides, interestingly, full-length AidB homologues are not present in many bacteria closely related to *E. coli* (*Klebsiella*, *Vibrio*, *Shewanella*, and *Photorhabdus*), but the closest hits are from some γ-proteobacteria such as *Pseudomonas, Azotobacter*, and *Acinetobacter*. This observation is intriguing because *Pseudomonas* species and closely related organisms are the most extensively studied and the most frequently used for bioremediation applications, owing to their ability to degrade numerous different contaminants (Wackett, L.P. 2003). The huge potential of the pseudomonads does not solely depend on a high proportion of genes responsible for the metabolism, transport and efflux of organic compounds (Nelson, K.E. *et al*., 2002), but also on broad capability of metabolic regulation: indeed, the control of gene expression is the key determinant of their flexibility and, in this respect, a variety of highly integrated regulatory mechanisms have been identified.
Taken together, these observations suggest that the characterization of the potential ability of AidB protein and its homologues to degrade alkylators, could allow the development of new successful strategies for the bioremediation of environments and industrial effluents contaminated by alkylating compounds.

7. Aim of the thesis

The aim of this research project was to explore new potential candidates for the bio-treatment of wastes and environments contaminated by alkylating agents. The study has been specifically focused on AidB protein due its potential ability to degrade certain alkylators.
Initially, the work has been aimed at investigate the role of AidB in the bacterial cell; given that the knowledge of the domain architecture is necessary for understanding the multifunctional properties of a protein, structural and functional characterization of domains present in AidB was performed. Successively, the mechanism by which this protein directly protects *E. coli* cells against alkylating compounds has been determined. Finally, taking into account the potential role played by AidB in the detoxification of alkylating molecules, this experimental work was targeted at identify as well as at characterize *E. coli* AidB homologues in bacteria used for bioremediation applications. Specifically, the study has been focused on the acyl-CoA dehydrogenase coded by the PP4780 gene from *Pseudomonas putida* KT2440. The involvement of this protein in the response to alkylation stress has been investigated in order to enrich the knowledge of this enzyme, thus improving its potential application in bioremediation strategies. In conclusion, the data obtained in this research project support the possibility of developing new successful strategies for the bioremediation of environments and wastes contaminated by alkylating compounds.

REFERENCES

- Battaile, K. P., Nguyen, T. V., Vockley, J., and Kim, J. J. *The Journal of biological chemistry* **279**(16), 16526–16534 (2004).
- Brockman, H. L. and Wood, W. A. *Journal of bacteriology* **124**(3), 1447–1453 (1975).
- Butler, C. S. and Mason, J. R. *Advances in microbial physiology* **38**, 47–84 (1997).
- Castro, C. E., Wade, R. S., and Belser, N. O. *Journal of Agricultural and Food Chemistry* **31**(6), 1184–1187 (1983).
- Cerdá-Olmedo, E. and Hanawalt, P. C. *Molecular and General Genetics* **101**(3), 191–202 (1968).
- Chaney, S. G. and Sancar, A. *J. Natl. Cancer Inst.* **88**(19), 1346–1360 (1996).
- Choi, J. Y. and Guengerich, F. P. *The Journal of biological chemistry* **279**(18), 19217–19229 (2004).
- Choi, J. Y., Angel, K. C., and Guengerich, F. P. *The Journal of biological chemistry* **281**(30), 21062–21072 (2006).
- De Lorenzo, V. *EMBO reports* **2**(5), 357–359 (2001).
- Delaney, J. C. and Essigmann, J. M. *Chemistry & Biology*, 743–753 (1999).
- Delaney, J. C. and Essigmann, J. M. *Proceedings of the National Academy of Sciences of the United States of America* **101**(39), 14051–14056 (2004).
- Demple, B., Sedgwick, B., Robins, P., Totty, N., Waterfield, M. D., and Lindahl, T. *Proceedings of the National Academy of Sciences of the United States of America* **82**(9), 2688–2692 (1985).
- Dua, M., Singh, A., Sethunathan, N., and Johri, A.K. *Applied Microbiology and Biotechnology* **59**(2), 143–152 (2002).
- Dungan, R. S., Gan, J., and Yates, S. R. *Water, Air, and Soil Pollution* **142**(1), 299–310 (2003).
- Ellis, L. B. *Current Opinion in Biotechnology* , 232–235 (2000).
- Evensen, G. and Seeberg, E. *Nature* **296**(5859), 773–775 (1982).
- Fournier, D., Hawari, J., Halasz, A., Streger, S. H., McClay, K. R., Masuda, H., and Hatzinger, P. B. *Appl. Environ. Microbiol.* **75**(15), 5088–5093 (2009).
- Gerson, S. L. and Willson, J. K. *Hematology/oncology clinics of North America* **9**(2), 431–450 (1995).
- Ghisla, S. and Thorpe, C. *European journal of biochemistry / FEBS* **271**(3), 494–508 (2004).
- Grant, R. A., Filman, D. J., Finkel, S. E., Kolter, R., and Hogle, J. M. *Nature structural biology* **5**(4), 294–303 (1998).
- Haggerty, H. G. and Holsapple, M. P. *Toxicology* **63**(1), 1–23 (1990).
- He, C., Hus, J. C., Sun, L. J., Zhou, P., Norman, D. P., Dötsch, V., Wei, H., Gross, J. D., Lane, W. S., Wagner, G., and Verdine, G. L. *Molecular cell* **20**(1), 117–129 (2005).
- Hix, S. *Free Radical Biology and Medicine* **19**(3), 293–301 (1995).
- Hurley, L. H. *Nature Reviews Cancer* **2**(3), 188–200 (2002).
- Ibekwe, A. M., Papiernik, S. K., Gan, J., Yates, S. R., Yang, C. H., and Crowley, D. E. *Appl. Environ. Microbiol.* **67**(7), 3245–3257 (2001).
- Jagerstad, M. and Skog, K. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **574**(1-2), 156–172 (2005).
- Jeggo, P., Defais, T. M., Samson, L., and Schendel, P. *Molecular and General Genetics* **157**(1), 1–9 (1977).
- Jiménez-Sánchez, A. and Cerdá-Olmedo, E. *Mutation research* **28**(3), 337–345 (1975).
- Kim, J. J., Wang, M., and Paschke, R. *Proceedings of the National Academy of Sciences of the United States of America* **90**(16), 7523–7527 (1993).
- Kim, J. J. and Miura, R. *European journal of biochemistry / FEBS* **271**(3), 483–493 (2004).

- Kondo, H., Nakabeppu, Y., Kataoka, H., Kuhara, S., Kawabata, S., and Sekiguchi, M. *The Journal of biological chemistry* **261**(33), 15772–15777 (1986).
- Landini, P. and Volkert, M. R. *Journal of bacteriology* **182**(23), 6543–6549 (2000).
- Landini, P., Hajec, L. I., and Volkert, M. R. *Journal of bacteriology* **176**(21), 6583–6589 (1994).
- Lawley, P. D. *Mutation research* **23**(3), 283–295 (1974).
- Lee, V. M., Keefer, L. K., and Archer, M. C. *Chemical research in toxicology* **9**(8), 1319–1324 (1996).
- Lindahl, T. *Nature* **259**(5538), 64–66 (1976).
- Lindahl, T., Sedgwick, B., Sekiguchi, M., and Nakabeppu, Y. *Annual review of biochemistry* **57**, 133–157 (1988).
- Lovley, D. R. *Nature Reviews Microbiology* **1**(1), 35–44 (2003).
- Ludlum, D. B. *Mutation research* **233**(1-2), 117–126 (1990).
- Margison, G. P., Cooper, D. P., and Brennand, J. *Nucleic acids research* **13**(6), 1939–1952 (1985).
- Margison, G. P., Santibanez Koref, M. F., and Povey, A. C. *Mutagenesis* **17**(6), 483–487 (2002).
- Mitch, W. A. and Sedlak, D. L. *Environmental science & technology* **38**(5), 1445–1454 (2004).
- Nagpal, A., Valley, M. P., Fitzpatrick, P. F., and Orville, A. M. *Biochemistry* **45**(4), 1138–1150 (2006).
- Nakabeppu, Y. and Sekiguchi, M. *Proceedings of the National Academy of Sciences of the United States of America* **83**(17), 6297–6301 (1986).
- Nakajima, Y., Miyahara, I., Hirotsu, K., Nishina, Y., Shiga, K., Setoyama, C., Tamaoki, H., and Miura, R. *J Biochem* **131**(3), 365–374 (2002).
- Neale, S. *Mutation research* **14**(2), 155–164 (1972).
- Nelson, K. E., Weinel, C., Paulsen, I. T., Dodson, R. J., Hilbert, H., Martins dos Santos, V. A., Fouts, D. E., Gill, S. R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R. T., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Chris Lee, P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J. A., Timmis, K. N., Düsterhöft, A., Tümmler, B., and Fraser, C. M. *Environmental microbiology* **4**(12), 799–808 (2002).
- Oktyabrsky, O. N., Golyasnaya, N. V., Smirnova, G. V., Demakov, V. A., Posokhina, N. K. h., and Kholstova, T. A. *Mutation research* **293**(3), 197–204 (1993).
- Oremland, R. S., Miller, L. G., Culbertson, C. W., Connell, T. L., and Jahnke, L. *Appl. Environ. Microbiol.* **60**(10), 3640–3646 (1994).
- Paul, D., Pandey, G., Pandey, J., and Jain, R. *Trends in Biotechnology* **23**(3), 135–142 (2005).
- Pegg, A. E. *Mutation research* **462**(2-3), 83–100 (2000).
- Rasche, M. E., Hyman, M. R., and Arp, D. J. *Appl. Environ. Microbiol.* **56**(8), 2568–2571 (1990).
- Rebeck, G. W. and Samson, L. *Journal of bacteriology* **173**(6), 2068–2076 (1991).
- Rieger, P. *Journal of Biotechnology* **94**(1), 101–123 (2002).
- Saget, B. M. and Walker, G. C. *Proceedings of the National Academy of Sciences of the United States of America* **91**(21), 9730–9734 (1994).
- Samson, L. and Cairns, J. O. H. N. *Nature* **267**(5608), 281–283 (1977).
- Samson, L. *Molecular microbiology* **6**(7), 825–831 (1992).
- Schendel, P. F. and Robins, P. E. *Proceedings of the National Academy of Sciences of the United States of America* **75**(12), 6017–6020 (1978).
- Sedgwick, B. and Lindahl, T. *Oncogene* **21**(58), 8886–8894 (2002).
- Sedgwick, B. and Robins, P. *Molecular and General Genetics* **180**(1), 85–90 (1980).

- Sedgwick, B. and Vaughan, P. *Mutation research* **250**(1-2), 211–221 (1991).
- Sedgwick, B. *Carcinogenesis* **18**(8), 1561–1567 (1997).
- Sedgwick, B. *Nature reviews. Molecular cell biology* **5**(2), 148–157 (2004).
- Sharp, J. O., Wood, T. K., and Alvarez-Cohen, L. *Biotechnology and Bioengineering* **89**(5), 608–618 (2005).
- Singer, B. and Kusmierek, J. T. *Annual review of biochemistry* **51**, 655–693 (1982).
- Singer, B. *Nature* **264**(5584), 333–339 (1976).
- Smirnova, G. V., Oktyabrsky, O. N., Moshonkina, E. V., and Zakirova, N. V. *Mutation research* **314**(1), 51–56 (1994).
- Sohn, O. S., Fiala, E. S., Requeijo, S. P., Weisburger, J. H., and Gonzalez, F. J. *Cancer Res* **61**(23), 8435–8440 (2001).
- Takahashi, K., Kawazoe, Y., Sakumi, K., Nakabeppu, Y., and Sekiguchi, M. *The Journal of biological chemistry* **263**(27), 13490–13492 (1988).
- Taverna, P. and Sedgwick, B. *Journal of bacteriology* **178**(17), 5105–5111 (1996).
- Taverna, P. and Sedgwick, B. *Journal of bacteriology* **178**(17), 5105–5111 (1996).
- Thomas, L., Yang, C. H., and Goldthwait, D. A. *Biochemistry* **21**(6), 1162–1169 (1982).
- Tiffany, K. A., Roberts, D. L., Wang, M., Paschke, R., Mohsen, A. W., Vockley, J., and Kim, J. J. *Biochemistry* **36**(28), 8455–8464 (1997).
- Timmis, K. *Trends in Biotechnology* **17**(5), 200–204 (1999).
- Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. *Nature* **419**(6903), 174–178 (2002).
- Vaughan, P., Sedgwick, B., Hall, J., Gannon, J., and Lindahl, T. *Carcinogenesis* **12**(2), 263–268 (1991).
- Volkert, M. R., Hajec, L. I., and Nguyen, D. C. *J. Bacteriol.* **171**(2), 1196–1198 (1989).
- Volkert, M. R., Nguyen, D. C., and Beard, K. C. *Genetics* **112**(1), 11–26 (1986).
- Wackett, L. *Current Opinion in Biotechnology* **11**(3), 229–231 (2000).
- Wackett, L. P. *Nat Biotech* **21**(2), 136–138 (2003).
- Yoshinari, T. and Shafer, D. *Canadian journal of microbiology* **36**(12), 834–838 (1990).
- Zhang, Y., Spokas, K., and Wang, D. *J Environ Qual* **34**(5), 1566–1572 (2005).

Results and discussion

The results obtained in this thesis are described in the following sections:

Additional protein functions involved in the biogenesis of ribosomes and the DNA repair mechanisms are associated with the transcriptional machinery gathered at the *Escherichia coli rrnB* P1 promoter

Angela Amoresano[1, 2], Claudia Cirulli[1,2], Valentina Rippa[1, 2], Carla Esposito[1], Rosanna Papa[1, 2] and Angela Duilio[1, 2*]

[1]Department of Organic Chemistry and Biochemistry, University Federico II of Naples, Naples, Italy
[2]School of Biotechnological Sciences, University Federico II of Naples, Naples, Italy

ABSTRACT

The transcriptional machinery at the ribosomal RNA *rrnB* P1 promoter of *Escherichia coli* was investigated by functional proteomics. Synthetic oligonucleotides encompassing the -35 sequence of the *rrnB* P1 promoter and its associated element upstream ("UP35") were incubated with the total protein extract. Proteins found to bind exclusively to the "UP35" oligonucleotide were identified by mass spectrometry and specific protein-protein interactions were confirmed by co-immunoprecipitation. Proteins specifically binding to the *rrnB* P1 promoter region belong to different functional groups: in addition to components of the transcription machinery, we also found proteins involved in translation, DNA repair and protection mechanisms and in cell metabolic processes. Presence of several ribosomal proteins strongly suggests that the ribosome can establish direct interaction with bacterial RNA polymerase. Interestingly, we found that non-specific DNA binding proteins, i.e. Dps and AidB, are recruited into the transcriptional complex. Functional hypotheses on the possible biological significance of the different functional groups were suggested.

Keywords: rRNA transcription, proteomics, cellular network, alkylation, fishing for partners, mass spectrometry.

Running title: Complex transcription machinery at *E. coli rrnB* P1 promoter.

*Corresponding Author:
Angela Duilio
Dipartimento di Chimica Organica e Biochimica,
Complesso Universitario di Monte Sant'Angelo
Via Cinthia 4
80126 Napoli (Italy).
Tel +39-081674314
Fax +39081674313;
email: anduilio@unina.it

INTRODUCTION

The availability of complete genome sequences for a constantly growing number of organisms has opened the door for novel proteomic approaches aimed at identifying components of complex protein machineries and thus to elucidate their function. Moreover, microarray technologies (1) providing information on global mRNA expression, clearly demonstrated that in many cases mRNA transcripts do not directly correlate with protein expression (2,3). Therefore, the field of proteomics is challenged with the task of providing both quantitative and functional data to further complement genomics (4).

In particular, functional proteomics has emerged as a new area of study aimed at isolating and identifying molecular components belonging to multi-protein complexes (5). It is now clear, in fact,

that fundamental biological mechanisms are carried out by a multitude of proteins that assemble to form large functional complexes (6) or specific cell pathways (7). A comprehensive description of cellular processes at the molecular level is then strictly dependent on the clear definition of the individual protein components involved in these functional entities. A key contribution to the disclosure of biological mechanisms can then be provided by functional proteomic strategies. Isolation of the entire multiprotein complexes can be essentially accomplished by either affinity based techniques (including "pull down" and "tap tag" experiments) (8-10) or immunoprecipitation procedures (11). Both methodologies take advantage from the availability of different (immuno)-affinity ligands endowed with high binding efficiency. The success of an affinity-based approach strictly depends on the absence of excessive nonspecific interactions that in turn is related to the specificity of the bait-partners recognition. When this binding specificity is extremely high, as in the case of DNA/RNA-binding proteins, a low level of false positives is expected.

Ribosomal RNA (rRNA) transcription constitutes a fundamental biological process in all living cells. Several studies demonstrated that transcription of rRNA is the rate-limiting step in ribosome synthesis (12-14). Consistent with this crucial role in ribosome biogenesis, the control of rRNA transcription initiation, the efficient elongation of RNA polymerase as well as all the other steps involved in the transcription machinery are tightly regulated by a complex network of protein components (14 and references therein). Ribosomal RNA accounts for >90% of total RNA present in the bacterial cell. Operons encoding ribosomal RNA are present in multiple copies in the genome, an unusual feature for bacterial genes; for instance, in *Escherichia coli* rRNA is encoded by 7 different operons. Ribosomal proteins are subject to various forms of regulation: proteins such as Fis or H-NS (15), as well as direct interaction between RNA polymerase and the ppGpp signal molecule (16,17) can modulate the activity of rRNA promoters. An important feature of ribosomal promoters is the presence of the so-called UP element, an A/T rich sequence located at 40 to 80 nucleotides upstream of the transcription start site that can stimulate transcription from the *rrnB* P1 promoter through direct interaction with the C-terminal domain of the RNA polymerase α subunit (18,19).

In this report, we have investigated the functional protein complex that gathers at the upstream elements of the ribosomal *rrnB* P1 promoter of *E. coli*. We find that several ribosomal proteins assemble at the *rrnB* P1 promoter, thus suggesting direct interaction between RNA polymerase and the ribosome. In addition, proteins involved in DNA repair/protection mechanisms are also found in the promoter-binding protein complex. Altogether, our findings strongly suggest that proteins with functions not directly related to transcription associate with RNA polymerase to form multiprotein complexes.

MATERIALS AND METHODS

Proteomic grade trypsin, dithiothreitol (DTT), HEPES, KCl, $MgCl_2$, glycerol, ammonium bicarbonate and triton were purchased from Sigma. Cyanogen bromide-activated agarose matrices were obtained from Pharmacia Biotech. All used solvents were of the highest purity available from Romil.

Escherichia coli K12 strain growth conditions and protein extract preparation.
E. coli K12 strain was grown in aerobic conditions at 37°C in LB medium (20). After 16 h, bacteria were harvested by centrifugation and resuspended in Buffer Z (25 mM HEPES pH 7.6, 50 mM KCl, 12.5 mM $MgCl_2$, 1 mM DTT, 20% glycerol, 0.1% triton) containing 1 μM phenyl methyl sulfonyl fluoride. Cells were disrupted by sonication. The suspension was centrifuged at 90,000 x g for 30 min at 4°C. After centrifugation the protein concentration of the extract was determined with Bradford assay (21).

Preparation of the oligonucleotide bait and coupling to Sepharose.
A 28-mer deoxynucleotide encompassing the upstream and the -35 sequence of the ribosomal promoter *rrnB* P1 and a randomised version of the same sequence were chemically synthesized (Table 1). The complementary oligonucleotides (500 μg) were annealed, 5'-phosphorylated and ligated as follows. Gel purified oligodeoxynucleotides were combined in 10 mM Tris pH 7.5, 50 mM NaCl, 1 mM EDTA pH 8. The total mixture was incubated at 90°C for 5 min, 65°C for 10 min, 37°C for 10 min and RT for 5 min. ATP (20 mM) and T4 polynucleotide kinase (200 units) were then added and the resulting solution was incubated at 37°C for 2 hours. The DNA was phenol extracted and then ethanol precipitated and the pellet was resuspended in H_2O and the ligation reaction was initiated by the addition of 20 mM ATP and T4 DNA ligase (15 units). The mixture was incubated at 16°C and the desired multimerization grade was monitored by agarose gel. The final DNA product was phenol extracted and ethanol precipitated. The polymerised oligonucleotides were covalently linked to cyanogen bromide-activated agarose beads following the procedure recommended by Amersham.

Electrophoretic shift mobility assays.
Annealed DNA oligonucleotides (fragments UP, UP35 and Neg) were ^{32}P labelled at 5' positions by using T4 polynucleotide kinase (20 units). Electromobility shift assays were performed in 20 μl reaction volume. Total protein extract (10 μg) or each individual recombinant protein (1 μg) were incubated with the ^{32}P-labelled DNA (2 ng, 50,000-100,000 cpm) in buffer Z, for 20 min at RT. Mixtures were then analysed by electrophoresis on 8% native polyacrylamide gel (29:1 cross-linking ratio) in 45 mM Tris pH 8.0, 45 mM boric acid, 1 mM EDTA. Electrophoreses were performed at room temperature at 200 V (20 V/cm). The gels were dried and analysed by autoradiography.

DNA affinity chromatography.
A 4 ml affinity chromatography column was prepared with the agarose derivatizated beads. The column was equilibrated with 30 ml of buffer Z. The entire cell extract (300 mg) was incubated with the bait overnight at 4°C. The column was washed with 50 ml of buffer Z and the proteins selectively retained from the DNA probe were eluted with the same buffer containing 1 M KCl. The fractions containing the proteins of interest eluted from the affinity chromatography were pooled and dialyzed in H_2O. The same affinity chromatography procedure was applied to the Neg oligonucleotide. Both the eluted and dialyzed fractions were loaded onto SDS-PAGE gel and visualized by colloidal Coomassie staining.

In situ digestion and MALDI analysis.
Protein bands stained with Coomassie brilliant blue were excised from the gel

and destained by repetitive washes with 0.1 M NH_4HCO_3 pH 7.5 and acetonitrile. Samples were reduced by incubation with 50 μl of 10 mM DTT in 0.1 M NH_4HCO_3 buffer pH 7.5 and carboxyamidomethylated with 50 μl of 55 mM iodoacetamide in the same buffer. Enzymatic digestion was carried out with trypsin (12.5 ng/μl) in 10 mM ammonium bicarbonate pH 7.8. Gel pieces were incubated at 4°C for 2 hours. Trypsin solution was then removed and a new aliquot of the digestion solution was added; samples were incubated for 18 hr at 37°C. A minimum reaction volume was used as to obtain the complete rehydratation of the gel. Peptides were then extracted by washing the gel particles with 10 mM ammonium bicarbonate and 1% formic acid in 50% acetonitrile at room temperature. The resulting peptide mixtures were desalted using ZipTip pipettes from Millipore, following the recommended purification procedure.

MALDI Analyses.
MALDI-TOF mass spectra were recorded using an Applied Biosystem Voyager DE-STR instrument. A mixture of analyte and matrix solution (alfa-cyano-hydroxycinnamic acid 10 mg/ml in 70% ACN, 0.1% citric acid, in MilliQ water) was applied to the metallic sample plate and dried at room temperature. Mass calibration was performed using external peptide standards. Raw data were analysed using the computer software provided with the instrument and reported as monoisotopic masses. Peptide masses of each digested protein were used to search for protein databases using the Mascot mass fingerprinting software (Matrix Science).

Liquid Chromatography Tandem Mass Spectrometry (LC/ MS/MS) analyses.
Tryptic peptide mixtures obtained from *in situ* digestions were also analysed by LC/MS/MS "on-line" using a 4000Q-Trap linear trap ion mass spectrometer (Applied Biosystems) coupled to an 1100 nano HPLC system (Agilent Technologies). The mixture was loaded on an Agilent reverse-phase pre-column cartridge (Zorbax 300 SB-C18, 5x0.3 mm, solvent 0.1% formic acid, loading time 5 min). Peptides were separated on a Agilent reverse-phase column (Zorbax 300 SB-C18, 150 mm), at flow rate of 0.3 μl/min with a 5 to 65% linear gradient in 60 min (A solvent 0.1% formic acid, 2% ACN in water; B solvent 0.1% formic acid, 2% water in ACN). Nanospray source was used at 2.5 kV with liquid coupling, with a declustering potential of 50 V, using an uncoated silica tip from NewObjectives (Ringoes, NJ) (O.D. 150 μm, I.D. 50 μm, Tip.Diameter. 15 μm). Spectra acquisition was based on a survey scan (EMS) using Linear Ion Trap scanned from m/z 400 to m/z 1600 at 4000 amu/sec. This scan mode was followed by an enhanced resolution experiment (ER) for the five most intense ions of interest and then by MS^2 acquisitions of the five most intense ions for peptide charge, from +2 to +3. MS^2 spectra were acquired using the best collision energy calculated on the bases of m/z values and charge state (rolling collision energy).
The acquired MS/MS spectra were transformed in Mascot generic file format and used for peptides identification with a licensed version of MASCOT, in a local database.

Comparison with IntAct data.
Validation of the proteomic results was performed by comparing our data with those provided by IntAct database (htpp://www.ebi.ac.uk/intact/search/do/search?search String={0}) linked in the corresponding Swiss-Prot entries (htpp://www.expasy.org/sprot/). Each protein candidate was submitted to the

IntAct databank obtaining a list of all proteins involved in interaction with the query protein.

Construction of expression vectors.
The bacterial strains and plasmids used in this work are all reported in Table 1. The *rluC*, *fabZ*, *tufA*, *srmB*, *aidB* and *hns* genes of *E. coli* K12 were amplified from host DNA by polymerase chain reaction (PCR) using the forward and reverse primers listed in Table 1. To obtain RluC, FabZ, EF-Tu and H-NS tagged with c-myc epitope, the corresponding amplification products were digested with *BamH*I and *Xho*I and cloned into the pET22b-c-myc vector, respectively. To obtain pET22b-c-myc, a *Nde*I/*BamH*I digested fragment corresponding to the c-myc epitope was inserted into the pET22b (+) expression plasmid (Novagen) linearized with *Nde*I and *BamH*I.
RluC, SrmB and H-NS tagged with T7 epitope were made by insertion of the corresponding amplified fragments into the *BamH*I and *Xho*I sites of pET28a (+) expression vector (Novagen). The plasmid pET28a-*aidB* was constructed by cloning *aidB* coding sequence into pET28a (+) digested with *BamH*I and *Hind*III.
All plasmids (Table 1) contain the coding sequence for the corresponding recombinant protein fused to a 6X histidine tag to facilitate protein purification by Ni^{2+} affinity chromatography. Plasmids construction was verified by automated DNA sequencing.

Production and purification of recombinant proteins.
The recombinant genes were separately expressed into the *E. coli* strain C41 (DE3) (22). For RluC, FabZ and SrmB production, the recombinant cells were grown in LB medium at 37°C without induction until the OD_{600} reached 3.0. The *aidB, tufA* and *hns* genes were expressed as follows: recombinant cells were grown at 37°C to an optical density at 600 nm of ~0.5, at which time 0.05 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) was added and the cultures were shifted down to 25°C for AidB production and to 28°C for EF-Tu and H-NS expression until the OD_{600} reached 3.0.
Selective antibiotics were used at the following concentrations: 100 μg/ml ampicillin; 50μg/ml kanamicin. After incubation, cells were harvested by centrifugation at 5000 x g for 15 min at 4°C, resuspended in 50 mM Na_2HPO_4, pH 7.4, disrupted by passage through a French Press and centrifuged at 14000 x g for 30 min at 4°C.
Recombinant proteins were purified by affinity chromatography on His-Select Nickel Affinity Gel fv (Sigma). The lysate was loaded onto His-Select Nickel Affinity Gel equilibrated with equilibration buffer (50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl, 10 mM imidazole). After 1 min of incubation at 4°C, the matrix was collected by centrifugation at 11000 x g for 1 min and washed 3 times with same equilibration buffer. The recombinant proteins were eluted with buffer containing 250 mM imidazole in 50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl.
Protein concentration was estimated with Bradford reagent (Bio-Rad protein assay) and protein content was checked by SDS-polyacrilamide gel electrophoresis (SDS-PAGE).

Co-immunoprecipitation and Western blotting.
For co-immunoprecipitations*, E. coli* strain C41 (DE3) was co-transformed with the following constructs:
pET28a-*aidB*/pET22b-c-myc-*fabZ*,
pET28a-*srmB*/pET22b-c-myc-*tuf*A,
pET28a-*srmB*/pET22b-c-myc-*rluC*,
pET28a-*rluC*/pET22b-c-myc-*tuf*A,

Table 1: Bacterial strains, plasmids and oligonucleotides

Strains/plasmids	Description	Reference or Source
Strains		
C41 (DE3)	Strain that derives from BL21 [F⁻ *ompT hsdSB* (r_B- m_B-) *gal dcm* (DE3)]. This strain has at least one uncharacterized mutation that prevents cell death associated with expression of many toxic recombinant proteins	ref. 22
Plasmids		
pET22b(+)	carries an N-terminal *pelB* signal sequence for potential periplasmic localization, plus an optional C-terminal His-tag sequence	Novagen
pET28a(+)	carries an N-terminal His•Tag®/thrombin/T7•Tag® configuration plus an optional C-terminal His•Tag sequence	Novagen
pET22b-c-myc	pET22bΔ(*Nde*I-*BamH*I)Ω(c-myc)	This work
pET22b-c-myc-*rluC*	pET22b-c-mycΔ(*BamH*I-*Xho*I)Ω(*rluC* gene)	This work
pET22b-c-myc-*fabZ*	pET22b-c-mycΔ(*BamH*I-*Xho*I)Ω(*fabZ* gene)	This work
pET22b-c-myc-*tufA*	pET22b-c-mycΔ(*BamH*I-*Xho*I)Ω(*tufA* gene)	This work
pET22b-c-myc-*hns*	pET22b-c-mycΔ(*BamH*I-*Xho*I)Ω(*hns* gene)	This work
pET28a-*aidB*	pET28aΔ(*BamH*I-*Hind*III)Ω(*aidB* gene)	This work
pET28a-*rluC*	pET28aΔ(*BamH*I-*Xho*I)Ω(*rluC* gene)	This work
pET28a-*srmB*	pET28aΔ(*BamH*I-*Xho*I)Ω(*srmB* gene)	This work
pET28a-*hns*	pET28aΔ(*BamH*I-*Xho*I)Ω(*hns* gene)	This work
Oligonucleotides		
UP Fw	5'-AGAAAATTATTTTAAATTTC-3'	
UP Rv	5'-GAAATTTAAAATAATTTTCT-3'	
UP35 Fw	5'-GAAAATTATTTTAAATTTCCTCTTGTCA-3'	
UP35 Rv	5'-TGACAAGAGGAAATTTAAAATAATTTTC-3'	
Neg Fw	5'-CAACTATAGTATAAGATATTCTTTCTTT-3'	
Neg Rv	5'-AAAGAAAGAATATCTTATACTATAGTTG-3'	
c-myc Fw	5'TTCATATGGAACAAAAACTCATCTCAGAAGAGGATCTGAATGGGGCCGCAGGATCCTAT3'	
c-myc Rv	5'ATAGGATCCTGCGGCCCCATTCAGATCCTCTTCTGAGATGAGTTTTTGTTCCATATGAAT3'	
rluC Fw	5'-ATAGGATCCATGAAAACAGAGACTCC-3'	
rluC-pET22b Rv	5'-ATACTCGAGTAAGCGCGCGTTACG-3'	
rluC-pET28a Rv	5'-ATACTCGAGTTAGCGCGCGTTACG-3'	
fabZ Fw	5'-ATAGGATCCTTGACTACTAACACTC-3'	
fabZ Rv	5'-TAACTCGAGTTAGGCCTCCCGGC-3'	
tufA Fw	5'-TATGGATCCGTGTCTAAAGAAAAATTTG-3'	
tufA Rv	5'-TATCTCGAGTAAGCCCAGAACTTTAGC-3'	
hns Fw	5'-TAAGGATCCATGAGCGAAGCACTTA-3'	
hns-pET22b Rv	5'-ATACTCGAGTAATTGCTTGATCAGG-3'	
hns-pET28a Rv	5'-ATACTCGAGTTATTGCTTGATCAGG-3'	
aidB Fw	5'-ATAGGATCCGTGCACTGGCAAACTCACACCG-3'	
aidB Rv	5'-CGCGAGCTCCGTAAGCTTTTACACACACACTC-3'	
srmB Fw	5'-ATAGGATCCATGACTGTAACGACTTTT-3'	
srmB Rv	5'-ATACTCGAGTTACTCTTCTGTCGTTTG-3'	

pET28a-*aidB*/pET22b-c-myc-*hns*,
pET28a-*hns*/pET22b-c-myc-*fabZ*,
pET28a-*hns*/pET22b-c-myc-*rluC*,
pET28a-*aidB*/pET22b-c-myc-*rluC*.
After expression of the recombinant genes without induction, cells were harvested, suspended in 50 mM Na_2HPO_4 (pH 7.4), disrupted by passage through a French Press and centrifuged at 14000 x g for 30 min at 4°C. The supernatants were used for the co-immunoprecipitation experiments.
Cell lysates (0.5 mg) were incubated with agarose-linked T7 antibody (Bethyl) and with agarose beads only (control of the experiment) at 4°C overnight. The beads were then collected by centrifugation. Precipitates were washed several times, the bound proteins were eluted with 1×SDS–PAGE sample buffer and subjected to SDS-PAGE followed by Western blot analysis that was performed by using anti-T7 mouse antibody (Novagen) and anti-c-myc mouse antibody (Calbiochem) as first antibodies and anti-mouse IgG conjugated to peroxidase as a secondary antibody (Calbiochem).

RESULTS

Interaction of Escherichia coli proteins with the rrnB P1 promoter upstream elements.

Escherichia coli cell extracts were tested for their ability to bind to the upstream element of the *rrnB* P1 promoter in electrophoresis mobility shift assays (EMSA). EMSA experiments were carried out using three different DNA oligonucleotides as probes. The first oligonucleotide (20 nucleotides, oligo "UP", Table 1) corresponds to the UP element alone, i.e. the DNA sequence directly involved in interaction with the α subunit of RNA polymerase and located immediately upstream of the -35 promoter element (23,24). The second oligonucleotide (28 nt, "UP35", Table 1) includes both the UP element and an additional 8 base pairs containing the TTGTCA sequence of the *rrnB* P1 -35 box. Finally, a third 28-nt oligonucleotide with the same nucleotide content as "UP35", but lacking any known promoter or regulatory elements ("Neg", Table 1) was also used as negative control in EMSA experiments (Table 1). Each probe was incubated with 10 μg of protein extracts from *E. coli* cells. At this protein concentration, a clear gel mobility shift band could only be detected with the "UP35" oligonucleotide, i.e. with the DNA fragment containing the upstream element and the -35 box (Fig. 1, lane 2). In contrast, no interaction was observed with either the "UP" DNA fragment containing only the upstream element or with the "Neg" oligonucleotide, suggesting that the presence of the -35 box is necessary for efficient protein binding at the concentration used (Fig. 1, lanes 1 and 3). The specific nature of protein binding to UP35 detected in EMSA was further confirmed by competition experiments: the radiolabelled "UP35" probe was incubated with the protein extract from *E. coli* cells and increasing quantities (100-500 fold) of either unlabelled "UP35" or "Neg" were added. Binding to the radiolabelled "UP35" by cell extract was reversed by an excess of unlabelled "UP35", while it was not affected by competition with identical concentrations of the "Neg" probe (Fig. 1, lanes 4-7).
These results demonstrated that the UP35 region could be recognised by specific protein factors and that binding to the upstream element was specific and dependent on the presence of the -35 box.

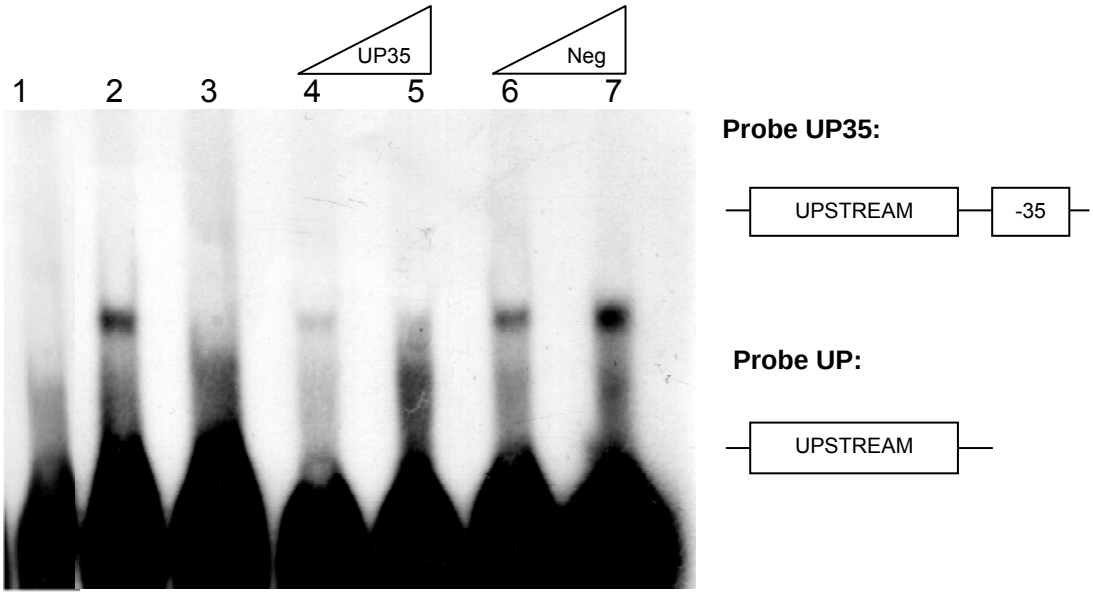


Figure 1. Electrophoresis mobility shift assays (EMSA) performed using either the UP or the UP35 elements and total protein extract from *E. coli* cells. Lane 1**:** *E. coli* protein extract incubated with UP; lane 2**:** *E. coli* protein extract incubated with UP35; lane 3**:** *E. coli* protein extract incubated with Neg; lane 4 - 5**:** Competition assay with UP35 (100x-500x); lane 6 - 7**:** Competition assay with Neg (100x-500x).

Identification of proteins specifically interacting with the oligonucleotide bait.

In order to identify which proteins are involved in binding to the *rrnB* P1 promoter upstream elements (UP element and -35 box), we performed functional proteomic assays using the "UP35" synthetic oligonucleotide linked to agarose beads ("baits"), as described in Materials and Methods. The total protein extract from *E. coli* K12 cells was first incubated with agarose beads to remove proteins that could bind the agarose matrix, then split in two fractions that were incubated with the agarose-linked DNA baits, i.e. the "UP35" and the "Neg" oligonucleotides, respectively. After extensive washing, the proteins specifically retained by the native and the random oligonucleotides were eluted with a strong ionic buffer containing 1 M KCl and visualized on SDS-PAGE. Figure 2 shows the corresponding Coomassie Blue stained gel displaying a number of discrete protein bands both in the sample incubated with "UP35" (lane A) and in the negative control (sample incubated with the "Neg" oligonucleotide, lane B). Proteins specifically interacting with the native double-strand DNA probe were selected by comparing the two electrophoretic patterns: bands solely occurring in the sample and absent in the control (samples 1 to 13 in Fig. 2) were selected for identification. In addition, we excised gel slices from lane B in the positions corresponding to the bands selected for identification (samples C1 to C13 in Fig. 2), and identified the proteins as well. Common proteins identified in both the sample and the control gel slices were eliminated, thus greatly decreasing the number of false positives.

Protein identification was carried out by MALDI-MS on peptides obtained by in-gel trypsin proteolytic digestion; peptides were assigned to specific proteins using the Mascot software.

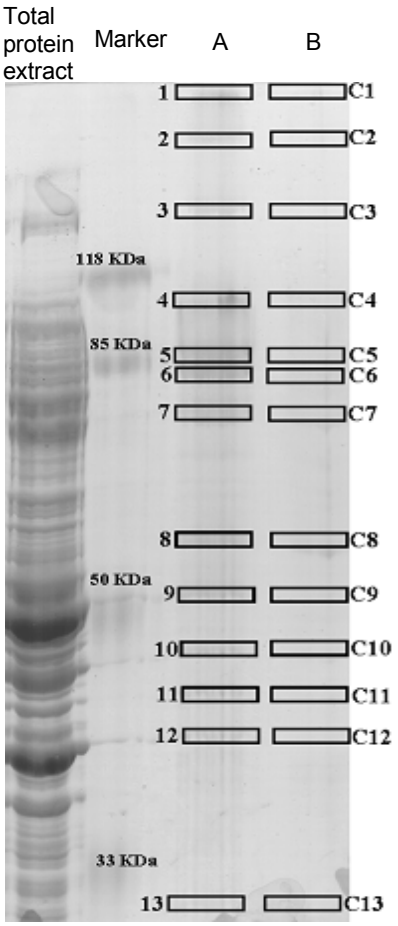


Figure 2. SDS PAGE fractionation of the protein complexes isolated by affinity chromatography using the UP35 bait. Coomassie Blue stained protein bands solely occurring in the sample (lane A) and absent in the control were submitted to mass spectral analyses (samples 1 to 13). Corresponding gel slices from the control lane B (samples C1 to C13) were also submitted to the identification procedure. Common proteins identified in both the sample and the control gel slices were eliminated.

Whenever the peptide mass fingerprinting procedure was unable to provide unambiguous identification of the proteins, the peptide mixtures were analysed by LC-MS/MS techniques generating sequence information on individual peptides leading to the identification of the protein components. The results obtained by the proteomic approach are summarised in Table 2. Proteomic analysis of the excised bands identified a total of 27 proteins solely present in the complex binding to the "UP35" oligonucleotide. In many cases, the peptides found in the mass spectral analyses led to a sequence coverage of about 50% or greater as for the elongation factor Tu or for tryptophanase. As expected, the α, β and β' subunits of the RNA polymerase were all found to be part of the "UP35"-complex. To our surprise, we could not find σ^{70} or any other σ factors, which are RNA polymerase subunits responsible for binding to core promoter elements such as the -35 sequence. Thus, it appears that, despite the need for the -35 sequence for specific binding (Fig. 1, lanes 1 and 2), RNA polymerase interaction with the "UP35" oligonucleotide mostly takes place via binding of the α subunit to the UP element. However, we cannot exclude that σ^{70} was not identified simply due to technical reasons. In addition to RNA polymerase, our results suggest that a large number of proteins can bind specifically to the *rrnB* P1 promoter region (Table 2). Not surprisingly, some of these additional proteins are either transcription regulators or known DNA-binding proteins (e.g. H-NS, Dps). However, a significant fraction of these additional proteins are part of the translation machinery: ribosomal proteins, proteins involved in RNA binding and modification (e.g. SrmB, DeaD) and in nascent polypeptides modifications (DnaJ), possibly suggesting that the protein synthesis machinery might assemble with RNA polymerase at the *rrnB* P1 promoter. Finally, some proteins are involved in cell energetic and catabolic metabolism (e.g. LpdA, TnaA), and their role as part of a transcription complex is less straightforward.

Table 2: Putative protein interactors identified by functional proteomic analysis

Gene name	Protein name	SwissProt protein identification number	Number of peptides found by MALDI-MS analysis	Number of peptides found by LC/MS/MS analysis
rpoB	RNA polymerase subunit beta (RNAP β)	**P0A8V2**		8
rpoC	RNA polymerase subunit beta' (RNAP β')	**P0A8T7**	22	
lon	ATP-dependent protease La	**P0A9M0**	21	
rnr	Ribonuclease R	**P21499**	22	12
deaD	Cold-shock DEAD box protein A (CsdA) (ATP-dependent RNA helicase DeaD)	**P0A9P6**		3
lpdA	Dihydrolipoyl dehydrogenase	**P0A9P0**	13	4
aidB	AidB	**P33224**	18	5
srmB	ATP-dependent RNA helicase SrmB	**P21507**		5
pssA	CDP-diacylglycerol--serine O-phosphatidyltransferase (Phosphatidylserine synthase)	**P23830**	17	
tnaA	Tryptophanase	**P0A853**	6	2
tufA	Elongation factor Tu (EF-Tu)	**P0A6N1**	17	3
rluC	Ribosomal large subunit pseudouridine synthase C (RluC)	**P0AA39**	33	781
ydbC	Putative oxidoreductase YdbC	**P25906**	9	185
rpoA	RNA polymerase subunit alpha	**P0A7Z4**	13	3
dnaJ	Chaperone protein DnaJ	**P08622**		3
fabZ	(3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase (FabZ)	**P0A6Q6**	15	9
rplI	50 S ribosomal protein L9	**P0A7R1**		6
rplQ	50 S ribosomal protein L17	**P0AG44**		4
rpsI	30 S ribosomal protein S9	**P0A7X3**		4
rplP	50 S ribosomal protein L16	**P0ADY7**		3
dps	DNA protection during starvation protein	**P0ABT2**		4
rpsF	30 S ribosomal protein S6	**P02358**		3
rpsL	30 S ribosomal protein S12	**P0A7S3**		2
rpsC	30 S ribosomal protein S3	**P0A7V4**		2
rpsG	30 S ribosomal protein S7	**P02359**		2
rpsE	30 S ribosomal protein S5	**P0A7W1**		2
hns	DNA-binding protein H-NS	**P0ACF8**		2

DNA binding properties of proteins part of the "UP35" binding complex.

In addition to recruitment by protein-protein interaction, proteins may be present in the complex by binding to the upstream elements of the *rrnB* P1 promoter through direct DNA binding. We tested some of the proteins identified by proteomic approach for their DNA binding ability, i.e. AidB, known to be a DNA-binding protein, two RNA-binding protein (EF-Tu and RluC) and FabZ, which is not known to bind any nucleic acid. To this aim, the corresponding genes, *aidB*, *tufA, rluC* and *fabZ* were cloned into a commercial expression vector of the pET series and the recombinant proteins were expressed as chimeric proteins bearing a 6 His tag.

The expressed proteins were then purified by affinity chromatography on Ni^{2+}-agarose beads and their homogeneity was tested by SDS-PAGE and mass fingerprinting analyses. EMSA experiments were carried out by incubating each individual protein with the radioactive "UP35" oligonucleotide probe, and the randomised version of the probe used as control (data not shown). As expected, addition of AidB retarded the migration of both the "UP35" DNA fragment and the "Neg" probe, thus confirming the ability of AidB to bind double stranded DNA in a non-specific fashion (25). In contrast, neither RluC nor FabZ and nor EF-Tu were able to bind the *rrnB* P1 promoter upstream region. This result is not surprising. It is conceivable that regulation of the *rrnB* operon is accomplished by a multi protein complex gathered at the promoter region in which only few proteins can effectively directly bind the DNA sequence whereas the complex is mainly stabilised by protein-protein interactions.

Validation of protein-protein interactions by database search.

In order to understand their role in the protein complexes formed at the "UP35" oligonucleotide, we verified protein-protein interactions by database search and by co-immunoprecipitation experiments. We searched each protein identified as part of the complex gathered at the upstream elements of the *rrnB* P1 promoter in the Swiss Prot databank (http://au.expasy.org/) and in related databases such as IntAct (http://www.ebi.ac.uk/intact/site/index.jsf). In addition to interactions among ribosomal proteins or RNA polymerase subunits, many examples of protein-protein interaction between transcription and translation machinery have already been reported; for instance, the β' subunit of RNA polymerase has been shown to interact with ribosomal proteins S5, S7, L17 (all found in our experiments) as well as with S4, S13, L2, L3, L4, L13 and L15. Even proteins involved in cell metabolism, such as FabZ (a β-hydroxyacyl-ACP dehydratase involved in fatty acid biosynthesis) are known to interact with ribosomal proteins (S3, S5, S7 and L9) as well as with Elongation factor Tu, another factor involved in protein synthesis. With the sole exceptions of tryptophanase (TnaA) and of the YdbC putative alcohol dehydrogenase, all other proteins were found to be involved in protein-protein interactions with one or more components of the "UP35" binding complex.

Validation of protein-protein interactions by co-immunoprecipitation experiments.

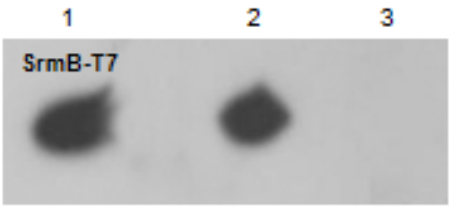
Databases screening provided initial support to our experimental results, suggesting that the proteins identified in the functional proteomic approach might indeed be involved in a complex network of interactions. However, in

order to validate the observed protein-protein interactions, we performed co-immunoprecipitation experiments, using the data obtained by database search to select which proteins to test in co-immunoprecipitation experiments. Initially, we focused on interactions between proteins belonging to the same functional groups; thus, we tested potential interactions between SrmB (an RNA helicase) with either EF-Tu or RluC (23S rRNA pseudouridine synthase), both RNA-binding proteins involved in either protein synthesis or ribosome biogenesis. T7-tagged SrmB was expressed in *E. coli* cells together with either c-myc tagged EF-Tu or RluC and the protein extract was immunoprecipitated by anti-T7 agarose-conjugated antibody, as described in Materials and Methods. The immunoprecipitate was run on SDS-PAGE and the SrmB protein was detected by Western blot using anti-T7 antibody (Fig. 3A, 3B, left panels). c-myc-labelled proteins that co-immunoprecipitate with T7-tagged SrmB could be detected by immunodetection using antibodies directed against the c-myc peptides (Fig. 3A, 3B, right panels). We could not detect any band when antibodies directed against either T7 or c-myc were used in Western Blot experiments using samples incubated with agarose beads alone (lane 3). The bands detected by Western blot in co-immunoprecipitates perfectly match c-myc tagged EF-Tu and RluC respectively used as control (lane 2). Analogous co-immunoprecipitation experiments showed interaction between RluC and EF-Tu (Fig. 3C). Thus, our findings strongly suggest that RluC, SrmB and EF-Tu can interact and might indeed be part of the same protein complex *in vivo*. Co-immunoprecipitation experiments were then performed to determine whether proteins identified as part of the "UP35"-binding protein complex, but belonging to different functional classes, could interact among each other. Indeed, we found that the transcription regulator H-NS was not only able to interact with AidB, a DNA-binding protein (Fig. 3D), but also with FabZ (involved in fatty acid metabolism, Fig. 3E) and with RluC (23S rRNA pseudouridine synthase, Fig. 3F). Likewise, AidB can establish protein-protein interactions with both FabZ (Fig. 3G) and RluC (Fig. 3H). These experimental results clearly indicated that the proteins identified in the functional proteomic experiments are involved in a complex network of interactions.

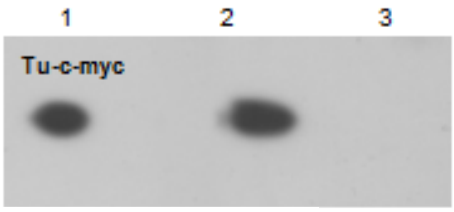
DISCUSSION

The ability of any organism to survive depends, in part, on mechanisms that enable it to modify its patterns of gene expression in response to extra- and intracellular signals. In the past, bacterial signalling was thought to function as a set of isolated, linear pathways. More recent studies, however, have demonstrated that many signalling pathways interact, thus forming an intricate network, which integrates both extracellular and intracellular signals, to ensure that the correct amount of the appropriate subset of genes is expressed at the proper time. Complete description of this complex signal transcription and transduction network and use of the network to predict the full range of cellular behaviours are major goals of system biology. One of the key issues in the contemporary genomic biology is to understand the structure and function of these cellular networks at different molecular levels. Among them, the transcriptional regulatory network (TRN) plays a crucial function, regulating cell physiology and metabolism through gene expression.

A

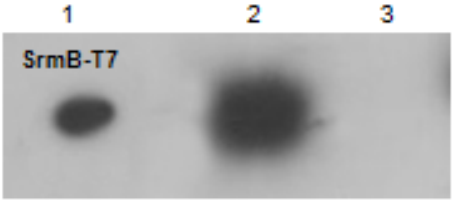


IP: anti-T7
WB: anti-T7

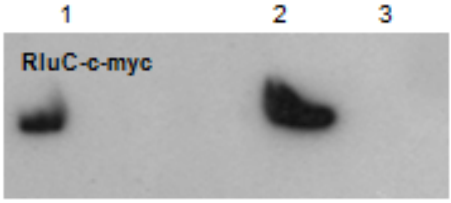


IP: anti-T7
WB: anti-c-myc

B

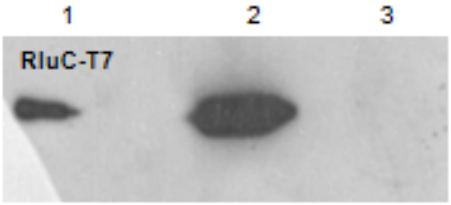


IP: anti-T7
WB: anti-T7

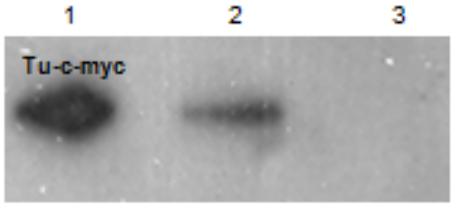


IP: anti-T7
WB: anti-c-myc

C

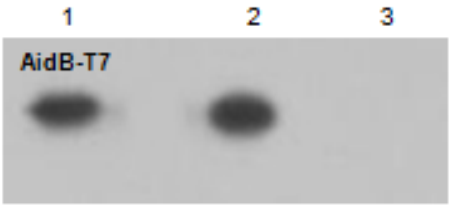


IP: anti-T7
WB: anti-T7

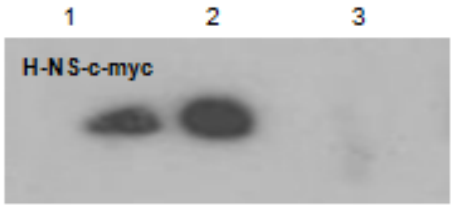


IP: anti-T7
WB: anti-c-myc

D



IP: anti-T7
WB: anti-T7



IP: anti-T7
WB: anti-c-myc

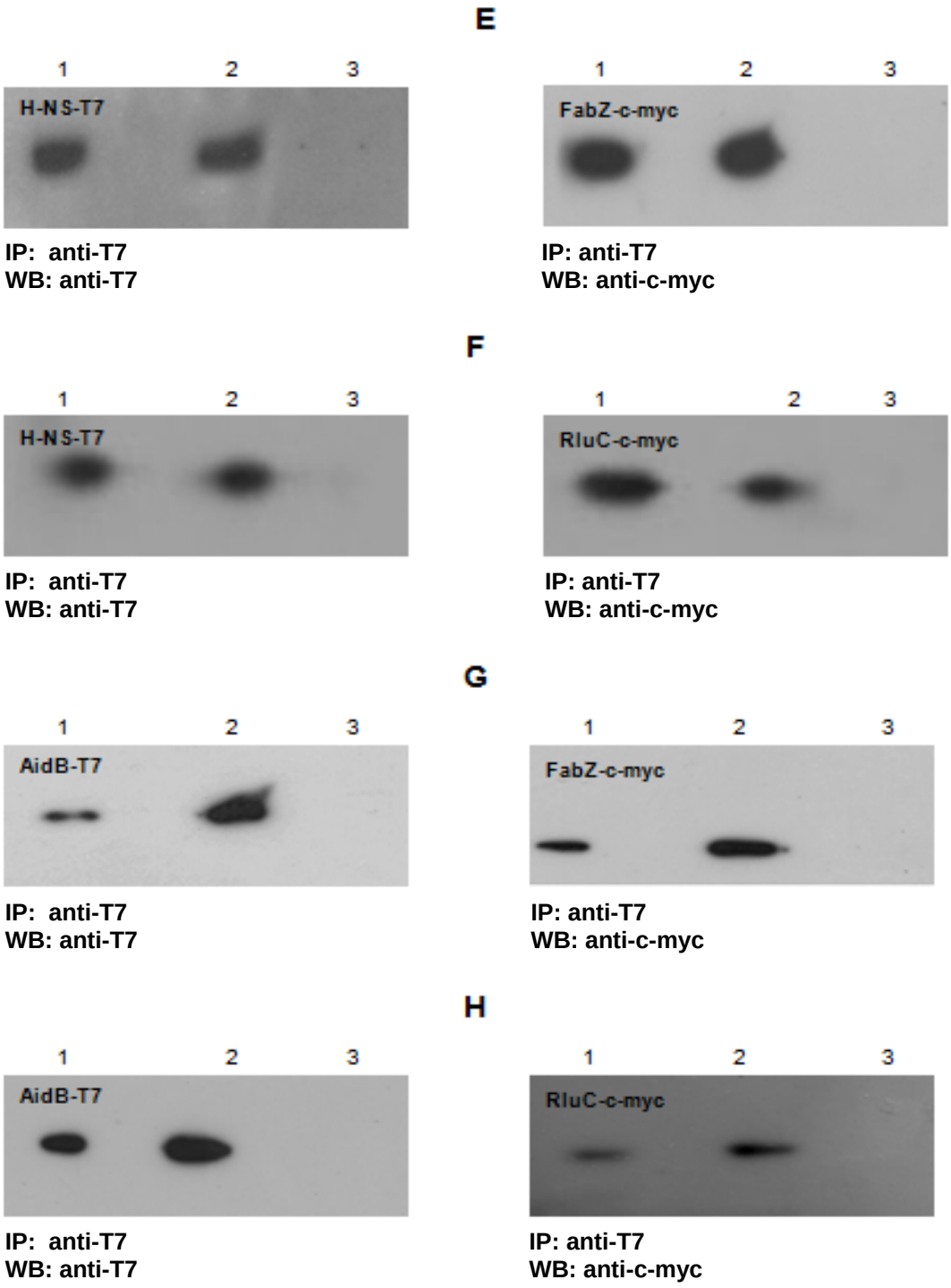


Figure 3. Co-immunoprecipitations experiments. Total protein extracts from *E. coli* strain C41 co-transformed with T7 and c-myc tagged version of individual proteins were subjected to immunoprecipitation with agarose-linked T7 antibody or with agarose beads only as control, followed by immunoblotting with anti-T7 and anti c-myc antibodies. In all cases, the left panel shows the control of the immunoprecipitation procedures (IP: anti-T7, Blot: anti-T7) whereas the right panel reports the interaction between the two proteins (IP: anti-T7, Blot: anti-c-myc).
In all the experiments, lane 1 corresponds to the fraction eluted from anti-T7 antibody beads, lane 3 shows the corresponding fraction eluted from agarose beads as negative control while lane 2 contains the purified T7 and/or c-myc tagged form of the individual proteins.

In bacterial cells, regulation of ribosomal RNA (rRNA), which accounts to more than 90% of total RNA, is crucial for metabolic and energetic balance. Consistent with such an important role, ribosomal RNA gene transcription is tightly regulated. However, it is likely that not all the factors controlling rRNA transcription initiation, processing and modification have been identified to date.

In the last decade, a considerable amount of attention has been given to an A+T-rich promoter sequence, the so-called UP element, located upstream of the core promoter region. In the *E. coli* ribosomal *rrnB* P1 promoter, the UP element is located in a region spanning the -40 and -60 positions, and is able to increase transcription from 30 to 70 fold by acting as a binding site for the α subunit of RNA polymerase (18). This element is a potential target for regulation of rRNA transcription by binding other putative factors able to modulate RNA polymerase activity. Indeed, a large number of promoters subjected to complex regulation possess strong α–binding sites (26-28).

In this work, we used a double stranded DNA oligonucleotide carrying the UP element and the -35 sequence of the *rrnB* P1 promoter as a bait to isolate the transcription complexes that gather at this promoter and to identify its components. Altogether, we were able to identify 27 different proteins by mass spectrometry methodologies. Interestingly, several of the identified proteins were already known to interact, either from data obtained from the literature, or from databases such as IntAct (http://www.ebi.ac.uk/intact/site/index.jsf), thus suggesting that they might be indeed involved in functional networks. According to their reported biological activities, the various promoter interactors can be grouped into three functional categories: DNA binding proteins and transcription factors, RNA-binding and translation factors, and cell metabolism-related proteins. As expected, we found the components of *E. coli* RNA polymerase, with the exception of σ^{70}, suggesting that, despite the presence of the -35 sequence, binding to the "UP35" oligonucleotide is mediated by core RNA polymerase, probably via interaction between the α subunit and the UP element. Despite the "UP35" oligonucleotide being an artifactual system, reports of RNA polymerase binding to promoter independent of σ^{70} interaction with core promoter elements would support this model (26,29 and unpublished data). However, it should be considered that σ^{70} might have also escaped identification in the proteomic experiment simply because of technical reasons.

An important component of the protein complex isolated at the upstream elements of the *rrnB* P1 promoter is represented by ribosomal proteins, which would suggest that the ribosome assembles with RNA polymerase at bacterial promoters. This model would be in agreement with various observations on direct protein-protein interaction between RNA polymerase subunits and ribosomal proteins (30,31), and with reports on multifunctional roles of protein synthesis factors (32). In addition, our results provide biochemical evidence that are consistent with known mechanisms of cross-talk between the ribosome and RNA polymerase, such as attenuation and stringent response. Indeed, direct interaction between transcription and translation machineries would greatly facilitate this genetic processes: for instance, ppGpp synthesis by ribosomal proteins RelA and SpoT would be relayed more

efficiently to RNA polymerase if both a ribosome and RNA polymerase are part of the same protein complex.
Enzymes active in cell metabolism (e.g. FabZ, LpdA), as well as proteins involved in RNA modification (e.g. RluC) or post-translational maturation of nascent peptides (e.g. DnaJ, La) might be recruited to the transcription complex either by transcription factors or by ribosomal proteins, as suggested by our co-immunoprecipitation experiments (Fig. 3).
Another interesting scenario is based on the nature of the DNA-binding proteins found in the "UP35"-binding complex. The H-NS protein is the only *bona fide* transcription regulator and it binds curved DNA regions with a strong preference for A/T rich stretches. The H-NS protein is a trans-acting regulatory factor that negatively modulates the transcription of several ribosomal RNA genes by binding to the upstream sequence of their respective promoters. This protein is overexpressed during the stationary phase when the synthesis of rRNAs has to be decreased (33). In contrast to H-NS, both AidB and Dps are non-specific DNA binding proteins (25,34) involved in the response to alkylation and oxidative damage, respectively. In this work, we suggest potential roles for the AidB protein in transcription complexes. AidB belongs to the adaptive response to alkylating agents (35,36 and references therein), which is defined by the Ada regulon including *ada* itself, *alkA*, *alkB* and *aidB* genes. The repair mechanism is governed by the Ada protein, that can repair O^6-methylated guanidine residues and acts upon self-methylation as a transcription activator of the *ada-alkB* operon and the *alkA* and *aidB* promoters (reviewed in 37). In contrast to the situation for DNA repair proteins Ada, AlkA, and AlkB, the role of AidB in the adaptive response is still uncharacterized. AidB has been proposed to repair alkylated DNA or detoxify alkylating reagents (25,38,39).
Findings reported here demonstrated that AidB can establish direct protein-protein interaction with other components of the *rrnB* P1 promoter binding complex, including H-NS. Interaction with H-NS might re-direct AidB binding towards curved DNA or A/T rich regions, thus recruiting AidB to transcription initiation complexes. The role of AidB might be related to repair of single strand DNA: indeed, single strand DNA has been demonstrated to undergo alkylation processes much quicker than double strand molecules even in the absence of oxidative stress (40). During transcription, the transcriptional fork generates single strand DNA molecules that might be a target for endogenous alkylating agents. In this respect, the recruitment of AidB and its related enzymes activities directly onto the promoter might prevent alkylation to damage the transcription process. Another intriguing, but perhaps more unlikely possibility could be that AidB might be recruited to this specific promoter to counteract alkylation of the nascent rRNA molecule. Both hypotheses are currently being investigated.

ACKNOWLEDGEMENTS
This work was supported by grants of Ministero dell'Università e della Ricerca Scientifica (Progetti di Rilevante Interesse Nazionale 2005, 2006; FIRB Rete Nazionale di Proteomica, RBRN07BMCT). Support from the National Center of Excellence in Molecular Medicine (MIUR - Rome) and from the Regional Center of Competence (CRdC ATIBB, Regione Campania – Naples) and from NIH grant CA100122 to M.R.V is gratefully acknowledged.

REFERENCES

1. Shalon,D., Smith,S.J., and Brown,P.O. (1996) A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res*., **6**, 639-645.
2. Anderson,L., and Seilhamer,J. (1997) A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis*, **18**, 533–537.
3. Gygi,S.P., Rochon,Y., Franza,B.R., and Aebersold,R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol*. *Cell*. *Biol*., **19**, 1720-1730.
4. Schmidt,F., Donahoe,S., Hagens,K., Mattow,J., Schaible,U.E., Kaufmann,S.H., Aebersold,R., and Jungblut,P.R. (2004) Complementary analysis of the *Mycobacterium tuberculosis* proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology. *Mol*. *Cell*. *Proteomics*, **3**, 24-42.
5. Monti,M., Orrù,S., Pagnozzi,D., and Pucci,P. (2005a) Functional proteomics. *Clin*. *Chim*. *Acta*, **357**, 140-150.
6. Michaud,G.A., and Snyder,M. (2002) Proteomic approaches for the global analysis of proteins. *Biotechniques*, **33**, 1308–1316.
7. Souchelnytskyi,S. (2002) Proteomics in studies of signal transduction in epithelial cells. *J*. *Mammary Gland Biol*. *Neoplasia*, **7**, 359 – 371.
8. Terpe,K. (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl*. *Microbiol*. *Biotechnol*., **60**, 523–533.
9. Reinders,J., Lewandrowski,U., Moebius,J., Wagner,Y., and Sickmann,A. (2004) Challenges in mass spectrometry-based proteomics. *Proteomics*, **4**, 3686-3703.
10. Burckstummer,T., Bennett,K.L., Preradovic,A., Schutze,G., Hantschel,O., Superti-Furga,G., and Bauch,A. (2006) An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nat*. *Methods,* **3**, 1013-1019.
11. Monti,M., Orrù,S., Pagnozzi,D., and Pucci,P. (2005b) Interaction proteomics. *Biosci Rep*., **25**, 45-56.
12. Schneider,D.A., Ross,W., and Gourse,R.L. (2003) Control of rRNA expression in *Escherichia coli*. *Curr*. *Opin*. *Microbiol*., **6**, 151-156.
13. Paul,B.J., Ross,W., Gaal,T., and Gourse,R.L. (2004) rRNA transcription in *Escherichia coli*. *Annu*. *Rev*. *Genet*., **38**, 749-770.
14. Dennis,P.P., Ehrenberg,M., and Bremer,H. (2004) Control of rRNA synthesis in *Escherichia coli*: a systems biology approach. *Microbiol*. *Mol*. *Biol*. *Rev*., **68**, 639-668.
15. Afflerbach,H., Schroder,O., and Wagner,R. (1999) Conformational changes of the upstream DNA mediated by H-NS and FIS regulate *E. coli rrnB* P1 promoter activity. *J. Mol. Biol*., **286**, 339–353.
16. Hernandez,J.V., and Cashel,M. (1995) Changes in conserved region 3 of *Escherichia coli* σ70 mediate ppGpp-dependent functions *in vivo*. *J*. *Mol*. *Biol*., **252**, 536–549.
17. Barker,M.M., Gaal,T., Josaitis,C.A., and Gourse,R.L. (2001) Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation *in vivo* and *in vitro*. *J*. *Mol*. *Biol*., **305**, 673–688.
18. Ross,W., Gosink,K.K., Salomon,J., Igarishi,K., Zou,C., Ishihama,A., Serinov,K., and Gourse,R.L. (1993) A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science*, **262**, 1407-1413.
19. Meng,W., Belyaeva,T., Savery,N.J., Busby,S.J., Ross,W.E., Gaal,T., Gourse,R.L., and Thomas,M.S. (2001) UP element-dependent transcription at the *Escherichia coli rrnB* P1 promoter: positional requirements and role of the RNA polymerase α subunit linker. *Nucleic Acids Res*., **29**, 4166-4178.

20. Sambrook, J., and Russell, D.W. (2001) *Molecular Cloning: a Laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
21. Bradford,M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal*. *Biochem*., **72**, 248-254.
22. Miroux,B., and Walker,J.E. (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J*. *Mol*. *Biol*., **19**, 289-298.
23. Ross,W., Schneider,D.A., Paul,B.J., Mertens,A., and Gourse,R.L. (2003) An intersubunit contact stimulating transcription initiation by *E coli* RNA polymerase: interaction of the α C-terminal domain and σ region 4. *Genes Dev*., **17**, 1293-1307.
24. Haugen,S.P., Ross,W., and Gourse,R.L. (2008) Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat*. *Rev*. *Microbiol*., **6**, 507-519.
25. Rohankhedkar,M.S., Mulrooney,S.B., Wedemeyer,W.J., and Hausinger,R.P. **(**2006) The AidB component of the *Escherichia coli* adaptive response to alkylating agents is a flavin-containing, DNA-binding protein. *J*. *Bacteriol*., **188**, 223-230.
26. Landini,P., and Volkert,M.R. (1995) RNA polymerase α subunit binding site in positively controlled promoters: a new model for RNA polymerase-promoter interaction and transcriptional activation in the *Escherichia coli ada* and *aidB* genes. *EMBO J*., **14**, 4329-4335.
27. Ross,W., Aiyar,S.E., Salomon,J., and Gourse,R.L. (1998) *Escherichia coli* promoters with UP elements of different strengths: modular structure of bacterial promoters. *J*. *Bacteriol*., **180**, 5375–5383.
28. Boucher,P.E., Maris,A.E., Yang,M-S., and Stibitz,S. (2003) The response regulator BvgA and RNA polymerase α subunit C-terminal domain bind simultaneously to different faces of the same segment of promoter DNA. *Mol*. *Cell*., **11**, 163–173.
29. Landini,P., Bown,J.A., Volkert,M.R., and Busby,S.J. (1998) Ada protein-RNA polymerase sigma subunit interaction and alpha subunit-promoter DNA interaction are necessary at different steps in transcription initiation at the *Escherichia coli ada* and *aidB* promoters. *J*. *Biol*. *Chem*., **273**, 13307-13312.
30. Butland,G., Peregrin-Alvarez,J.M., Li,J., Yang,W., Yang,X., Canadien,V., Starostine, A., Richards,D., Beattie,B., Krogan, N., *et al*. (2005) Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature*, **433**, 531-537.
31. Arifuzzaman,M., Maeda,M., Itoh,A., Nishikata,K., Takita,C., Saito,R., Ara,T., Nakahigashi,K., Huang,H.C., Hirai,A. *et al*. (2006) Large-scale identification of protein–protein interaction of *Escherichia coli* K-12. *Genome Res*., **16**, 686-691.
32. Hausmann,C.D., and Ibba,M. (2008) Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. *FEMS Microbiology Reviews*, **32**, 705-721.
33. Gralla,J.D. (2005) *Escherichia coli* ribosomal RNA transcription: regulatory roles for ppGpp, NTPs, architectural proteins and a polymerase-binding protein. *Mol*. *Microbiol*., **55**, 973-977.
34. Martinez,A., and Kolter,R. (1997) Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. *J*. *Bacteriol*., **179**, 5188–5194.
35. Volkert,M.R. (1988) Adaptive response of *Escherichia coli* to alkylation damage. *Environ*. *Mol*. *Mutagen*., **11**, 241-255.
36. Sedgwick,B., Bates,P.A., Paik,J., Jacobs,S.C., and Lindahl,T. (2007) Repair of alkylated DNA: Recent advances. *DNA Repair*, **6**, 429-442.

37. Volkert,M.R., and Landini,P. (2001) Transcriptional responses to DNA damage. *Curr*. *Opin*. *Microbiol*., **4**, 178-185.
38. Landini,P., Hajec,L.I., and Volkert,M.R. (1994) Structure and transcriptional regulation of the *Escherichia coli* adaptive response gene *aidB*. *J*. *Bacteriol*., **176**, 6583-6589.
39. Bowles,T., Metz,A.H., Quin,J.O., Wawrzak,Z., and Eichman,B.F. (2008) Structure and DNA binding of alkylation response protein AidB. *PNAS*, **105**, 15299-15304.
40. Bodell,W.J., and Singer,B. (1979) Influence of hydrogen bonding in DNA and polynucleotides on reaction of nitrogens and oxygens toward ethylnitrosourea. *Biochemistry*, **18**, 2860-2863.

Role of *Escherichia coli* AidB protein in the transcriptional regulation

Valentina Rippa[1, 2], Carla Esposito[1], Michael Volkert[3] and Angela Duilio[1, 2*]

[1]Department of Organic Chemistry and Biochemistry, University Federico II of Naples, Naples, Italy
[2]School of Biotechnological Sciences, University Federico II of Naples, Naples, Italy
[3]Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts, USA

ABSTRACT

Upon exposure to alkylating agents, *Escherichia coli* increases expression of *aidB* along with three genes (*ada*, *alkB*, and *alkA*) that encode DNA repair proteins. Despite extensive efforts, the molecular mechanism by which AidB protects against the mutagenic effects of DNA alkylators remains elusive. In this study, we assigned a biological role to AidB protein in the cell not exposed to alkylating agents. AidB was shown to possess a high affinity for *ada/alkB*, *alkA* and *aidB* promoters containing an upstream element, an AT rich transcriptional enhancer sequence. The physiological significance of this specific binding was investigated by *in vivo* transcription assays. The results clearly demonstrated that AidB represses its own synthesis during normal cell growth. Another objective of this study was to define and characterize the domain architecture of AidB. The N-terminal region consisting of the first 439 residues was shown to exhibit isovaleryl-CoA dehydrogenase activity and to purify as a tetramer; the C-terminal region corresponding to residues 440-541 was determined to possess DNA binding activity, to purify as a monomer and to function as a transcriptional repressor *in vivo*. Thus, we demonstrated that the AidB protein is a modular transcription factor which requires the short C-terminal region for its regulatory function.

Keywords: AidB protein, adaptive response genes, upstream element, transcription regulation.

Running title: Transcriptional role of AidB protein in *Escherichia coli.*

*Corresponding Author:
Angela Duilio
Dipartimento di Chimica Organica e Biochimica,
Complesso Universitario di Monte Sant'Angelo
Via Cinthia 4
80126 Napoli (Italy).
Tel +39-081674314
Fax +39081674313;
email: anduilio@unina.it

INTRODUCTION

Transcription regulation is one of the principal strategies used by bacteria to respond to external stimuli and to adapt to a changing environment. Exposure of *Escherichia coli* to sublethal concentrations of alkylating agents such as methyl methanesulfonate (MMS) stimulates the expression of four genes, *ada*, *alkB*, *alkA*, and *aidB*. The activation of these genes confers increased cellular resistance to the mutagenic and cytotoxic effects of alkylating agents and is known as the adaptive response (Karran, P. *et al*., 1982; Kataoka, H. *et al*.,1983; Nakabeppu, Y. *et al*., 1984; Volkert, M.R. and Nguyen, D.C. 1984). The Ada protein is the key enzyme of this process; Ada acts both as a methyltransferase able to remove methyl groups from damaged DNA and as a transcriptional activator for the

adaptive response genes (Teo, I. *et al*., 1986; Nakabeppu, Y. *et al*., 1984; Landini, P. and Volkert, M.R. 1995). AlkA is a DNA glycosylase that catalyzes the base excision repair of alkylpurines (Nakabeppu, Y. *et al*., 1984). AlkB is an α-ketoglutarate-Fe(II)-dependent DNA dioxygenase that repairs 1-methyladenine and 3-methylcytosine lesions by oxidative demethylation (Trewick, S.C. *et al*., 2002). Despite detailed understanding of Ada, AlkA, and AlkB, the mechanism by which AidB protects against DNA damage in the adaptive response remains elusive. It has been suggested that AidB may act as a detoxification enzyme to destroy certain alkylating reagents or as a DNA repair enzyme (Landini, P. *et al*., 1994; Landini, P. and Volkert, M.R. 1995).

AidB is a protein of 541 amino acids that is related in sequence to acyl-CoA dehydrogenases (ACADs), a family of enzymes which use a flavin adenine dinucleotide (FAD) to catalyze the α,β-dehydrogenation of acyl-CoA conjugates (Landini, P. *et al*., 1994; Rohankhedkar, M.S., *et al*., 2006). Recently, AidB has been shown to bind stoichiometric amounts of redox active FAD and to exhibit weak isovaleryl-CoA dehydrogenase (IVD) activity (Landini, P. *et al*., 1994; Rohankhedkar, M.S., *et al*., 2006). Subsequent structural studies revealed a distinctive FAD active site that provides a rationale for AidB's limited acyl-CoA dehydrogenase activity. AidB was also shown to bind double strand DNA independently of its sequence (Rohankhedkar, M.S., *et al*., 2006). However, recently we demonstrated that AidB is part of the protein complex that binds to the -35 and upstream elements of the ribosomal *rrnB* P1 promoter but not to sequences lacking the UP element (Amoresano *et al*., submitted for publication), an AT rich transcriptional enhancer sequence. This result suggested that AidB might possess sequence-specific binding. Since the -60 to -40 regions of *ada/alkB*, *alkA* and *aidB* promoters resemble the *rrnB* P1 UP element in A/T content and location relative to the core promoter (Landini, P. *et al*., 1998), we supposed that AidB protein might specifically interact with the promoters of the Ada-regulated adaptive response.

In this work, we show that AidB displays a high affinity for UP regions of P*ada*/*alkB*, P*alkA* and P*aidB*. The functional significance of this specific binding was investigated by *in vivo* transcription assays using *lacZ* as reporter gene. AidB was shown to repress its own synthesis during normal cell growth.

Another objective of this study was to define and characterize the domain architecture of AidB. The N-terminal region consisting of the first 439 residues was shown to exhibit isovaleryl-CoA dehydrogenase activity; the C-terminal region corresponding to residues 440-541 was determined to possess DNA binding activity and to function as a transcriptional repressor *in vivo*. Thus, the results reported here showed that the AidB protein is a modular transcription factor which requires the short C-terminal region for its regulatory function.

MATERIALS AND METHODS

Bacterial strains and plasmids.

The bacterial strains, the plasmids and the oligonucleotides used in this work are listed in Table 1. Plasmids pMV132H, pSL101 and pSL112 were a kind gift from P. Landini (University of Milan, Italy). MV1161, MV5924, MV1601, MV6608 MV1571 and MV6607 *E. coli* strains were a kind gift from M. Volkert (University of Massachusetts, Worcester, MA).

Media and chemicals.
Luria-Bertani or nutrient broths (for bacterial cultures and plating) and suspension medium (for bacterial dilutions) were used as described by Miller (Miller, J.H. 1972). Ampicillin and kanamycin (Sigma) were used at 100 and 50 μg/ml, respectively.

Construction of the expression plasmids.
The *aidB*, *aidB*ΔCt and Ct-*aidB* genes were amplified from DNA genomic of *E. coli* K12 by polymerase chain reaction (PCR) using the forward and reverse primers listed in Table 1. The amplified fragments were digested with the restriction enzymes underlined in Table 1 and cloned into the expression vector pET28a (Novagen) which was linearized with the same restriction enzymes. The resulting plasmids, designated as listed in Table 1, were verified by automated DNA sequencing.
For the complementation experiments, the *aidB*, *aidB*ΔCt and Ct-*aidB* genes were cloned into the expression vector pET28a-P*lac*. The *lac* promoter was amplified from DNA genomic of *E. coli* by PCR using the primers listed in Table 1. The amplification product was digested with *Sph*I and *BamH*I and cloned into the pET28a (+) plasmid. Then the *aidB*, *aidB*ΔCt and Ct-*aidB* genes were positioned downstream of the *lac* promoter. Plasmid construction was verified by automated DNA sequencing.

Production and purification of recombinant proteins.
The constructs pET28a-*aidB*, pET28a-*aidB*ΔCt, pET28a-Ct-*aidB* (Table 1) were individually transformed into the *E. coli* strain C41 (DE3). The recombinant cells were grown at 25°C to an optical density at 600 nm of ~0.5, at which time 0.05 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) was added and the cultures were allowed to grow until the OD_{600} reached 3.0. Cells were harvested by centrifugation at 5000 x g for 15 min at 4°C, resuspended in buffer A (50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl, 10 mM imidazole, 1 mM PMSF), disrupted by passage through a French Press and centrifuged at 14000 x g for 30 min at 4°C.

Table 1: Bacterial strains, plasmids and oligonucleotides

Strains/plasmids	Description	Reference or source
***E. coli* strains**		
C41 (DE3)	Strain that derives from BL21 [F^- *ompT hsdS*$_B$ (r_B- m_B-) *gal dcm* (DE3)]. This strain has at least one uncharacterized mutation that prevents cell death associated with expression of many toxic recombinant proteins	Miroux and Walker, 1996
MV1161	*rfa-550* derivative of AB1157 that carries the following additional markers: *argE3 hisG4 leu-6 proA2 thr-1 ara-14 galK2 lacY1 mtl-1 xyl-5 rpsL31 supE44 tsx-6 rfa-550*	Volkert
MV5924	Strain derivative of MG1655 in which the *aidB* gene had been inactivated by a tetracycline resistance cassette	Volkert
MV1601	Strain carrying a *lacZ* transcriptional insertion within the chromosomal *alkB* gene	Volkert
MV6608	Strain harboring the *aidB* mutation in the a*lkB*::*lacZ* background	Volkert
MV1571	Strain harboring *lacZ* fragment in the chromosomal *alkA* gene	Volkert
MV6607	Strain harboring the *aidB* mutation in the a*lkA*::*lacZ* background	Volkert

pET28a(+)	Plasmid carrying an N-terminal His•Tag®/thrombin/T7•Tag® configuration plus an optional C-terminal His•Tag sequence	Novagen
pET28a-*aidB*	pET28aΔ(*BamH*I-*Hind*III)Ω(*aidB* gene)	This work
pET28a-*aidB*ΔCt	pET28aΔ(*BamH*I-*Hind*III)Ω(*aidB*ΔCt gene)	This work
pET28a-Ct-*aidB*	pET28aΔ(*BamH*I-*Xho*I)Ω(Ct-*aidB* gene)	This work
pET28a-P*lac*	pET28aΔ(*Sph*I-*BamH*I)Ω(*lac* promoter)	This work
pET28a-P*lac*-*aidB*	pET28a-P*lac*Δ(*BamH*I-*Hind*III)Ω(*aidB* gene)	This work
pET28a-P*lac*-*aidB*ΔCt	pET28a-P*lac*Δ(*BamH*I-*Hind*III)Ω(*aidB*ΔCt gene)	This work
pET28a-P*lac*-Ct-*aidB*	pET28a-P*lac*Δ(*BamH*I-*Xho*I)Ω(Ct-*aidB* gene)	This work
pMV132H	Plasmid carrying *lacZ* fusion under the control of the wild type P*aidB*	Landini
pSL101	Plasmid derivative of pJCD01 carrying a 242 bp fragment of the wild type P*aidB*	Landini
pSL112	Plasmid derivative of pJCD01 carrying a 238 bp fragment of the mutant P*aidB* in which -12C has been substituted with the residue T	Landini

Oligonucleotides

UP35 P*aidB* Fw	5'-GATAAGAATGTTTTAGCAATCTCTTTCTGTCA-3'
UP35 P*aidB* Rv	5'-TGACAGAAAGAGATTGCTAAAACATTCTTATC-3'
UP35 P*ada* Fw	5'-GCGAAAAAAATTAAAGCGCAAGATTGTTGGTT-3'
UP35 P*ada* Rv	5'-AACCAACAATCTTGCGCTTTAATTTTTTTCGC-3'
UP35 P*alkA* Fw	5'-TTGCCGTCGCGACAACCGGAATATGAAAGCAA-3'
UP35 P*alkA* Rv	5'-TTGCTTTCATATTCCGGTTGTCGCGACGGCAA-3'
Neg P*aidB* Fw	5'-ATGAATGTACTGTATCTAGTCTACTAGTATCT-3'
Neg P*aidB* Rv	5'-AGATACTAGTAGACTAGATACAGTACATTCAT-3'
Neg P*ada* Fw	5'-TAGACTTGATCAGATGAGATGAAGTACAGATA-3'
Neg P*ada* Rv	5'-TATCTGTACTTCATCTCATCTGATCAAGTCTA-3'
Neg P*alkA* Fw	5'-AGTCACTGTCAGTCACAGTACAGACGACGAGA-3'
Neg P*alkA* Rv	5'-TCTCGTCGTCTGTACTGTGACTGACAGTGACT-3'
P*leu*A Fw	5'-GGGTTGACATCCGTTTTTGTATCCAGTAACTC-3'
P*leu*A Rv	5'-GAGTTACTGGATACAAAAACGGATGTCAACCC-3'
aidB Fw	5'-ATAGGATCCGTGCACTGGCAAACTCACACCG-3'
aidB Rv	5'-CGCGAGCTCCGTAAGCTTTTACACACACACTC-3'
*aidB*ΔCt Fw	5'-ATAGGATCCGTGCACTGGCAAACTC-3'
*aidB*ΔCt Rv	5'-AATAAGCTTTTAGAGAACGCGCAACACATC-3'
Ct-*aidB* Fw	5'-ATAGGATCCAATAAGCAAGCGGGCG-3'
Ct-*aidB* Rv	5'-TTACTCGAGTTACACACACACTCCCC-3'
P*lac* Fw	5'-TAAAGCATGCTTTACACTTTATGCTTCCGGTCGTATGTTGTGTGGAAAGCTTTTA-3'
P*lac* Rv	5'-TAAAAGCTTTCCACACAACATACGACCGGAAGCATAAAGTGTAAAGCATGCTTTA-3'

The recombinant proteins were purified by affinity chromatography on His-Select Nickel Affinity Gel (Sigma). The lysate was loaded onto His-Select Nickel Affinity Gel equilibrated with equilibration buffer (50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl, 10 mM imidazole). After 1 min of incubation at 4°C, the matrix was collected by centrifugation at 11000 x g for 1 min and washed 3 times with same equilibration buffer. The recombinant proteins were eluted with buffer containing 250 mM imidazole in 50 mM Na_2HPO_4, pH 7.4,

0.3 M NaCl. Protein concentration was estimated with Bradford reagent (Bio-Rad protein assay) and protein content was checked by SDS-polyacrilamide gel electrophoresis (SDS-PAGE).

In vivo transcription assays.
MV1161 (wild type) and MV5924 (Δ*aidB*) *E. coli* strains were transformed with reporter plasmid pMV132H carrying the *lacZ* gene under the control of P*aidB*. The complementation experiments were performed by transforming MV5924 containing pMV132H with the following constructs: pET28a-P*lac*-*aidB*, pET28a-P*lac*-*aidB*ΔCt, pET28a-P*lac*-Ct-*aidB*. These bacterial cultures grown overnight in LB medium at 37°C, were diluted 1:100 in fresh medium. At an $A_{600\ nm}$ of 0.4, the cultures were divided in two aliquots, and one was supplemented with 0.04% MMS to activate the adaptive response. Cellular pellets were collected during the exponential growth phase. The cells were resuspended in 50 mM Na_2HPO_4 (pH 7.4), disrupted by passage through a French Press and centrifuged at 14000 x g for 30 min at 4°C. The supernatant was collected and protein concentration was determined with the Bio-Rad protein assay (Bradford, M.M. 1976), using bovine serum albumine as standard. β-galactosidase activity was determined by measuring ONPG-hydrolysis, as described by Miller (Miller, J.H., 1972). *In vivo* transcription from the *ada* promoter region was measured using both MV1601 and MV6608 strains; to measure *in vivo* transcription from the *alkA* promoter we used both MV1571 and MV6607 strains.

In vitro transcription assays.
Transcription experiments were performed on two different DNA templates: plasmid pSL101 containing wild type P*aidB* and plasmid pSL112 carrying mutant P*aidB* with a C-T substitution at position -12. Plasmids (5 nM) and RNA polymerase holoenzyme (120 nM) (Epicentre Technology) were incubated for 20 min at 37°C, in a solution of 40 mM Hepes pH 8.0, 10 mM magnesium chloride, 200 mM potassium glutamate, 4 mM dithiothreitol and 100 µg/ml bovine serum albumin in the absence and in the presence of increasing quantities of the AidB protein (0.1-0.4 µg). Elongation step was started by the addition of a pre-warmed mixture containing nucleotides and heparin (final concentrations were 500 µM ATP, GTP and CTP, 30 µM UTP, 1 µCi of [α^{32}P]UTP and 500 µg/ml heparin) to the template-polymerase mix and allowed to proceed for 10 min at 37°C. The reactions were stopped by the addition of stop solution (10 mM EDTA, 0.5% bromophenol blue, 0.025% xylene cyanol). After heating to 65°C, samples were subjected to electrophoresis on a 7% denaturing polyacrilamide gel in 0.5 X Tris borate-EDTA. Transcripts were detected by exposure to X-ray film overnight at -80°C.

Native molecular mass of AidBΔCt and Ct-AidB.
Size exclusion chromatography was performed by using a Superdex 200 PC 3.2/30 (for AidBΔCt) and a Superdex 75 PC 3.2/30 (for Ct-AidB) column (GE Healthcare) equilibrated in buffer containing 50 mM Tris-HCl, 150 mM NaCl, pH8. The molecular mass of the native proteins was estimated by comparing their retention time to those of molecular mass standards (thyroglobulin, 670,000 Da; bovine γ-globulin, 158,000 Da; chicken ovalbumin, 44,000 Da; equine myoglobin, 17,000 Da; vitamin B_{12}, 1,350 Da; Bio-Rad).

Isovaleryl-CoA dehydrogenase activity assay.
Isovaleryl-CoA dehydrogenase activity assays were carried out at room temperature in 200 mM phosphate buffer, pH 8.0, and using purified recombinant proteins that had been dialyzed to remove imidazole. For routine assays, 2 mM isovaleryl-CoA (Sigma) was used as the substrate and 0.1 mM 2,6-dichlorophenolindophenol (DCPIP) was used as the terminal electron acceptor in a final volume of 300 µl. The change in absorbance at 600 nm was monitored by using a Beckman DU 7500 spectrophotometer, and the enzyme activity was calculated by assuming an extinction coefficient of 20.6 mM^{-1} cm^{-1} for DCPIP (Engel, P.C., 1981).

Electrophoretic shift mobility assays.
Annealed DNA oligonucleotides (fragments UP35 P*aidB*, UP35 P*ada*, UP35 P*alkA*, Neg P*aidB*, Neg P*ada*, Neg P*alkA*) were ^{32}P labelled at 5' positions by using T4 polynucleotide kinase (20 units). Electromobility shift assays were performed in 20 µl reaction volume. AidB protein (2 pmol) was incubated with the ^{32}P-labelled DNA (2 ng, 50,000-100,000 cpm) in buffer Z (25 mM HEPES pH 7.6, 50 mM KCl, 12.5 mM $MgCl_2$, 1 mM DTT, 20% glycerol, 0.1% triton), for 20 min at RT. Mixtures were then analysed by electrophoresis on 8% native polyacrylamide gel (29:1 cross-linking ratio) in 45 mM Tris pH 8.0, 45 mM boric acid, 1 mM EDTA. Electrophoreses were performed at room temperature at 200 V (20 V/cm). The gels were dried and analysed by autoradiography.
For EMSA performed by using biotinylated UP35 P*aidB*, protein-DNA complexes were separated on 8% nondenaturing polyacrylamide gels. After electrophoresis, gels were transferred to nylon membranes and transferred DNA was immobilized by UV crosslinking for 3 min. For detection of bound DNA, membranes were blocked using blocking buffer followed by the addition of Streptavidin-HRPO and blots were developed according to the manufacturer's instructions (Amersham, Arlington Heights, IL, U.S.A.).

RESULTS
AidB shows a high affinity for upstream regulatory regions.
In a previous report (Rohankhedkar *et al*., 2006), it was shown that AidB binds double strand DNA independently of its sequence. However, our recent work demonstrated that AidB is part of the protein complex that gathers at the -35 and UP elements of the *rrnB* P1 promoter but not at identical sequences lacking the UP element (Amoresano *et al*., submitted for publication). This result suggested that AidB might possess sequence-specific binding. Since the -60 to -40 regions of *ada/alkB*, *alkA* and *aidB* promoters resemble the *rrnB* P1 UP element in A/T content and location relative to the core promoter (Landini *et al*., 1998), we supposed that AidB protein might specifically interact with the promoters of the Ada-regulated adaptive response. This hypothesis was investigated by electrophoresis mobility shift gel assays (EMSA). To this aim, the *aidB* gene was cloned into a commercial expression vector of the pET series and the recombinant protein was expressed as chimeric protein bearing a 6 His tag at the C-terminus. The expressed protein was then purified by affinity chromatography on Ni^{2+}-agarose beads and its homogeneity was tested by SDS-PAGE and mass fingerprinting analyses. EMSA experiments were carried out using three different

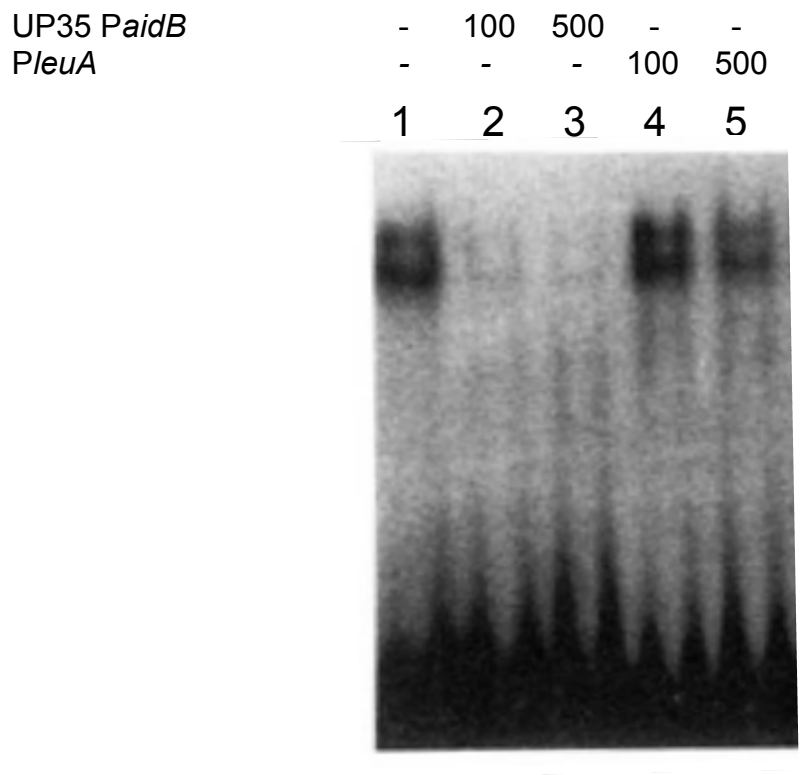


Fig. 1. Gel retardation experiments performed incubating the AidB protein with the UP35 P*aidB*; competitors were included as indicated. Lane 1**:** AidB protein incubated with UP35 P*aidB*. Lane 2-3**:** Competition assay with UP35 P*aidB* (100x-500x) as a specific competitor. Lane 4-5**:** Competition assay with P*leuA* as a non specific competitor (100x-500x).

radiolabelled DNA fragments corresponding to the UP region and the -35 box of the *aidB*, *ada/alkB* and *alkA* promoters ("UP35 P*aidB*", "UP35 P*ada*", "UP35 P*alkA*", Table 1). The randomised version of each UP35 probe, i.e. an oligonucleotide with the same nucleotide content as UP35, but lacking any known promoter or regulatory elements ("Neg P*aidB*", "Neg P*ada*", "Neg P*alkA*", Table 1), was also used in EMSA experiments. As expected, addition of AidB retarded the migration of both the UP35 DNA fragments and the "Neg" probes (data not shown). The specificity of AidB binding to the three UP35 probes was tested by competition experiments: the radiolabelled "UP35" probes were incubated with the AidB protein and increasing quantities (100-500 fold) of either unlabelled "UP35" (used as a specific competitor) or P*leuA* (a promoter lacking the UP element used as a non specific competitor), were added to the reaction mixture. The Fig. 1 shows that the binding to the radiolabelled "UP35 P*aidB*" by AidB was reversed by an excess of unlabelled "UP35" (Lanes 2-3), while it was not affected by competition with identical concentrations of P*leuA* (Lanes 4-5). Analogous experiments showed the specific interaction of AidB with "UP35 P*ada*" and "UP35 P*alkA*" (data not shown). Altogether, these results demonstrated that AidB possesses high binding affinity for the upstream regions of the adaptive response genes.

AidB acts as a transcriptional repressor in vivo.

In order to determine whether the

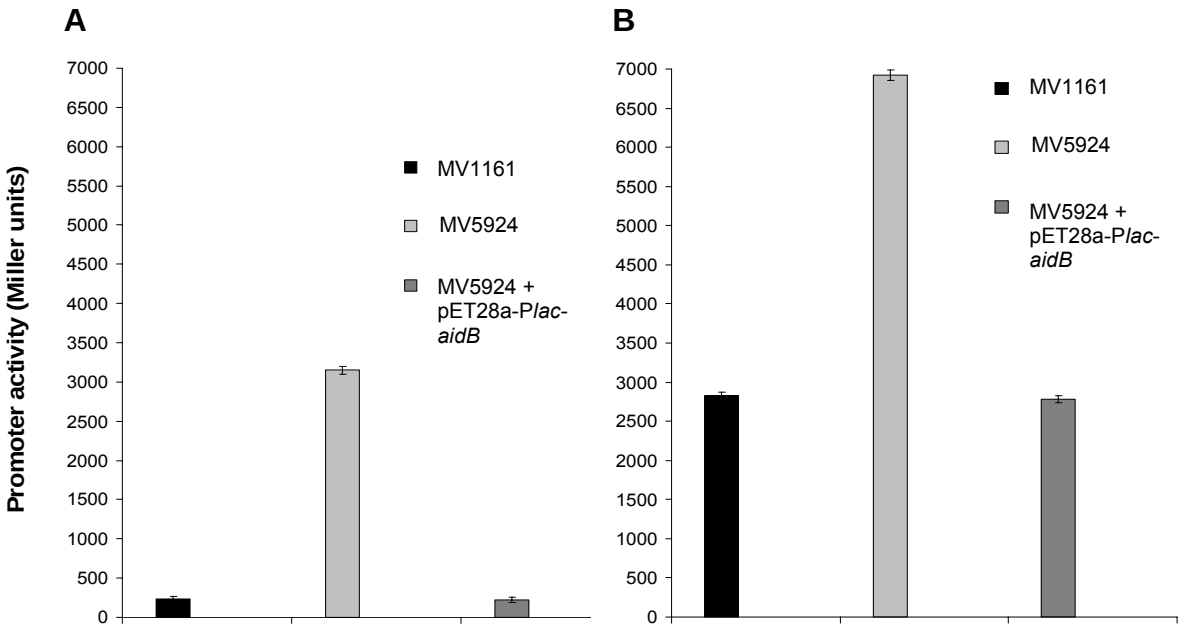


Fig. 2. *In vivo* transcription from the *aidB* promoter. Vector pMV132H was introduced into MV1161 (wild type) and MV5924 (Δ*aidB*) *E. coli* strains and the β-galactosidase specific activity was determined in the absence (A) and in the presence (B) of 0.04% MMS as alkylating agent. MV5924 strain was also transformed with plasmid pET28a-P*lac-aidB* producing functional AidB protein and the corresponding β-galactosidase activity was evaluated. Means and standard deviations have been calculated from four independent assays.

specific binding of AidB to P*aidB*, P*ada* and P*alkA* might be of biological significance, we tested its effect on transcription from these promoters by *in vivo* β-galactosidase reporter assays. The putative AidB activity on its promoter was investigated by transforming both MV1161 (wild type) and MV5924 (Δ*aidB*) *E. coli* strains with plasmid pMV132H carrying the gene coding for the β-galactosidase under the control of the *aidB* promoter. These strains were grown in LB medium, in the absence and in the presence of sub-inhibitory concentrations of the alkylating agent methyl methane sulfonate (MMS) and the β-galactosidase activity was monitored during the exponential growth phase. As shown in Fig. 2A, during the growth in LB medium, the Δ*aidB* strain exhibited levels of β-galactosidase activity 13-fold higher than that observed in the wild type cells, suggesting that transcription from the *aidB* promoter is repressed by AidB during normal cell growth. The complementation of the Δ*aidB* mutation by pET28a-P*lac-aidB* restored transcription levels from the *aidB* promoter to wild type levels. These data clearly demonstrated that in the absence of DNA alkylation damage, the AidB protein is required for repression of its own expression. Exposure to sub-inhibitory concentrations of MMS, as expected, resulted in an increase of *aidB* promoter activity (comparison between Figs. 2A and 2B). Interestingly, the transcription levels in the Δ*aidB* strain were slightly higher (by almost 2.5-fold) than that observed for the wild type cells suggesting that the activator methylated Ada, in the absence of AidB, is more efficient in activating transcription of the *aidB* gene. This effect could be due to the competition between AidB and methylated Ada for binding to same site on the *aidB*

promoter. In fact, gel retardation experiments performed in this work have demonstrated that AidB is able to bind to a region of P*aidB* extending from the upstream element to the -35 box and previous studies (Landini, P. and Volkert, M.R. 1995) have found that methylated Ada protects the residues from -62 to -38 in the *aidB* promoter.

Successively, we evaluated the AidB activity on the *ada* promoter (which controls expression of the *ada* and *alkB* genes) and the *alkA* promoter. *In vivo* transcription from the *ada* promoter region was measured using both MV1601 strain carrying a *lacZ* transcriptional insertion within the chromosomal *alkB* gene and MV6608 strain harboring the *aidB* mutation in the a*lkB*::*lacZ* background. To measure *in vivo* transcription from the *alkA* promoter we used MV1571 strain which has a chromosomal *lacZ* operon fused to the *alkA* promoter and its *aidB* mutant derivative MV6607. We did not observe any effect of the *aidB* deletion on *lacZ* expression driven by the *ada* and *alkA* promoters: during normal cell growth, the β-galactosidase catalytic activities determined in the Δ*aidB* strains were low and similar to that resulting from the wild type strains (data not shown). Thus, we could not assign any role to AidB protein in transcription regulation of *ada, alkB* and *alkA* genes, although gel retardation assays have shown specific interaction between AidB and promoters under study.

In vitro regulatory activity of AidB.

The transcriptional activity of the AidB protein on its promoter was further investigated through *in vitro* transcription assays using *E. coli* RNA polymerase holoenzyme. Although RNA polymerase σ^{70} ($E\sigma^{70}$) is able to bind to the *aidB* promoter region in the absence of additional factors, it can initiate transcription efficiently only in the presence of the activator protein Ada (Landini, P. and Volkert, M.R. 1995). Studies performed by Lacour *et al.,* (2002) have shown that $E\sigma^{70}$, in the absence of others factors, is able to activate transcription from a mutant form of the *aidB* promoter in which the residue C at position -12 has been substituted with the nucleotide T. For this reason, *in vitro* transcription assays were carried out by using two different DNA templates: plasmid pSL101 containing wild type P*aidB* and vector pSL112 carrying mutant P*aidB* with a C-T substitution at position -12. The results of these experiments are shown in Fig. 3. Two distinct gel bands were detected when the RNA polymerase σ^{70} was incubated with the plasmid pSL112 containing mutant P*aidB* (Fig. 3, Lane 2). The upper band is relative to the *aidB* transcript while the lower band corresponds to the transcript of the RNA I promoter used as an internal control. The upper band is rather weak when the RNA polymerase σ^{70} was incubated with the plasmid pSL101 (Fig. 3, Lane 1) because, as described above, $E\sigma^{70}$ requires the activator Ada protein to initiate transcription efficiently from the wild type *aidB* promoter. When increasing quantities (0.1-0.4 μg) of the AidB protein were added to the transcription reaction containing vector pSL112, the band corresponding to the *aidB* transcript clearly disappeared (Fig. 3, Lanes 3-6) while the transcript of the RNA I promoter was not affected by the presence of AidB. In agreement with the results of the *in vivo* transcription experiments, AidB was shown to be involved in negative regulation of its own gene.

Characterization of the functional domains in AidB.

Knowledge of the domain architecture of the AidB flavoprotein is necessary

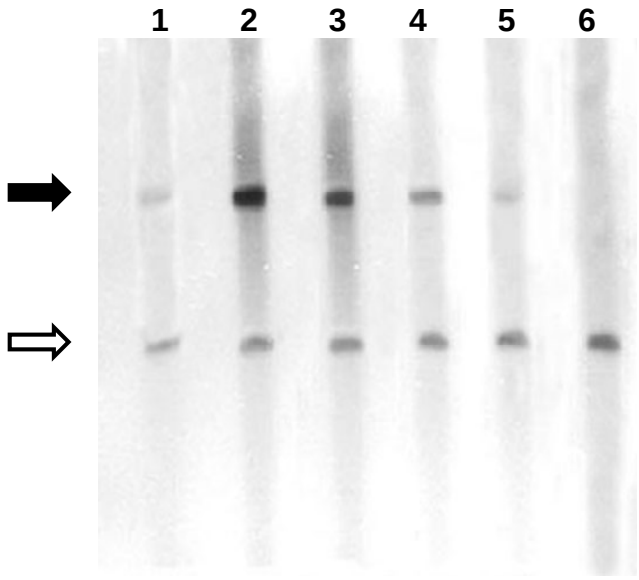


Fig. 3. *In vitro* transcription experiments with RNA polymerase at the *aidB* promoter.
RNA polymerase σ^{70} was incubated with the plasmid pSL101 containing wild type P*aidB* or with the plasmid pSL112 containing mutant P*aidB*, in the absence and in the presence of AidB. Lane 1: pSL101. Lane 2: pSL112. Lanes 3-6: pSL112 incubated with 0.1, 0.2, 0.3 and 0.4 µg of AidB. The filled arrow head shows the position of the transcript from the *aidB* promoter. The open arrow head indicates the position for the transcript of the RNA I promoter used as an internal control.

for understanding its potential multifunctional properties. Sequence analysis showed that the N-terminal 439 residues of AidB are homologous to the acyl-CoA dehydrogenases (ACADs) (Ghisla, S. and Thorpe, C. 2004) and that the C-terminal region involving residues 440-541 exhibits similarity to a DNA binding domain of human topoisomerase I. This region of AidB has been defined by the x-ray crystal structure as its putative DNA binding site. On the basis of these observations, we focused our study on the identification of the catalytic and DNA-binding domains that are required for AidB function *in vivo*. To this aim, we synthesized two deletion derivatives of the AidB protein: AidBΔCt consisting of the first 439 residues and Ct-AidB corresponding to residues 440-541. The PCR products corresponding to two deletion mutants were cloned into a commercial expression vector of the pET series and the recombinant products were expressed as chimeric proteins bearing a 6 His tag at the C-terminus. The expressed proteins were then purified by affinity chromatography on Ni^{2+}-agarose beads and their homogeneity was tested by SDS-PAGE and mass fingerprinting analyses. The two mutant proteins were then characterized by size-exclusion chromatography. On the basis of its retention time and comparison to standard proteins, AidBΔCt was estimated to possess a molecular mass of 196 KDa: given that the predicted molecular mass is around 50 kDa (including the C-terminal hexahistidine tag), AidBΔCt was determined to be a tetramer. In contrast, Ct-AidB was shown to purify as a monomer.

Successively, the functional characterization of these deletion mutants was performed examining their DNA binding and enzymatic properties. To identify the catalytic domain, the isovaleryl-CoA dehydrogenase (IVD) activity of AidBΔCt and Ct-AidB was examined

Table 2: Isovaleryl-CoA dehydrogenase activity of the full length and mutant AidB proteins was assayed as described in Materials and Methods following the isovaleryl-CoA-dependent reduction of the electron acceptor DCPIP. The data shown are averages of 10 measurements. Nd=not detected.

AidB protein	Isovaleryl-CoA dehydrogenase activity
	μmol min^{-1} [mg protein]$^{-1}$
full lenght	0.12 ± 0
AidBΔCt	0.12 ± 0
Ct-AidB	Nd

and compared with that of the full length AidB. AidBΔCt mutant displayed low levels of isovaleryl-CoA catalytic activity that are identical to that exhibited by the full-length AidB protein (Table 2) whereas Ct-AidB mutant had no detectable enzymatic activity.

Therefore, as predicted by the structural analysis, we have shown that the N-terminal region of AidB is responsible for catalytic activity.

To gain insight into the domain that interacts with DNA, AidBΔCt and Ct-AidB were assayed *in vitro* for DNA binding activity by gel retardation assays. These experiments were performed using the biotin-labelled DNA fragment corresponding to the UP region and the -35 box of the *aidB* promoter ("UP35 P*aidB*", Table 1). As shown in Fig. 4, Ct-AidB exhibited DNA binding activity (Lane 4) while AidBΔCt was not able to bind to DNA (Lane 2). These experiments clearly demonstrated that the C-terminal domain alone is sufficient for DNA binding activity.

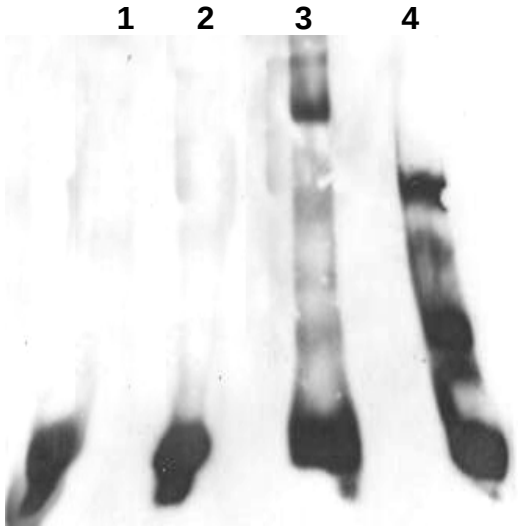


Fig. 4. Gel retardation experiments carried out by incubating the full length and mutant AidB proteins with the "UP35 P*aidB*" probe. Lane 1: UP35 P*aidB* probe. Lanes 2-4: UP35 P*aidB* fragment incubated with AidBΔCt, AidB and Ct-AidB proteins, respectively.

Finally, the ability of the two mutant proteins to repress the transcription of the *aidB* gene was tested by *in vivo* β-galactosidase reporter assays. The Δ*aidB* strain was transformed with reporter plasmid pMV132H together with either pET28a-P*lac*-*aidB*ΔCt or pET28a-P*lac*-Ct-*aidB*.

E. coli cells were grown in LB medium and the β-galactosidase activity was monitored and compared with that detected in the Δ*aidB* strain as well as in the Δ*aidB* strain carrying *aidB* gene on the plasmid pET28a. As shown in Table 3, Ct-AidB repressed *lacZ* expression with levels of β-galactosidase activity decreased to 7% compared with control cells that lacked AidB protein. In contrast, AidBΔCt was not able to repress the P*aidB*-*lacZ* transcription. These reporter assays clearly demonstrated that Ct-AidB acts as a transcriptional repressor of the *aidB* gene *in vivo*.

Table 3: β-galactosidase activity from *E. coli* strain MV5924 harbouring the *lacZ* reporter construct pMV132H and various *aidB* gene constructs. Means and standard deviations have been calculated from four independent assays.

AidB protein	β-galactosidase activity	
	Miller units	**%**
None	3150 ± 46	100
AidB	222 ± 32	7
AidBΔCt	3162 ± 48	100
Ct-AidB	225 ± 31	7

DISCUSSION

Understanding protein functions as well as unravelling cellular mechanisms at the molecular level constitutes a major challenge in modern biology. In this paper, we report a new biological role for AidB, a component of the *E. coli* adaptive response to alkylating agents. While many questions about its function in the Ada-regulated response remain, we assigned a role to AidB in the cell not exposed to alkylating agents. In this study, we provided evidence that AidB shows a high affinity for the upstream regions of P*ada*/*alkB*, P*alkA* and P*aidB*. This observation lead us to a line of inquiry concerning its possible role in the transcription mechanism. This hypothesis was investigated *in vivo* by performing β-galactosidase reporter assays in both wild type and Δ*aidB E. coli* strains. In the absence of alkylation damage, a large increase in the *lacZ* expression driven by *aidB* promoter was observed in the *E. coli* cells lacking AidB as compared to the wild type control. In addition, the complementation of the Δ*aidB* mutation restored *lacZ* transcription to wild type levels. Taken together, these results clearly demonstrated that AidB represses its own synthesis during normal cell growth. *In vitro* transcription experiments performed by using a DNA template containing the *aidB* promoter and in the presence and absence of AidB protein confirmed these data. Autoregulation of the *aidB* gene might play an important physiological role, because to some extent mitigates the energetic burden of bacterial cell not exposed to alkylating agents. In contrast, we were not able to assign any role to AidB in transcription regulation of *ada, alkB* and *alkA* genes.

Another goal of this study was to identify and characterize the functional domains of AidB. The N-terminal region consisting of the first 439 residues was shown to exhibit the same levels of isovaleryl-CoA dehydrogenase (IVD) activity as the full length protein. The level of IVD activity observed in AidB is quite low (Table 2) compared to other acyl-CoA dehydrogenases: for comparison, human isovaleryl-CoA dehydrogenase exhibits specific activity of 8.2 to 11.7 μmol min^{-1} (mg protein)$^{-1}$ (Battaile, K.P. *et al*., 1998; Mohsen, A. W. *et al*., 1998). Recent structural studies revealed several unique features that distinguish AidB's FAD cavity from the ACAD active sites despite the conservation of their general properties (Bowles, T. *et al*., 2008). These observations provided a rationale for AidB's limited acyl-CoA dehydrogenase activity, suggesting that fatty acyl-CoA are not substrates for the enzyme (Rohankhedkar, M.S. *et al*., 2006). The crystal structural of AidB has also identified a putative DNA binding site located in its C-terminal

region. In this work, we verified that the C-terminal residues 440-541 comprise a domain that is responsible for the DNA binding activity. Interestingly, this short region was also shown to function as a transcriptional repressor *in vivo*. Thus, the results reported here showed that the AidB protein is a modular transcription factor which requires its C-terminal region for regulatory function. The identification of two structurally and functionally independent domains suggests that AidB protein might possess multifunctional properties.

Our finding that AidB has transcriptional activity does not provide any precise information on the possible role of the protein in the defence against alkylation. AidB has been hypothesized to detoxify certain alkylating reagents or to repair alkylated DNA (Landini, P. *et al*., 1994; Rohankhedkar, M.S. *et al*., 2006). However, numerous and similar examples of proteins which have more than a biological function have already been reported in the literature, i.e. Meddows *et al*., 2005 demonstrated that the transcription factor DksA is also involved in repair of DNA double-strand breaks.

Another intriguing possibility is that AidB may stimulate the transcription of genes whose products are directly responsible for alkylation resistance. Therefore, it would be useful to perform microarray experiments to identify potential novel *E. coli* genes that are regulated by AidB protein in response to alkylation damage.

ACKNOWLEDGEMENTS

The authors wish to thank Prof. Paolo Landini from the University of Milan, Italy, for pMV132H, pSL101, and pSL112 plasmids.

This work was supported by grants of Ministero dell'Università e della Ricerca Scientifica (Progetti di Rilevante Interesse Nazionale 2004, 2005, 2006; FIRB 2001), of Programma Nazionale di Ricerche in Antartide 2004, and of Regione Campania L.R. 05/03. Support from the National Center of Excellence in Molecular Medicine (MIUR - Rome) and from the Regional Center of Competence (CRdC ATIBB, Regione Campania – Naples), Istituto Nazionale Biomolecole e Biosistemi (INBB) is gratefully acknowledged.

REFERENCES

- Amoresano, A., Cirulli, C., Rippa, V., Esposito, C., Papa, R., and Duilio, A. Submitted to *Molecular Microbiology*.
- Battaile, K. P., McBurney, M., Van Veldhoven, P. P., and Vockley, J. *Biochimica et biophysica acta* **1390**(3), 333–338 (1998).
- Bowles, T., Metz, A. H., O'Quin, J., Wawrzak, Z., and Eichman, B. F. *Proceedings of the National Academy of Sciences of the United States of America* **105**(40), 15299–15304 (2008).
- Bradford, M. M. *Analytical biochemistry* **72**, 248–254 (1976).
- Engel, P. C. (1981). Butyryl-CoA dehydrogenase from *Megasphaera elsdenii. Methods Enzymol* **71**, 359-366.
- Ghisla, S. and Thorpe, C. *European journal of biochemistry/FEBS* **271**(3), 494-508 (2004).
- Karran, P., Hjelmgren, T., and Lindahl, T. *Nature* **296**(5859), 770–773 (1982).
- Kataoka, H., Yamamoto, Y., and Sekiguchi, M. *J. Bacteriol.* **153**(3), 1301–1307 (1983).
- Lacour, S., Kolb, A., Boris Zehnder, A. J., and Landini, P. *Biochemical and biophysical research communications* **292**(4), 922–930 (2002).
- Landini, P., Hajec, L. I., and Volkert, M. R. *Journal of*

bacteriology **176**(21), 6583–6589 (1994).
- Landini, P. and Volkert, M. R. *The Journal of biological chemistry* **270**(14), 8285–8289 (1995).
- Landini, P., Bown, J. A., Volkert, M. R., and Busby, S. J. *The Journal of biological chemistry* **273**(21), 13307–13312 (1998).
- Meddows, T. R., Savory, A. P., Grove, J. I., Moore, T., and Lloyd, R. G. *Molecular microbiology* **57**(1), 97–110 (2005).
- Miller, J. H. *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Pr (1972).
- Miroux, B., and Walker, J. E. *Journal of Molecular Biology* **260**(3), 289-298 (1996).
- Mohsen, A. W., Anderson, B. D., Volchenboum, S. L., Battaile, K. P., Tiffany, K., Roberts, D., Kim, J. J., and Vockley, J. *Biochemistry* **37**(28), 10325–10335 (1998).
- Nakabeppu, Y., Kondo, H., and Sekiguchi, M. *Journal of Biological Chemistry* **259**(22), 13723–13729 (1984).
- Rohankhedkar, M. S., Mulrooney, S. B., Wedemeyer, W. J., and Hausinger, R. P. *Journal of bacteriology* **188**(1), 223–230 (2006).
- Teo, I., Sedgwick, B., Kilpatrick, M. W., McCarthy, T. V., and Lindahl, T. *Cell* **45**(2), 315–324 (1986).
- Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. *Nature* **419**(6903), 174–178 (2002).
- Volkert, M. R. and Nguyen, D. C. *Proceedings of the National Academy of Sciences of the United States of America* **81**(13), 4110–4114 (1984).

Preferential protection from DNA alkylation by the *Escherichia coli* AidB protein

Valentina Rippa[1, 2], Michael Volkert[3] and Angela Duilio[1, 2#].

[1]Department of Organic Chemistry and Biochemistry, University Federico II of Naples, Naples, Italy
[2]School of Biotechnological Sciences, University Federico II of Naples, Naples, Italy
[3]Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts, USA

ABSTRACT
Exposure of *Escherichia coli* to alkylating agents activates expression of *aidB* along with three genes (*ada*, *alkA*, and *alkB*) that encode DNA repair proteins. Despite extensive efforts, the molecular function of AidB in the response to alkylation damage remains uncharacterized. AidB was recently shown to be a flavoprotein that binds to dsDNA, implicating it as a flavin-dependent DNA repair enzyme. Subsequent structural studies suggested that AidB might protect DNA from attack by alkylating agents. In this study, we demonstrate that AidB prevents DNA damage and that it does so in a preferential manner, protecting genes adjacent to promoters carrying an upstream element, an AT rich transcriptional enhancer region. These include many of the housekeeping genes required for basic metabolic processes, such as the *rrn* operons, suggesting that *E. coli* and possibly other organisms preferentially protect subsets of genes from DNA damage.

Keywords: AidB protein, alkylating agents, DNA protection, UP element.

Running title: Protective role of AidB in *Escherichia coli*.

INTRODUCTION
Alkylating agents present in the cell and in the environment chemically modify DNA to produce cytotoxic and mutagenic lesions. Alkylation damage to DNA therefore poses a severe threat to the stability of the genome and, in mammals, can lead to genetic diseases and cancer. As a safeguard against DNA alkylation damage, all organisms have evolved multiple DNA repair mechanisms to remove these modifications and restore DNA to an undamaged state. In addition, bacteria employ an inducible response that serves to protect cells from changing levels of mutagens. In *Escherichia coli*, exposure to sublethal doses of alkylating agents such as methyl methanesulfonate (MMS), *N*-methylnitrosourea (MNU), *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) and *N*-ethyl-*N'*-nitro-*N*-nitrosoguanidine (ENNG) stimulates the expression of four genes, *ada*, *alkB*, *alkA*, and *aidB*. The activation of these genes confers increased cellular resistance to the mutagenic and cytotoxic effects of alkylating agents and is known as the adaptive response (Kataoka, H. *et al*.,1983; Nakabeppu, Y. *et al*., 1984; Volkert, M.R. and Nguyen, D.C. 1984). The Ada protein is the key enzyme of this process; Ada acts both as a methyltransferase able to remove

Corresponding Author:
Angela Duilio
Dipartimento di Chimica Organica e Biochimica,
Complesso Universitario di Monte Sant'Angelo
Via Cinthia 4
80126 Napoli (Italy).
Tel +39-081674314
Fax +39081674313;
email: anduilio@unina.it

methyl groups from damaged DNA and as a transcriptional activator for the adaptive response genes (Teo, I. *et al*., 1986; Nakabeppu, Y. *et al*., 1984; Landini, P. and Volkert, M.R. 1995). AlkA is a double-stranded DNA glycosylase that catalyzes the base excision repair of a wide variety of alkylated bases (Nakabeppu, Y. *et al*., 1984); AlkB is an α-ketoglutarate-Fe(II)-dependent DNA dioxygenase that repairs 1-methyladenine and 3-methylcytosine lesions by oxidative demethylation (Trewick, S.C. *et al*., 2002). Despite detailed understanding of Ada, AlkA, and AlkB, the mechanism by which AidB protects against DNA damage in the adaptive response remains elusive. AidB is related in sequence to acyl-CoA dehydrogenases (ACADs), which use a flavin adenine dinucleotide (FAD) to catalyze the α,β,-dehydrogenation of acyl-CoA conjugates (Landini, P. *et al*., 1994) and it was predicted to function in detoxification of alkylating agents. AidB was shown to be a flavoprotein and to bind to dsDNA implicating it as a flavin-dependent DNA repair enzyme (Rohankhedkar, M.S., *et al*., 2006). Recently, the crystal structure of AidB revealed that its DNA binding site lies within an interior channel, while its flavin binding site is accessible only from the exterior of the protein and is spatially distant from the DNA binding region (Bowles, T. *et al*., 2008). Based on these observations, it was suggested that AidB might bind to DNA and protect it by inactivating alkylators before they are able to react with their DNA target.

The goal of this study was to determine if AidB does in fact protect *E. coli* cells from DNA damage by alkylating agents. Here we demonstrate that AidB prevents DNA damage and that it does so in a preferential manner, protecting genes transcribed from promoters carrying an upstream element, an AT rich transcriptional enhancer region. These include many of the housekeeping genes required for basic metabolic processes, such as the *rrn* operons, suggesting that *E. coli* and possibly other organisms preferentially protect subsets of genes from DNA damage.

MATERIALS E METHODS

Bacterial strains and plasmids.

The bacterial strains, the plasmids and the oligonucleotides used in this work are listed in Table 1. MG1655 and MV5924 *E. coli* strains were a kind gift from M. Volkert (University of Massachusetts, Worcester, MA).

Media and chemicals.

Luria-Bertani or nutrient broths (for bacterial cultures and plating) and suspension medium (for bacterial dilutions) were used as described by Miller (Miller, J.H. 1972). Ampicillin (Sigma) was used at 100 μg/ml.

Construction of the fusion plasmids for transcription assays.

The *rrnB* promoter P1 with and without its UP element, the *leuA* and the *ompF* promoter were amplified from DNA genomic of *E. coli* by PCR using the primers listed in Table 1. The amplification products were digested with *Sph*I and *Hind*III and cloned into the reporter pET22b-*lacZ* plasmid. To obtain pET22b-*lacZ*, the *lacZ* gene was inserted into the pET22b (+) expression plasmid (Novagen) linearized with *Hind*III and *Xho*I. The resulting plasmids, designated as listed in Table 1, were verified by automated DNA sequencing.

In vivo transcription assays.

MG1655 and MV5924 *E. coli* strains were individually transformed with pET22b-*lacZ*, pET22b-P*rrnB*(+UP)-*lacZ*, pET22b-P*rrnB*(-UP)-*lacZ*, pET22b-P*leuA*-*lacZ* and pET22b-

Table 1: Bacterial strains, plasmids and oligonucleotides.

Strains/plasmids	Description	Reference or source
***E. coli* strains**		
C41 (DE3)	Strain that derives from BL21 [F⁻*ompT hsdS*$_B$ (r_B-m_B-) *gal dcm* (DE3)]. This strain has at least one uncharacterized mutation that prevents cell death associated with expression of many toxic recombinant proteins	Miroux *et al.* (1996)
MG1655	F⁻ wild-type K-12 strain	Volkert
MV5924	Strain derivative of MG1655 in which the *aidB* gene had been inactivated by a tetracycline resistance cassette	Volkert
Plasmids		
pET22b(+)	carries an N-terminal *pelB* signal sequence for potential periplasmic localization, plus an optional C-terminal His-tag sequence	Novagen
pET22b-*lacZ*	pET22bΔ(*Hind*III-*Xho*I)Ω(*lacZ* gene)	This work
pET22b-P*rrnB*(+UP)-*lacZ*	pET22b-*lacZ*Δ(*Sph*I-*Hind*III)ΩP*rrnB*(+UP)	This work
pET22b-P*rrnB*(-UP)-*lacZ*	pET22b-*lacZ*Δ(*Sph*I-*Hind*III)ΩP*rrnB*(-UP)	This work
pET22b-P*leuA*-*lacZ*	pET22b-*lacZ*Δ(*Sph*I-*Hind*III)ΩP*leuA*	This work
pET22b-P*ompF*-*lacZ*	pET22b-*lacZ*Δ(*Sph*I-*Hind*III)ΩP*ompF*	This work
Oligonucleotides		
lacZ Fw	5'-TGTAAGCTTATAACAATTTCACACAGGAA-3'	
lacZ Rv	5'-CGGCTCGAGTTATTTTTGACACCAGAC-3'	
P*rrnB*(+UP) Fw	5'-TAAAGCATGCATGTTGCGCGGTCAG-3'	
P*rrnB*(-UP) Fw	5'-ATTTGCATGCCCTCTTGTCAGGCC-3'	
P*rrnB* Rv	5'-ATTAAGCTTAGGAGAACCCCGCTGA-3'	
P*leuA* Fw	5'-ATAAGCATGCGGGTTGACATCCGTT-3'	
P*leuA* Rv	5'-AAGAAGCTTGATAAAGCGAACGATGTG-3'	
P*ompF* Fw	5'-ATTTGCATGCACAAAGTTCCTTAAATTTTA-3'	
P*ompF* Rv	5'-TAAAAGCTTAATAAAAATTTACGGAACTATTG-3'	

P*ompF-lacZ* plasmids. These bacterial cultures grown overnight in LB medium at 30°C, were diluted 1:100 in fresh medium. At an $A_{600\ nm}$ of 0.4, the cultures were divided in five aliquots: one was not supplemented and the other four aliquots were supplemented with MNNG (5µg/ml), ENNG (5µg/ml), MMS 0.04%, and MNU 0.04%, respectively. Cellular pellets were collected during the exponential growth phase. β-galactosidase activity from the promoters-*lacZ* fusions was determined by measuring ONPG-hydrolysis, as described by Miller (Miller, J.H. 1972) and was comparable with the activity obtained by a promoterless *lacZ* gene.

Production and purification of AidB *protein.*

The construct pET28a-*aidB* (Amoresano *et al*., submitted for publication) was transformed into the *E. coli* strain C41 (DE3). The recombinant cells were grown at 25°C to an optical density at 600 nm of ~ 0.5, at which time 0.05 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) was added and the culture was

allowed to grow until the OD_{600} reached 3.0. Cells were harvested by centrifugation at 5000 x g for 15 min at 4°C, resuspended in buffer A (50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl, 10 mM imidazole, 1 mM PMSF), disrupted by passage through a French Press and centrifuged at 14000 x g for 30 min at 4°C. The recombinant protein was purified by affinity chromatography on His-Select Nickel Affinity Gel (Sigma). The lysate was loaded onto His-Select Nickel Affinity Gel equilibrated with equilibration buffer (50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl, 10 mM imidazole). After 1 min of incubation at 4°C, the matrix was collected by centrifugation at 11000 x g for 1 min and washed 3 times with same equilibration buffer. The recombinant protein was eluted with buffer containing 250 mM imidazole in 50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl. Protein concentration was estimated with Bradford reagent (Bio-Rad protein assay) and protein content was checked by SDS-polyacrilamide gel electrophoresis (SDS-PAGE).

Isolation of plasmid DNA and damage assay.

The MG1655 and MV5924 *E. coli* strains bearing pET22b-*lacZ* were grown overnight in LB medium at 30°C; these bacterial cultures were then diluted 1:100 in fresh medium. At an $A_{600\ nm}$ of 0.4, the cultures were divided in five aliquots: one was not supplemented and the other four aliquots were supplemented with MNNG (5µg/ml), ENNG (5µg/ml), MMS 0.04%, MNU 0.04%, respectively. After addition of alkylating agent, the bacterial cells were allowed to grow for 3 h; the plasmid DNA was isolated and served as a probe for the estimation of alkylated bases. The plasmids were divided into 2 aliquots, one of which was treated with the *E. coli* AlkA (a kind gift from Patrick J. O'Brien) and endonuclease IV (NEB); the other aliquot did not receive further treatment (control). Treatment with AlkA was performed in 70 mM MOPS, pH 7.5, 1 mM EDTA, 1mM DTT, 5% glycerol for 30 min at 37°C, followed by treatment with Endo IV for 1 h at 37°C. Then the samples were subjected to electrophoresis in 0.8% agarose gel for ~1 h at 80 V using 40 mM Tris, pH 7.8, 1 mM EDTA buffer.

In vitro DNA binding and damage assays.

An *in vitro* DNA binding assay was conducted incubating pET22b-*lacZ* plasmid (100 ng) with varying concentrations of AidB in buffer B (10 mM Tris, pH 7.8, 1 mM EDTA, 20% glycerol), for 1 h at 30°C. Mixtures were then analysed by electrophoresis in a 0.8% agarose gel using 40 mM Tris, pH 7.8, 1 mM EDTA buffer. AidB and DNA were mixed at the following ratios (w/w): 3:1, 15:1, 30:1. In DNA damage assay, AidB (3 µg) was incubated with pET22b-*lacZ* (100 ng) for 1 h at 30°C, either before or after the addition of 300 mM MMS to the reaction mixture. The samples were then treated with AlkA and Endo IV and subjected to electrophoresis in 0.8% agarose gel for ~1 h at 80 V using 40 mM Tris, pH 7.8, 1 mM EDTA buffer.

Determination of DNA damage on lacZ gene.

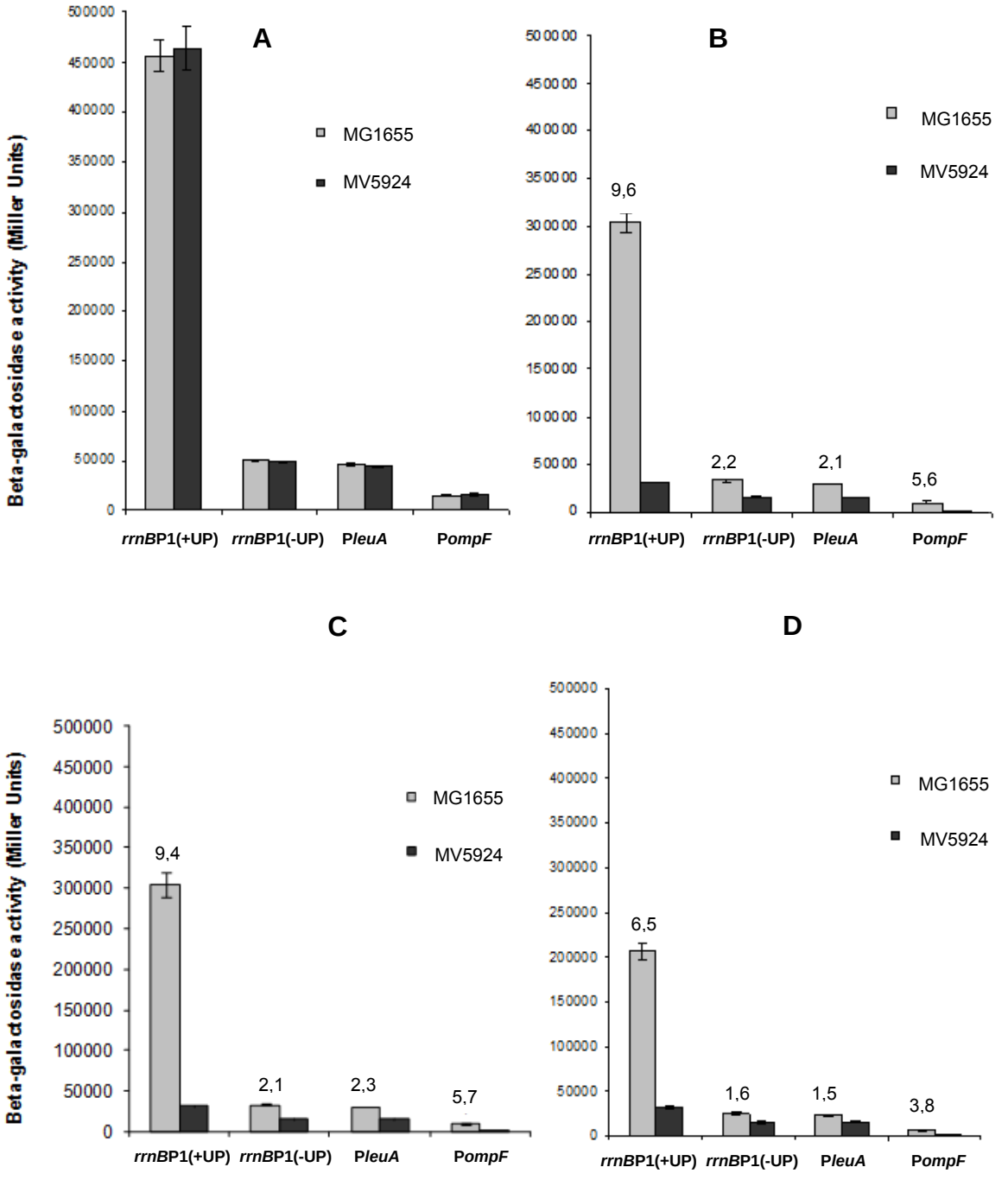
MG1655 and MV5924 *E. coli* strains were individually transformed with pET22b-*lacZ*, pET22b-P*rrnB*(+UP)-*lacZ*, pET22b-P*rrnB*(-UP)-*lacZ.* These bacterial cultures grown overnight in LB medium at 30°C, were diluted 1:100 in fresh medium. At an $A_{600\ nm}$ of 0.4, the cultures were supplemented with 0.04% MMS to activate the adaptive response and the bacterial cells were allowed to grow for 3h. Then, the plasmids under study were isolated from these bacterial cells and were

digested with *Hind*III and *Xho*I to release the *lacZ* fragment. To estimate the presence of alkyl lesions, the DNA fragments were treated and not with the AlkA and Endo IV proteins. The samples were then subjected to electrophoresis on alkaline agarose gel in 30 mM NaOH, 1 mM EDTA, pH 8 buffer, at 60 V for 3 hours at 25°C. The gel was neutralized by soaking in a solution containing 1.5 M NaCl and 1 M Tris-HCl, pH 7.6 for 1 h. Finally, the gel was stained in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.4) containing SYBR® Gold for 30 min at 25°C and the samples were then analysed for single-strand DNA breaks.

RESULTS

Functional analysis of AidB during transcription.

Recently, using the UP element and the -35 sequence from the *rrnB* P1 promoter as bait and a sequence



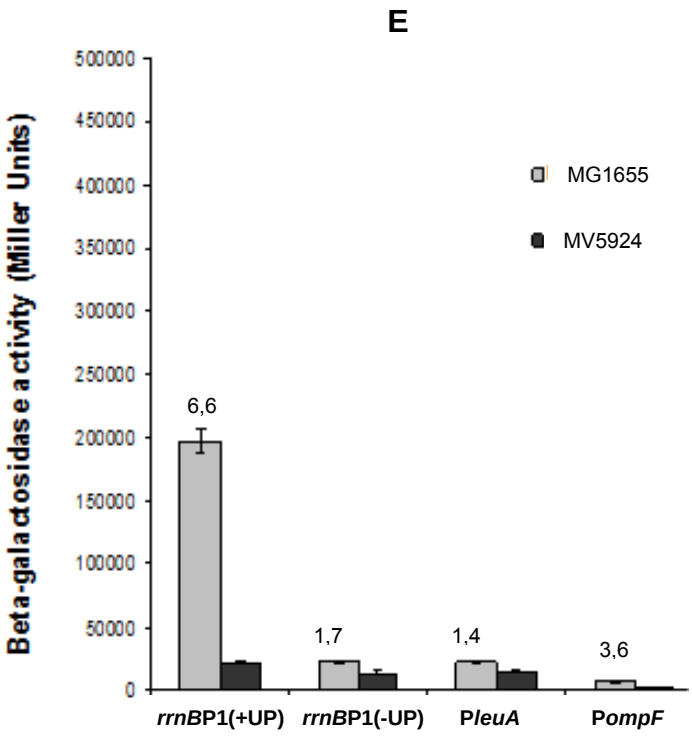


Figure 1. *In vivo* transcription from the promoters-*lacZ* fusions. The pET22b-P*rrnB*(+UP)-*lacZ*, pET22b-P*rrnB*(-UP)-*lacZ*, pET22b-P*leu*A-*lacZ* and pET22b-P*omp*F-*lacZ* plasmids were individually introduced into MG1655 (wild type) and MV5924 (Δ*aidB*) *E. coli* strains and the β-galactosidase specific activity was determined in the absence (A) and in the presence of MNNG (5μg/ml) (B), ENNG (5μg/ml) (C), MMS 0.04% (D), MNU 0.04% (E) as alkylating agents. The activities of promoters are reported in Miller units; the activity obtained by a promoterless *lacZ* gene was subtracted. Numbers above bars refer to the ratio of the β-galactosidase activity of the promoter measured in the wild type cells to the activity of that same promoter in the *aidB* mutant strain. Means and standard deviations have been calculated from four independent assays.

lacking the UP element for comparison, we isolated proteins on the basis of their preferential affinity for the UP region. Curiously, AidB was found as one of several proteins that preferentially bind to the UP element containing DNA (Amoresano *et al*., submitted for publication). In order to determine whether the presence of AidB at the *rrnB* P1 promoter might be of biological significance, we tested its effect on transcription from this promoter by *in vivo* transcription assays. As a control, we also investigated the putative AidB activity on other several promoter fragments: the ribosomal promoter *rrnB* P1 without its UP sequence, the promoter from an unrelated gene which lacks an UP element (P*leuA*), and the promoter from an unrelated gene which contains an UP region (*ompF*). The DNA fragments under study were individually fused to a promoterless *lacZ* gene contained in the reporter plasmid pET22b-*lacZ*. Both MG1655 (wild type) and MV5924 (Δ*aidB*) *E. coli* strains were then transformed with the fusion plasmids and grown in LB medium, either in the absence or in the presence of alkylating agents (MMS, MNU, MNNG, ENNG) and the β-galactosidase activity was monitored during the exponential growth phase. As shown in Fig. 1A, in all cases, wild type and mutant cells not exposed to alkylators showed identical levels of β-galactosidase activity indicating that the presence of AidB has no effect on transcription during normal cell growth. In contrast, upon treatment with alkylating agents, the transcription levels in the Δ*aidB* strain were lower than that observed for the wild type strain (Fig. 1B-E), suggesting a functional role for AidB on transcription during alkylation stress. Interestingly, the results from β-galactosidase assays showed that AidB strongly affects the activity of promoters containing an UP element: indeed, the mutant cells harbouring *rrnB* P1 and P*ompF*, upon exposure to N-alkyl-nitrosoguanidines (MNNG, ENNG) (Fig. 1B, 1C), exhibited almost a 10-fold and 6-fold decrease of promoter

activity, respectively. When the promoter regions without UP sequence were analyzed, we find that the mutant cells exhibit transcription levels slightly lower (approximately 2-fold) than that observed for the wild type cells. We also observed that AidB allows more efficient transcription in *E. coli* cells exposed to MNNG and ENNG (Fig. 1B, 1C) rather than to other alkylators such as MMS or MNU (Fig. 1D, 1E), supporting the hypothesis that AidB might be involved in deactivation of nitrosoguanidines.

Taken together, these data strongly demonstrated that AidB affects transcription during alkylation stress and that it has more pronounced effect on the activity of promoters containing an upstream element. These results raise the possibility that AidB may prevent alkyl damage, thereby allowing more efficient transcription.

AidB reduces the level of alkylation damage in DNA.

The aim of our study was to determine whether AidB might be able of preventing DNA alkylation damage. This issue was investigated by a plasmid damage assay: the plasmid DNA (pET22b-*lacZ*) was isolated from wild type and *aidB* mutant cells grown both in the absence and in the presence of alkylators (MMS, MNU, MNNG, ENNG) and served as a probe for the estimation of alkylated bases on DNA. The plasmids were divided into 2 aliquots, one of which was treated with the *E. coli* AlkA (a kind gift from Patrick J. O'Brien) and endonuclease IV (Endo IV) proteins; the other aliquot did not receive further treatment and served as a control. AlkA glycosylase recognizes and removes a wide variety of alkylated bases (O'Brien, P. J. and Ellenberger, T., 2004); Endo IV is an

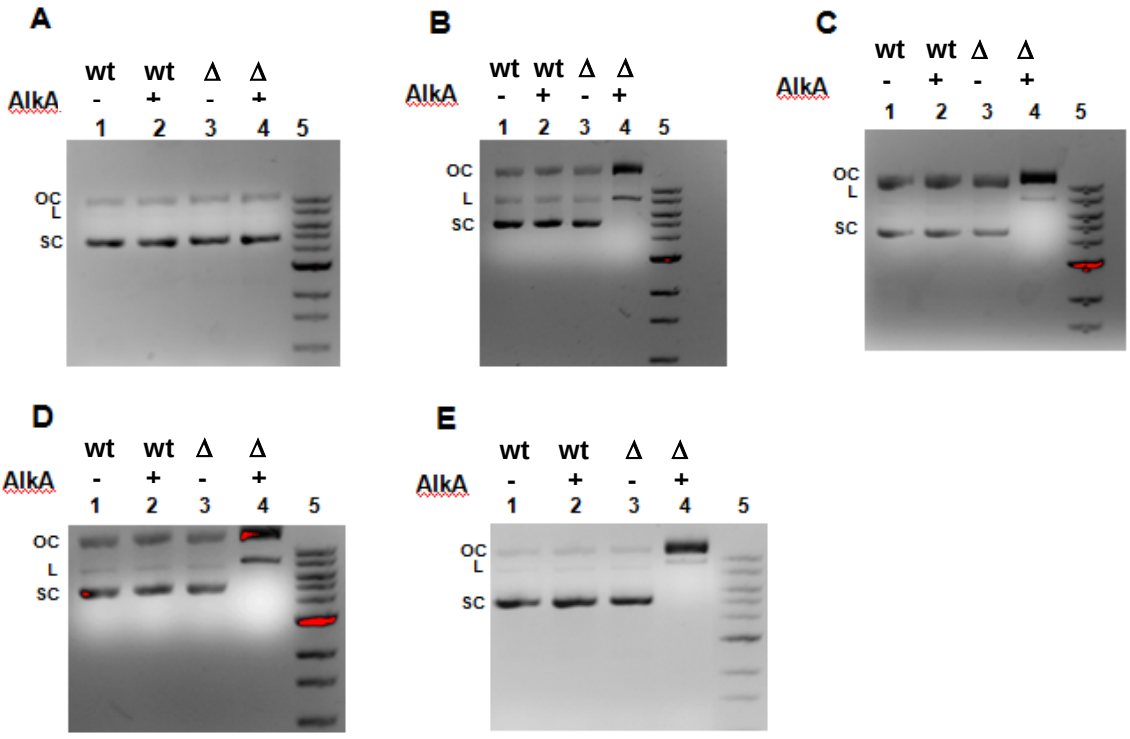


Figure 2. Plamid damage assay. The pET22b-*lacZ* DNA was isolated from wild type (Lanes 1-2) and Δ*aidB* (Lanes 3-4) *E. coli* strains grown in the absence (A) or in the presence of MNNG (5 μg/ml) (B), ENNG (5 μg/ml) (C), MMS 0.04% (D), MNU 0.04% (E), digested (lanes 2, 4) or not (lanes 1, 3) with AlkA and Endo IV and subjected to agarose gel electrophoresis. Lane 5, 1 Kb DNA marker (NEB). OC: open circular; L: linear; SC: supercoiled.

apurinic/apyrimidinic (AP) endonuclease that converts abasic sites produced by AlkA to nicks (Ljungquist, S. 1977); therefore, the combined action of these two enzymes on a damaged plasmid results in the conversion of the covalently closed circular (supercoiled) form to the open circular form. AlkA treated and untreated plasmids were then subjected to electrophoresis on agarose gels and tested to look for loss of the supercoiled form. As shown in Fig. 2A, alkyl lesions were not detected in plasmids isolated from bacteria grown in LB medium; when plasmid DNA from cells exposed to alkylating agents was analyzed (Fig. 2B-E), we find that treatment with AlkA and Endo IV had no effect on DNA isolated from wild type cells (Lanes 2), but resulted in a complete loss of supercoiled plasmid from the *aidB* mutant (Lanes 4). These results strongly indicated that the presence of AidB reduces the level of alkylation damage in plasmid DNA. While the damage assay results are consistent with a role for AidB in predicted protection of DNA, its function in repair of alkyl lesions is not ruled out.

AidB binds and protects plasmid DNA.
In vitro experiments were performed to demonstrate conclusively that AidB binds DNA to protect it from alkylation damage. To this aim, the recombinant AidB protein bearing a His tag was produced, purified by affinity chromatography on Ni^{2+}-agarose beads and used in a plasmid damage assay. First, gel mobility shift assays were performed to determine the amount of AidB required to completely bind the plasmid DNA. The saturation of binding was achieved at a ratio protein:DNA of 30:1 (w/w) (data not shown). Using this ratio, AidB was then incubated with plasmid DNA either before or after the addition of MMS to the reaction mixture. MMS was used for this assay because it does not require metabolic conversion to alkylate DNA (Adler, I. D. 1980). To detect the presence of alkyl lesions, the purified plasmid DNA was cleaved with AlkA and Endo IV. Figure 3 shows that plasmid remains in the supercoiled form when AidB is added prior to MMS treatment (Lane 4), indicating that AidB protein functions *in vitro* and that the His-tagged form is fully active. However, when the sequence is reversed and AidB is added after MMS treatment, no effect of AidB presence is detected (Lane 8). An identical experiment using BSA protein served as a control (data not shown).
Since the order of addition should not affect an AidB repair function, but protection should require the presence

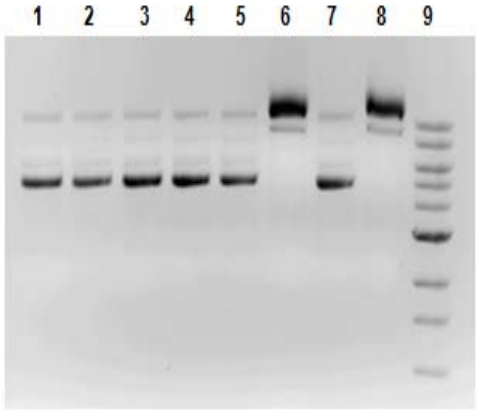


Figure 3. AidB protects DNA from alkylation damage *in vitro*. Lane 1-2: pET22b-*lacZ* plasmid not digested or digested with AlkA and Endo IV, respectively. Lanes 3-4: The pET22b-*lacZ* plasmid was incubated with AidB for 1 h at 37°C; the complex AidB-DNA was treated with MMS and not digested (Lane 3) or digested (Lane 4) with AlkA and Endo IV. Lanes 5-6: The pET22b-*lacZ* was treated with MMS; the DNA methylated was not digested or digested with AlkA and Endo IV, respectively. Lanes 7-8: The pET22b-*lacZ* was treated with MMS; the DNA methylated was incubated with AidB and this sample was not digested or digested with AlkA and Edo IV, respectively. Lane 9, 1 kb DNA marker (NEB).

of AidB prior to damaging treatments, we conclude that AidB protein acts by preventing damage from occurring.

AidB preferentially protects DNA regions containing an UP element.

Since AidB specifically binds to UP sequence containing promoters, and it affects strongly transcription at these promoters during exposure to alkylating agents, we supposed that AidB protection might be preferentially targeted to genes with an UP element. To address this question, we analyzed the effect exhibited by AidB *in vivo* on a portion of plasmid DNA: the *lacZ* fragment. In this experiment, we used plasmids carrying a *lacZ* gene that lacks a promoter, *lacZ* fused to the *rrnB* P1 promoter containing its UP element and *lacZ* fused to *rrnB* P1 in which UP sequence was deleted. We investigated whether the presence of AidB might affect the content of alkyl lesions within the *lacZ* sequences. To this aim, the plasmids under study were isolated from wild type and *aidB* mutant cells treated with MMS and were digested with restriction enzymes to release the *lacZ* fragment. To estimate the presence of alkyl lesions, the *lacZ* DNA fragments were gel purified, extracted and treated or not with the AlkA and Endo IV proteins. The samples were then subjected to electrophoresis on alkaline agarose gels to denature DNA and separate nicked from full-length ssDNA fragments. The *lacZ* integrity was then determined by the relative abundance of full length *lacZ* restriction fragments and nicked forms. Figure 4 shows that the mutant cells are not able to protect the *lacZ* gene (Lanes 8, 10, 12), since no full length *lacZ* fragments were detected. This confirms that the presence of AidB is required for the protection against alkyl damage. Analysis of the DNA isolated from wild type cells revealed several interesting features of AidB specificity. The *lacZ* sequence fused to the UP element containing *rrnB* P1 was fully protected since all DNA exposed to AlkA remains as full length (Fig. 4, comparison between Lanes 3 and 4). When *lacZ* fused to the UP-less ribosomal promoter was tested, we find that the sample treated with AlkA (Lane 6) shows a decrease in the amount of full length fragment, as compared with the untreated control DNA (Lane 5). Interestingly, the treatment of promoterless *lacZ* with AlkA caused an almost complete loss of full length fragment (Lane 2).

These results demonstrate that AidB preferentially protects the DNA of genes transcribed from an UP element containing promoter. However, it also protects DNA in general, since it does provide some protection to the non-

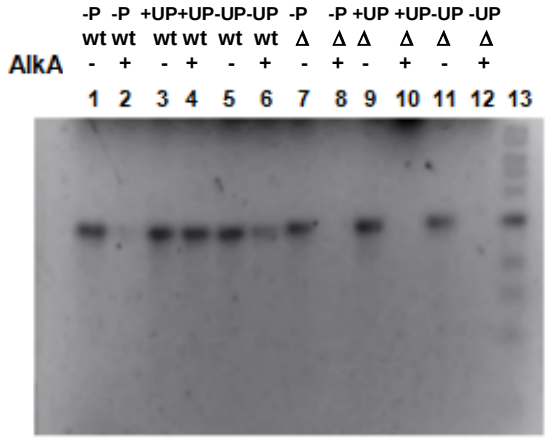


Figure 4. AidB preferentially protects DNA regions containing an UP element. The pET22b-*lacZ*, pET22b-P*rrnB*(+UP)-*lacZ* and pET22b-P*rrnB*(-UP)-*lacZ* plasmids were isolated from wild type (Lanes 1-6) and Δ*aidB* (Lanes 7-12) *E. coli* strains grown in the presence of MMS 0.04% and digested to release the *lacZ* fragment. The digested DNA samples were untreated (Lanes 1,3,5,7,9,11) or treated (Lanes 2,4,6,8,10,12) with AlkA and Endo IV and subjected to electrophoresis on alkaline agarose gels. Lanes 1,2,7,8: *lacZ* lacking a promoter (-P); Lanes 3,4,9,10: *lacZ* fused to the *rrnB* P1 promoter with its UP element (+UP); Lanes 5,6,11,12: *lacZ* fused to the *rrnB* P1 promoter without its UP element (-UP); Lane 13: 1Kb DNA marker (NEB).

transcribing *lacZ* restriction fragment and the *lacZ* gene fused to the *rrnB* P1 promoter lacking the UP element. Moreover, the result that AidB protects different sequences to varying degrees is consistent with the initial observation (Fig. 2) that AidB protects whole plasmids from alkylation. The observation that the entire *lacZ* sequence adjacent to UP element containing promoters is protected in a better manner suggests that the promoter may function to nucleate AidB protein binding which may then polymerise along the DNA molecule. The structural studies of Bowles *et al*., 2008, suggest that a single AidB protein molecule can bind from 25-30 base pairs of dsDNA. Thus to protect the genome, a minimum of 160,000 AidB molecules would be required. Since AidB has a relatively weak promoter, it is unlikely that such a level would ever be achieved. Thus we suspect that the preferential binding to genes such as *rrnB*, and potentially to other UP element containing DNA sequences such as *aidB* itself and its regulatory gene *ada* (Landini, P. *et al*., 1998), *recA*, *polA* and other highly transcribed genes (Estrem, S.T. *et al*., 1999) may preferentially protect these DNA sequences from alkylation damage.

DISCUSSION

In a previous study, we found the AidB protein to be part of the protein complex that binds to the -35 and the UP elements of the *rrnB* P1 promoter, but not to a sequence lacking the UP element. In this work, we investigated whether the presence of AidB at the *rrnB* P1 promoter might be of biological significance; to this aim, we evaluated its possible role in rRNA expression by β-galactosidase reporter assays. In addition, we also tested the effect of AidB on transcription from other promoters fused to the *lacZ* reporter gene. As appropriate controls, the following promoters were chosen: the ribosomal promoter *rrnB* P1 without its UP sequence, the promoter from an unrelated gene which lacks an UP element (P*leuA*) and the promoter from an unrelated gene which contains an UP region (P*ompF*). The fusion plasmids were individually introduced into both wild type and Δ*aidB E. coli* strains which were then treated and not with alkylating agents. Upon alkylation stress, a decrease in the *lacZ* expression was observed in the *E. coli* cells lacking AidB as compared to the wild type control. Interestingly, AidB was found to have a more pronounced effect on the activity of promoters containing an upstream element. Our results clearly demonstrated that a functional *aidB* allele is necessary for efficient transcription during exposure to alkylating agents. These data lead us to a line of inquiry concerning the possible role of AidB in the transcription initiation. Its function might be related to repair of alkylated DNA given that, during transcription, the transcriptional fork generates single stranded DNA molecules that have been demonstrated to undergo alkylation processes much quicker than double stranded molecules. Another possibility is that AidB might prevent alkyl damage by protecting DNA, thereby allowing more efficient transcription. In this respect, a recent report on the three dimensional structure of AidB bound to double strand DNA provided a rationale for its possible role in damage prevention (Bowles, T. *et al*., 2008).

The goal of our study was thus to determine the mechanism by which AidB protects against alkylating agents. First, through a plasmid damage assay we demonstrated that AidB *in vivo* reduces the level of alkylation damage in DNA: indeed,

upon exposure to alkylators, the DNA plasmid isolated from wild type cells was shown to be undamaged while alkyl lesions were detected in plasmid DNA extracted from mutant cells. While these results are consistent with a role for AidB in predicted protection of DNA, we could not rule out its possible function in repair of alkyl lesions. To unravel this issue, *in vitro* studies were performed incubating AidB with plasmid DNA either before or after the addition of MMS; an effect of AidB presence was detected only in the former case. Therefore, as predicted by structural analysis, we have demonstrated conclusively that AidB binds DNA to protect it from alkylation damage rather than repair it. A similar role has been observed for the Dps protein, which protects DNA in starved *E. coli* against oxidative damage (Almiron, M. *et al*., 1992; Martinez, A. and Kolter, R. 1997). Both AidB and Dps are up-regulated during stationary phase and are *rpoS*-dependent (Volkert, M.R. *et al*., 1994; Altuvia, S. *et al*., 1994). Interestingly, endogenous methylating agents such as nitrosamines are formed as by-products of stationary-phase metabolism (Taverna, P. and Sedgwick, B, 1996), leading to an accumulation of alkyl lesions on DNA. Induction of AidB expression, therefore, could serve to prevent endogenous stationary-phase alkylation damage in a manner similar to Dps protection of oxidative damage.

Finally, since AidB was shown to specifically bind to UP sequence containing promoters and to affect strongly transcription at these promoters upon exposure to alkylating agents, we asked if AidB protection might be preferentially targeted to genes with an UP element. To address this question, we investigated whether the presence of AidB might affect the content of alkyl lesions within *lacZ* sequences fused to the *rrnB* P1 promoter containing and not its UP element and judged by comparison with the *lacZ* fragment lacking a promoter. Our experiments demonstrated that AidB preferentially protects DNA sequences transcribed from promoters carrying an UP element.

In conclusion, in this work we characterized the role of AidB in the *E. coli* adaptive response to alkylating agents: AidB functions to protect DNA from alkylation damage, by binding it and, presumably, by inactivating alkylators before they are able to react with their DNA target. In addition, we demonstrated that AidB preferentially protects DNA sequences adjacent to UP element-containing promoters, which include many of the housekeeping genes required for basic metabolic processes, such as the sequences encoding the ribosomal rRNA or the DNA polymerase I that is involved in DNA replication and repair. Further experiments are required to address the question whether the preferential protection exhibited by AidB is limited to transcriptionally active DNA. It would also be useful to examine if a similar protective mechanism is conserved and active in other organisms.

ACKNOWLEDGEMENTS

This work was supported by grants of Ministero dell'Università e della Ricerca Scientifica (Progetti di Rilevante Interesse Nazionale 2005, 2006), of Programma Nazionale di Ricerche in Antartide 2004, and of Regione Campania L.R. 05/03. Support from the National Center of Excellence in Molecular Medicine (MIUR - Rome) and from the Regional Center of Competence (CRdC ATIBB, Regione Campania – Naples), Istituto Nazionale Biomolecole e Biosistemi (INBB) is gratefully acknowledged.

REFERENCES

- Adler, I. D. *Mutation research* **74**(2), 77–93 (1980).
- Almirón M., Link, A. J., Furlong, D., and Kolter, R. *Genes & Development* **6**(12b), 2646–2654 (1992).
- Altuvia, S., Almirón, M., Huisman, G., Kolter, R., and Storz, G. *Molecular microbiology* **13**(2), 265–272 (1994).
- Bouziane, M., Miao, F., Ye, N., Holmquist, G., Chyzak, G., and O'Connor, T. R. *Acta biochimica Polonica* **45**(1), 191–202 (1998).
- Bowles, T., Metz, A. H., O'Quin, J., Wawrzak, Z., and Eichman, B. F. *Proceedings of the National Academy of Sciences of the United States of America* **105**(40), 15299–15304 (2008).
- Kataoka, H., Yamamoto, Y., and Sekiguchi, M. *J. Bacteriol.* **153**(3), 1301–1307 (1983).
- Landini, P., Hajec, L. I., and Volkert, M. R. *Journal of bacteriology* **176**(21), 6583–6589 (1994).
- Landini, P. and Volkert, M. R. *The Journal of biological chemistry* **270**(14), 8285–8289 (1995).
- Landini, P., Bown, J. A., Volkert, M. R., and Busby, S. J. *The Journal of biological chemistry* **273**(21), 13307–13312 (1998).
- Ljungquist, S. *Journal of Biological Chemistry* **252**(9), 2808–2814 (1977).
- Martinez, A. and Kolter, R. *J. Bacteriol.* **179**(16), 5188–5194 (1997).
- Miller, J. H. *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Pr.
- Nakabeppu, Y., Kondo, H., and Sekiguchi, M. *Journal of Biological Chemistry* **259**(22), 13723–13729 (1984).
- O'Brien, P. J. and Ellenberger, T. *Journal of Biological Chemistry* **279**(26), 26876–26884 (2004).
- Rohankhedkar, M. S., Mulrooney, S. B., Wedemeyer, W. J., and Hausinger, R. P. *Journal of bacteriology* **188**(1), 223–230 (2006).
- Taverna, P. and Sedgwick, B. *Journal of bacteriology* **178**(17), 5105–5111 (1996).
- Teo, I., Sedgwick, B., Kilpatrick, M. W., McCarthy, T. V., and Lindahl, T. *Cell* **45**(2), 315–324 (1986).
- Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. *Nature* **419**(6903), 174–178 (2002).
- Volkert, M. R. and Nguyen, D. C. *Proceedings of the National Academy of Sciences of the United States of America* **81**(13), 4110–4114 (1984).
- Volkert, M. R., Hajec, L. I., Matijasevic, Z., Fang, F. C., and Prince, R. *J. Bacteriol.* **176**(24), 7638–7645 (1994).

Potential role for the *Escherichia coli* AidB and the *Pseudomonas putida* PP4780 as detoxification enzymes.

INTRODUCTION

Alkylating agents comprise a broad class of highly reactive chemical compounds that introduce alkyl groups into biologically active molecules and prevent normal functioning. These compounds mostly react with DNA molecule to produce cytotoxic and mutagenic lesions. Alkylation damage to DNA therefore poses a severe threat to the stability of the genome and, in mammals, can lead to genetic diseases and cancer. Living organisms are continuously exposed to alkylating agents that are among the most abundant environmental pollutants influencing health of the persons exposed: these molecules that include nitrosamines and nitrosoureas, are released into the ecosystem at high concentrations, as a consequence of human activities (Vaughan, P., *et al*., 1991; Taverna, P. and Sedgwick, B. 1996). Since alkylating molecules are ubiquitous and hence unavoidable, all organisms (eubacteria, archaebacteria, and eukaryotes) have evolved multiple DNA repair strategies as a safeguard against alkylation damage. In addition, many bacteria, to defend against fluctuating levels of environmental alkylators, mount an inducible response that enhances cellular resistance to alkylation damage and is known as the adaptive response. In *Escherichia coli*, exposure to sublethal doses of alkylating agents such as methyl methanesulfonate (MMS), *N*-methylnitrosourea (MNU), *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG), and *N*-ethyl-*N'*-nitro-*N*-nitrosoguanidine (ENNG) increases the expression of *aidB* along with three genes (*ada*, *alkA*, *alkB*) that encode DNA repair proteins (Kataoka, H. *et al*., 1983; Nakabeppu, Y. *et al*., 1984; Volkert, M.R. and Nguyen, D.C. 1984). The role of AidB in the response to alkylating agents was elusive for many years. AidB was shown to be a flavoprotein that has weak isovaleryl-CoA dehydrogenase activity and that binds to double strand DNA (Landini, P. *et al*., 1994; Rohankhedkar, M.S. *et al*., 2006). Our recent work demonstrated that AidB prevents alkylation damage by protecting DNA and, presumably, by inactivating alkylators before they are able to react with their target. Interestingly, a recent report on the three dimensional structure of AidB bound to dsDNA (Bowles, T. *et al*., 2008) supported this model, revealing that the protein is well equipped to sterically occlude DNA from attack by damaging agents. Importantly, the unique chemical environment of FAD active site provided a rationale for a possible role of AidB in deactivation of nitrosoguanidines or their reactive intermediates. In this respect, AidB's FAD cavity was shown to be rich in thiol and methylsulfide groups and since MNNG, unlike MMS or MNU, is activated by thiols to produce highly reactive methylation agents (e.g., methyldiazonium ion) (Lawley, P. D. 1974), it is intriguing to speculate that flavin binding channel may act as a sink for reactive MNNG derivatives. Coupled with structural analysis, our recent results help to support the hypothesis that AidB may act as a detoxification enzyme: indeed, we demonstrated that AidB allows more efficient transcription during stress alkylation and that it has a more pronounced effect in *E. coli* cells exposed to nitrosoguanidines (MNNG, ENNG) rather than to other alkylators such as MMS or MNU. Taking into account these observations, AidB represents a promising tool for the treatment of sites contaminated by alkylating agents. Since, ideally, the design of successful strategies for the bioremediation would require the knowledge of the microorganisms present in the polluted environments, their metabolic abilities, and how they respond to changes in environmental conditions, the challenge of this experimental work was to identify *E. coli* AidB homologues in bacteria used for bioremediation applications

and to investigate the possible involvement of these gene products in the protection against alkylating agents. Particularly, the our attention has been focused on the search for AidB homologues in pseudomonads that are among extensively studied workhorses of environmental bioremediation owing to their ability to degrade numerous different contaminants (Wackett, L.P. 2003). The huge potential of the pseudomonads does not solely depend on a high proportion of genes responsible for the metabolism, transport and efflux of organic compounds, but also on broad capability of metabolic regulation (Nelson, K.E. *et al* ., 2002): indeed, the control of gene expression is the key determinant of their flexibility and, in this respect, a variety of highly integrated regulatory mechanisms have been identified. Hence, on the basis of the potential degradative capability offered by AidB and the great versatility of *Pseudomonas* species, the aim of this work was to identify new potential candidates for the bio-treatment of wastes and environments contaminated by alkylating compounds.

MATERIALS AND METHODS

Bacterial strains and plasmids.

The bacterial strains, the plasmids and the oligonucleotides used in this work are listed in Table 1. MG1655 and MV5924 *E. coli* strains were a kind gift from M. Volkert (University of Massachusetts, Worcester, MA).

Media and chemicals.

Luria-Bertani or nutrient broths (for bacterial cultures and plating) and suspension medium (for bacterial dilutions) were used as described by Miller (Miller, J.H. 1972). Kanamycin and tetracycline (Sigma) were used at 50 and 5 μg/ml, respectively.

Construction of the expression plasmids.

The PP4780 gene was amplified from DNA genomic of *P. putida* KT2440 by polymerase chain reaction (PCR) using the forward and reverse primers listed in Table 1. The amplified fragment was digested with the restriction enzymes underlined in Table 1 and cloned into the expression vector pET28a (Novagen) which was linearized with the same restriction enzymes. The resulting plasmid, pET28a-PP4780, was verified by automated DNA sequencing. For the complementation experiments, the PP4780 gene was positioned downstream of the *lac* promoter into the expression vector pET28a-P*lac*, generating pET28a-P*lac*-PP4780.

Production and purification of recombinant proteins.

The constructs pET28a-*aidB* (Amoresano *et al*., submitted for publication) and pET28a-PP4780 (Table 1) were individually transformed into the *E. coli* strain C41 (DE3). The recombinant cells were grown at 25°C to an optical density at 600 nm of ~0.5, at which time 0.05 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) was added and the cultures were allowed to grow until the OD_{600} reached 3.0. Cells were harvested by centrifugation at 5000 x g for 15 min at 4°C, resuspended in buffer A (50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl, 10 mM imidazole, 1 mM PMSF), disrupted by passage through a French Press and centrifuged at 14000 x g for 30 min at 4°C. The recombinant proteins were purified by affinity chromatography on His-Select Nickel Affinity Gel (Sigma). The lysate was loaded onto His-Select Nickel Affinity Gel equilibrated with equilibration buffer (50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl, 10 mM imidazole). After 1 min of incubation at 4°C, the matrix was collected by centrifugation at 11000 x g for 1 min and washed 3 times with same equilibration

Table 1: Bacterial strains, plasmids and oligonucleotides.

Strains/plasmids	Description	Reference or source
***E. coli* strains**		
C41 (DE3)	Strain that derives from BL21 [F⁻*ompT hsdS*$_B$ (r_B- m_B-) *gal dcm* (DE3)]. This strain has at least one uncharacterized mutation that prevents cell death associated with expression of many toxic recombinant proteins	Miroux *et al.* (1996)
MG1655	F⁻ wild-type K-12 strain	Volkert
MV5924	Strain derivative of MG1655 in which the *aidB* gene had been inactivated by a tetracycline resistance cassette	Volkert
Plasmids		
pET28a(+)	Plasmid carrying an N-terminal His•Tag®/thrombin/T7•Tag® configuration plus an optional C-terminal His•Tag sequence.	Novagen
pET28a-P*lac*-*aidB*	pET28a-P*lac*Δ(*BamH*I-*Hind*III)Ω(*aidB* gene)	Previous work
pET28a-P*lac*-PP4780	pET28a-P*lac*Δ(*BamH*I-*Hind*III)Ω(PP4780 gene)	This work
pET28a-*aidB*	pET28aΔ(*BamH*I-*Hind*III)Ω(*aidB* gene)	Amoresano *et al*., 2009
pET28a-PP4780	pET28a(*BamH*I-*Xho*I)Ω(PP4780 gene)	This work
Oligonucleotides		
PP4780 Fw	5'-ATAGGATCCATGAGCCTGCACCAG-3'	
PP4780 Rv	5'-TTACTCGAGCAACAGGGGCCAGG-3'	
PP4780-compl Rv	5'-TTAAAGCTTCAACAGGGGCCAGG-3'	

buffer. The recombinant proteins were eluted with buffer containing 250 mM imidazole in 50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl. Protein concentration was estimated with Bradford reagent (Bio-Rad protein assay) and protein content was checked by SDS-polyacrilamide gel electrophoresis (SDS-PAGE).

Growth profiles.

MV1161 (wild type) and MV5924 (Δ*aidB*) *E. coli* strains were not transformed or transformed, individually, with pET28a-P*lac*-PP4780 and pET28a-P*lac*-*aidB*. These bacterial cultures grown overnight in LB medium at 37°C, were diluted 1:100 in fresh medium. At an $A_{600\ nm}$ of 0.2, the cultures were divided in four aliquots supplemented with MMS 0.04%, MNU 0.04%, MNNG (5µg/ml) and ENNG (5µg/ml), respectively. After addition of alkylating agent, the bacterial cells were allowed to grow overnight. For all experiments, each growth profile was reproduced three times.

Isovaleryl-CoA dehydrogenase activity assay.

Isovaleryl-CoA dehydrogenase activity assays were carried out at room temperature in 200 mM phosphate buffer, pH 8.0, and using purified recombinant proteins that had been dialyzed to remove imidazole. For routine assays, 2 mM isovaleryl-CoA (Sigma) was used as the substrate and 0.1 mM 2,6-dichlorophenolindophenol (DCPIP) was used as the terminal electron acceptor in a final volume of 300 µl. The change in absorbance at 600 nm was monitored by using a Beckman DU 7500 spectrophotometer, and the enzyme activity was calculated by assuming an extinction coefficient of 20.6 mM^{-1} cm^{-1} for DCPIP.

RESULTS

AidB increases cell survival after exposure to nitrosoguanidines.

Inducible resistance of *E. coli* to the cytotoxic and mutagenic effects of alkylating agents involves the increased expression of the *ada-alkB* operon, *alkA* and *aidB* genes (Lindahl, T. *et al*., 1988). The DNA repair mechanisms of Ada, AlkA, and AlkB have been structurally and functionally characterized (Landini, P. and Volkert, M.R. 1995; Nakabeppu, Y. *et al*., 1984; Trewick, S. C. *et al*., 2002); recently, we demonstrated that AidB prevents alkylation damage rather than repair it. In order, to completely define the role of AidB in preventing alkylators-induced toxicity, the effect of its deletion on sensitivity to different alkylating compounds was tested. Specifically, we investigated the AidB response to four different mutagens; these include alkyl sulfonates (MMS), nitrosoureas (MNU), and nitrosoguanidines (MNNG or ENNG). To this aim, both MG1655 (wild type) and MV5924 (∆*aidB*) *E. coli* strains were grown in LB medium, in the presence of sub-inhibitory concentrations of the mutagens described above and their growth behaviour was analyzed and compared. As shown in Figs. 1A and 1B, the *aidB* mutant cells don't display any change in sensitivity to MMS or MNU as compared with the wild type cells; in contrast, the inactivation of *aidB* gene was shown to decrease resistance to lethal and mutagenic effects of MNNG and ENNG (Figs. 1C and 1D). Interestingly, the complementation of the ∆*aidB* mutation by pET28a-P*lac-aidB* plasmid restored the cellular resistance to nitrosoguanidines; in fact, the complementated strain was shown to have a growth profile identical to wild type cells (Figs. 1C and 1D). Taken together, these results suggested that AidB plays a critical role in preventing toxicity induced by MNNG and ENNG. Since sensitivity to only nitrosoguanidines was affected in MV5924 strain, we hypothesized that AidB may be involved, by a dehydrogenase mechanism, in a detoxification pathway of this class of alkylators that constitute one of the most important environmental contaminants influencing health of the persons exposed. This result is consistent with previous structural studies performed by Bowles *et al*., 2008. Hence, on the basis of these observations, AidB was shown to represent a promising tool for the bio-treatment of sites contaminated by alkylating agents.

AidB homologues in Pseudomonas species.

Since, ideally, developing successful strategies for the bioremediation requires the knowledge of the microorganisms present in the polluted environments, their metabolic abilities, and how they respond to changes in environmental conditions, this experimental work was targeted at identify as well as at characterize *E. coli* AidB homologues in bacteria used for bioremediation applications. To this aim, the deduced amino sequence of the 60.5 kDa AidB was submitted to Swiss Protein and GenBank-EMBL databases. Interestingly, we found that AidB homologues are not present in many bacteria closely related to *E. coli* (*Klebsiella*, *Vibrio*, *Shewanella*, and *Photorhabdus*), but the closest hits are from some γ-proteobacteria (*Pseudomonas, Azotobacter*, and *Acinetobacter*), β-proteobacteria (*Burkholderia*, *Ralstonia*, *Bordetella*), and even gram-positive bacteria (*Mycobacterium*, *Nocardia*). This observation was intriguing because *Pseudomonas* species and closely related organisms are the most extensively studied and the most frequently used for bioremediation applications (Wackett, L.P. 2003), owing to their ability to degrade many several organic pollutants. Particularly, the homology search revealed that AidB shows 46% identity with the putative acyl-CoA dehydrogenase coded by the PP4780 gene from *Pseudomonas putida* KT2440.

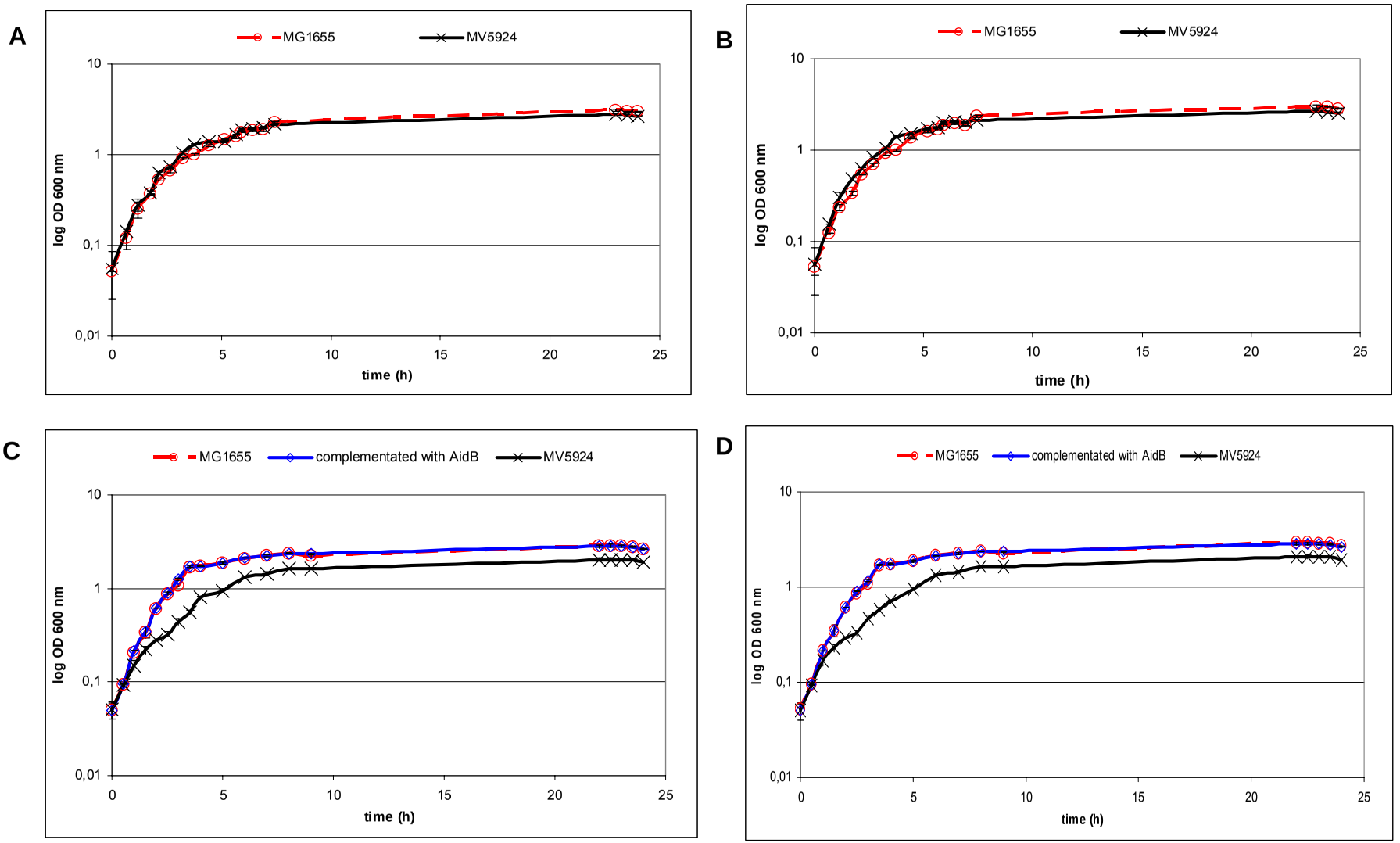


Figure 1: Growth profiles of MG1655 (wild type) and MV5924 (Δ*aidB*) *E. coli* strains grown in the presence of MMS 0.04% (A), MNU 0.04% (B), MNNG 5 µg/ml (C) and ENNG 5 µg/ml (D), respectively. In Fig. 1C and 1D, it is also reported the growth curve of MV5924 transformed with pET28a-P*lac-aidB* plasmid with the aim to create an *aidB*-proficient background.

Interestingly, the conservation of Glu425, that corresponds to the catalytic residue in most of the short- and medium- chain acyl-CoA dehydrogenases, suggested that a dehydrogenase or oxidase activity is essential to the physiological function of the *P. putida* protein. Strain KT2440, whose genome has recently been sequenced (Nelson, K.E. *et al*., 2002), is one of the best characterized pseudomonads; its broad metabolic versatility, genetic plasticity, and ability to colonize many different environments such as bulk soil and the rhizosphere make this bacterium an ideal candidate for bioremediation applications (Clarke, P. 1982; Molina, L. *et al*., 1998; Espinosa-Urgel, M. *et al*., 2002; Timmis, K. N. 2002; Jiménez, J. I. *et al*., 2002; Canovas, D. *et al*., 2003; Santos, P. M. *et al*., 2004; Ramos-Gonzalez, M. I. *et al*., 2005; Kurbatov, L. *et al*., 2006; Nogales, J. *et al*., 2008). *P. putida* was shown to adapt to a variety of stress conditions and interestingly, the changes that may occur in soil include the fluctuation in levels of alkylating agents. This bacterium has also been certified as a biosafety system, which means that it can be used as a host for containment systems (Molin, S. *et al*., 1993) both for applications in biotechnological production and release into the environment. On the basis of these observations, it was intriguing to investigate the possible involvement of the acyl-CoA dehydrogenase from *P. putida* in the response to alkylation stress.

Cloning and over-expression of the PP4780 gene in E. coli.
To characterize the putative acyl-CoA dehydrogenase from *P. putida*, the first step of our work was to clone the PP4780 gene into a commercial expression vector of the pET series and to over-express the recombinant product in *E. coli* cells; the protein bearing a 6 His tag at the C-terminus was purified by affinity chromatography on Ni^{2+}-agarose beads and its homogeneity was tested by SDS-PAGE and mass fingerprinting analyses. Successively, a preliminary characterization of the purified protein was performed examining its potential enzymatic properties. To this aim, the isovaleryl-CoA dehydrogenase (IVD) activity of the *P. putida* protein was examined and compared with that of *Ec*AidB. The product coded by the PP4780 gene displayed low levels of isovaleryl-CoA catalytic activity that are identical to that exhibited by AidB protein (Table 2). It is intriguing to speculate that, such for AidB, fatty acyl-coA are not substrates for the *P. putida* enzyme but that its active site may act as a sink for certain alkylating compounds or their reactive intermediates.

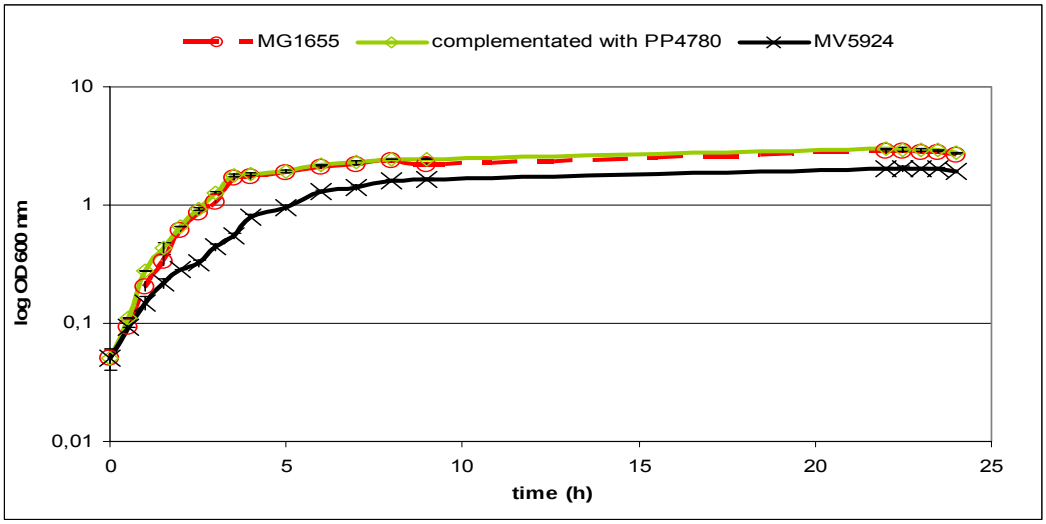
Table 2: Isovaleryl-CoA dehydrogenase activity of the *P. putida* protein coded by the PP4780 gene and the *E. coli* AidB was assayed as described in Materials and Methods following the isovaleryl following the isovaleryl-CoA-dependent reduction of the electron acceptor DCPIP. The data shown are averages of 10 measurements.

Protein	Isovaleryl-CoA dehydrogenase activity
	μmol min^{-1} [mg protein]$^{-1}$
AidB	0.12 ± 0
PP4780	0.13 ± 0

PP4780 protects the bacterial cell from alkylation damage.

The putative involvement of PP4780 in cellular protection against alkylating agents was investigated by transforming MV5924 (Δ*aidB*) *E. coli* strain with a plasmid carrying the PP4780 gene and by analyzing the effect of this gene product on sensitivity of *aidB* mutant cells to nitrosoguanidines. To this aim, the strain complementated was exposed to sub-inhibitory concentrations of MNNG or ENNG and its growth profile was compared with that determined for the wild type MG1655 and Δ*aidB* MV5924 strains treated with nitrosoguanidines. As shown in Figs. 2A and 2B, the PP4780 gene suppressed the increased nitrosoguanidines-sensitivity caused by *aidB* mutation; in fact the complementated strain was shown to have a growth profile identical to wild type cells.

A



B

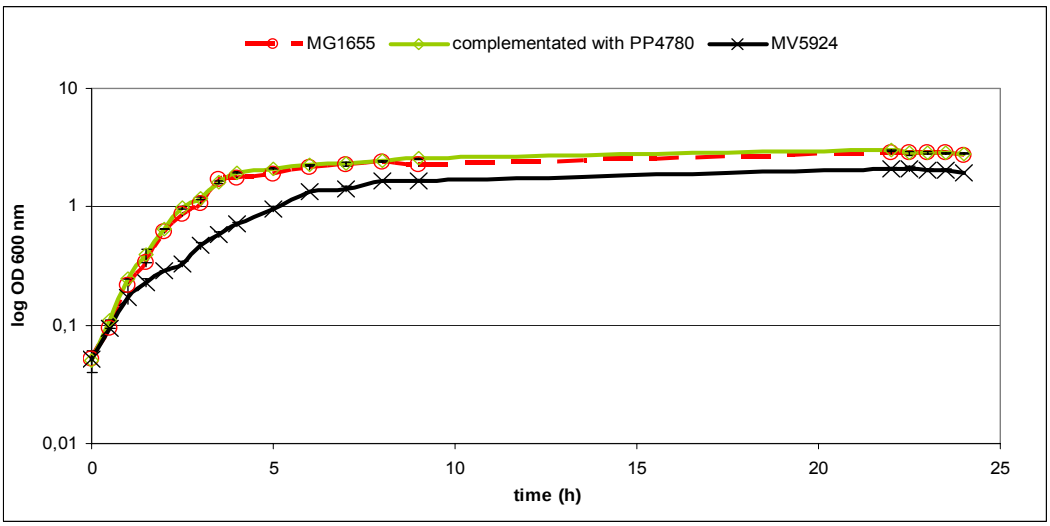
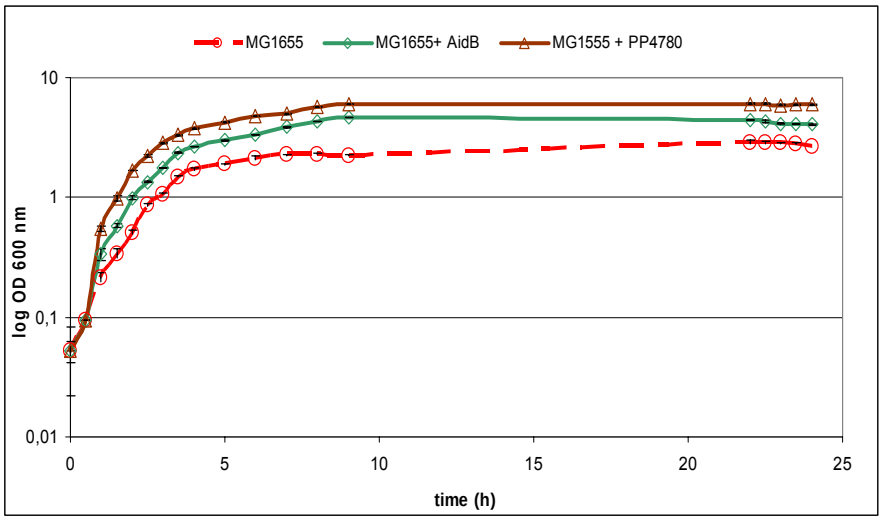


Figure 2: Growth profiles of MG1655 (wild type), MV5924 (Δ*aidB*) and MV5924 transformed with pET28a-P*lac*-PP4780 plasmid, grown in the presence of MNNG 5 μg/ml (A) and ENNG 5 μg/ml (B), respectively.

This finding clearly demonstrated that PP4780 expression produced a functional protein that is involved in the resistance against alkylation damage, presumably by acting as a detoxification enzyme which inactivates nitrosoguanidines.

In order, to further define the *P. putida* protein role in preventing toxicity after nitrosoguanidines challenge, we also tested whether PP4780 over-expression might confer increased resistance to *E. coli* cells treated with both MNNG and ENNG. Interestingly, recombinant cells that over-express PP4780 exhibited to be more resistant to nitrosoguanidines as compared with wild type and *aidB*-overexpressing cells (Figs. 3A and 3B).

A



B

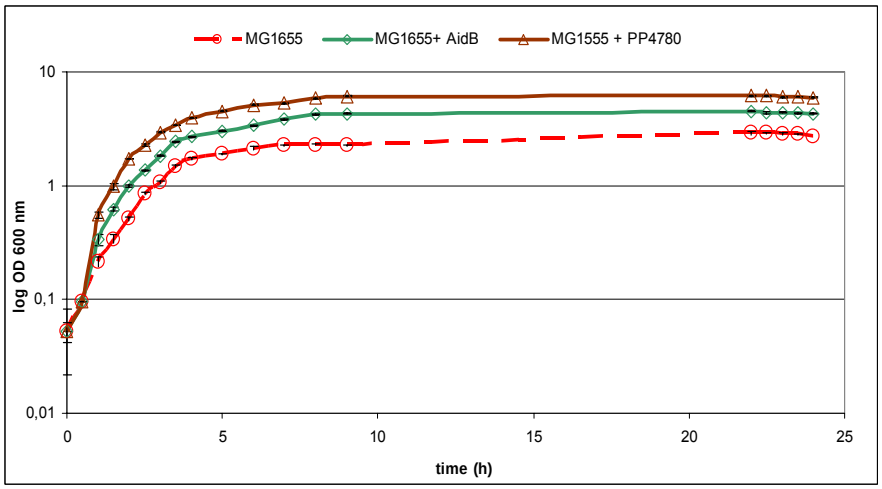


Figure 3: Growth profiles of MG1655 (wild type), MG1655 over-expressing AidB and MG1655 over-expressing PP4780, grown in the presence of MNNG 5 µg/ml (A) and ENNG 5 µg/ml (B), respectively.

Taken together these results, we clearly demonstrated that the *P. putida* protein plays a critical role in defending the bacterial cell against alkylation stress; the putative involvement of this protein in detoxification of nitrosoguanidines was considered as a possible explanation for the phenotypes observed.
Clearly, further studies will be required to elucidate whether *E. coli* AidB and *P. putida* protein effectively act as detoxification enzymes to destroy nitrosoguanidines that are recognised as one of the most abundant environmental pollutants. However, the data obtained in this research project strongly support the possibility of developing new successful strategies for the bioremediation of environments and industrial effluents contaminated by alkylating compounds.

REFERENCES

- Bowles, T., Metz, A. H., O'Quin, J., Wawrzak, Z., and Eichman, B. F. *Proceedings of the National Academy of Sciences* **105**(40), 15299–15304 (2008).
- Cánovas, D., Cases, I., and de Lorenzo, V. *Environmental microbiology* **5**(12), 1242–1256 (2003).
- Clarke, P. *Antonie van Leeuwenhoek* **48**(2), 105–130 (1982).
- Espinosa-Urgel, M., Kolter, R., and Ramos, J. L. *Microbiology (Reading, England)* **148**(Pt 2), 341–343 (2002).
- Jiménez, J. I., Miñambres, B., García, J. L., and Díaz, E. *Environmental microbiology* **4**(12), 824–841 (2002).
- Kataoka, H., Yamamoto, Y., and Sekiguchi, M. *J. Bacteriol.* **153**(3), 1301–1307 (1983).
- Kurbatov, L., Albrecht, D., Herrmann, H., and Petruschka, L. *Environmental Microbiology* **8**(3), 466–478 (2006).
- Landini, P. and Volkert, M. R. *The Journal of biological chemistry* **270**(14), 8285–8289 (1995).
- Landini, P., Hajec, L. I., and Volkert, M. R. *Journal of bacteriology* **176**(21), 6583–6589 (1994).
- Lawley, P. D. *Mutation research* **23**(3), 283–295 (1974).
- Lindahl, T., Sedgwick, B., Sekiguchi, M., and Nakabeppu, Y. *Annual review of biochemistry* **57**, 133–157 (1988).
- Miller, J. H. *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Pr.
- Molin, S., Boe, L., Jensen, L. B., Kristensen, C. S., Givskov, M., Ramos, J. L., and Bej, A. K. *Annual review of microbiology* **47**, 139–166 (1993).
- Molina, L., Ramos, C., Ronchel, M.-C., Molin, S., and Ramos, J. L. *Appl. Environ. Microbiol.* **64**(6), 2072–2078 (1998).
- Nakabeppu, Y., Kondo, H., and Sekiguchi, M. *Journal of Biological Chemistry* **259**(22), 13723–13729 (1984).
- Nelson, K. E., Weinel, C., Paulsen, I. T., Dodson, R. J., Hilbert, H., dos Santos, M., Fouts, D. E., Gill, S. R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R. T., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Lee, P. C., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J., Timmis, K. N., Düsterhöft, A., Tümmler, B., and Fraser, C. M. *Environmental Microbiology* **4**(12), 799–808 (2002).
- Nogales, J., Palsson, B., and Thiele, I. *BMC Systems Biology* **2**(1), 79+ (2008).
- Ramos-Gonzalez, M. I., Campos, M. J., and Ramos, J. L. *J. Bacteriol.* **187**(12), 4033–4041 (2005).
- Rohankhedkar, M. S., Mulrooney, S. B., Wedemeyer, W. J., and Hausinger, R. P. *Journal of bacteriology* **188**(1), 223–230 (2006).
- Santos, P. M., Benndorf, D., and Sá-Correia, I. *Proteomics* **4**(9), 2640–2652 (2004).
- Taverna, P. and Sedgwick, B. *Journal of bacteriology* **178**(17), 5105–5111 (1996).
- Timmis, K. N. *Environmental microbiology* **4**(12), 779–781 (2002).
- Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. *Nature* **419**(6903), 174–178 (2002).
- Vaughan, P., Sedgwick, B., Hall, J., Gannon, J., and Lindahl, T. *Carcinogenesis* **12**(2), 263–268 (1991).
- Volkert, M. R. and Nguyen, D. C. *Proceedings of the National Academy of Sciences of the United States of America* **81**(13), 4110–4114 (1984).
- Wackett, L. P. *Nat Biotech* **21**(2), 136–138 (2003).

Conclusions

CONCLUSIONS

The strategy of using and manipulating detoxification abilities of microorganisms to reduce environmental contamination, such as in soils and waste waters, appears to be feasible alternative to physicochemical methods. This work has been aimed at exploring new potential candidates for the bio-treatment of wastes and environments contaminated by alkylating compounds that are among abundant pollutants present in the ecosystem. The study has been specifically focused on the AidB protein, a component of the adaptive response to alkylating agents in bacterium *Escherichia coli*.

Initially, the role of AidB in the bacterial cell was investigated through structural and functional characterization of its protein domains. Successively, the mechanism by which AidB directly protects against alkylation stress has been determined: AidB prevents DNA damage, presumably by inactivating certain alkylators before they are able to react with their DNA target.

Finally, taking into account its potential role in the detoxification of alkyl-nitrosoguanidines, this experimental work was targeted at identify as well as at characterize *E. coli* AidB homologues in bacteria used for bioremediation applications. Specifically, the study has been focused on the acyl-CoA dehydrogenase coded by the PP4780 gene from *Pseudomonas putida* KT2440. The involvement of this protein in preventing toxicity induced by nitrosoguanidines has been demonstrated: the complementation of the *aidB* mutation by PP4780 gene was shown to decrease the cellular sensitivity to nitrosoguanidines and the over-expression of PP4780 was shown to confer increased resistance to lethal and mutagenic effects of these damaging agents. The potential role for the *P. putida* protein in detoxification of nitrosoguanidines was considered as a possible explanation for the phenotypes observed. Clearly, characterizing the degradative properties of the acyl-CoA dehydrogenase from *E. coli* and *P. putida* KT2440 will be required to allow their potential application in bioremediation field. Besides, further work should be done aimed at exploring the entire metabolic pathways necessary for the complete degradation of alkylating compounds. In light of these observations, the current research should be focused on both identification of enzymatic activities involved in the catabolic pathway of alkylators degradation, as well as on testing their catalytic performances.

PUBLICATIONS

- Papa R, Rippa V, Sannia G, Marino G, Duilio A (2007) An effective cold inducible expression system developed in *Pseudoalteromonas haloplanktis* TAC125. *J Biotechnol* 127, 199-210.
- Papa R, Rippa V, Duilio A (2009) Identification of the transcription factor responsible for L-malate-dependent regulation in the marine Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. *FEMS Microbiol Lett* 295, 177-186.
- Amoresano A, Cirulli C, Rippa V, Esposito C, Papa R, Duilio A (2009) Additional protein functions involved in the biogenesis of ribosomes and the DNA repair mechanisms are associated with the transcriptional machinery gathered at the *Escherichia coli rrnB* P1 promoter. *Mol Microbiol*-submitted.
- Rippa V, Cirulli C, Di Palo B, Doti N, Amoresano A, Duilio A (2009) The ribosomal protein L2 interacts with the RNA polymerase α subunit and acts as a transcription modulator in *Escherichia coli. J Bacteriol*-submitted.

COMMUNICATIONS

- Cirulli C, Amoresano A, Rippa V, Esposito C, Doti N, Papa R, Duilio A. A proteomic approach in the investigation of transcriptional regulatory network in *E. coli*. HIUPO 2006, Pisa, Italy.
- Rippa V, Amoresano A, Cirulli C, Esposito C, Duilio A. Investigation on transcriptional regulatory network in *Escherichia coli*: role of the AidB protein. Congresso Prokaryotes 2007, Cortona, Italy.
- Amoresano A, Cirulli C, Rippa V, Papa R, Esposito C, Duilio A. The transcriptional machinery gathered at the *E. coli rrnB* P1 promoter includes proteins involved in the biogenesis of ribosome and the DNA repair mechanisms. ItPA 2007, Catania, Italy.
- Rippa V, Esposito C, Cirulli C, Duilio A. Regolazione della trascrizione in *Escherichia coli*: ruolo della proteina AidB. Giornate scientifiche interpolo 2007, Napoli, Italy.
- Rippa V, Amoresano A, Duilio A. Transcriptional role of *E. coli* AidB protein. Congresso Prokaryotes 2008, Cortona, Italy.
- Di Palo B, Rippa V, Esposito C, Amoresano A, Duilio A. Transcriptional role of ribosomal protein L2 in *Escherichia coli.* Congresso Prokaryotes 2008, Cortona, Italy.
- Rippa V, Di Procolo P, Esposito C, Duilio A. Characterisation of transcriptional factor AidB in *Escherichia coli*. Giornate scientifiche interpolo 2008, Napoli, Italy.

INTERNATIONAL COURSES

- **OSSIBE 4**, the 4th Oulu Summer School in Bioprocess Engineering “Protein production- what can go wrong and haw can you improve it fast?" Oulu, Finland, June 11-15th 2007.

RESEARCH ACTIVITY IN FOREIGN LABORATORIES

- From July 15th to October 31th 2009, my research activity was performed in the Laboratory of Prof. Michael Volkert, Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester (Massachusetts).

Other publications

The ribosomal protein L2 interacts with the RNA polymerase α subunit and acts as a transcription modulator in *Escherichia coli*

Valentina Rippa[1], Claudia Cirulli[1], Benedetta Di Palo[1], Nunzianna Doti[1], Angela Amoresano[1,] and Angela Duilio[1]*

[1]Department of Organic Chemistry and Biochemistry, School of Biotechnological Sciences, Federico II University of Naples - Naples – Italy.

Abstract

Identification of interacting proteins in stable complexes is essential to understand the mechanisms that regulate cellular processes at the molecular level. Transcription initiation in prokaryotes requires coordinated protein-protein and protein-DNA interactions often involving one or more transcription factors in addition to RNA polymerase (RNAP) subunits. The RNAPα subunit is a key regulatory element in gene transcription which functions through direct interaction with other proteins to control all stages of the process. A clear description of the RNAPα protein partners will greatly benefit the understanding of transcription modulation.

A functional proteomic approach was employed to investigate protein components specifically interacting with RNAPα. A tagged form of the *Escherichia coli* RNAPα was used as a bait to fish its molecular partners out from the entire cellular extract. Among other interactors, the 50S ribosomal protein L2 (RPL2) was clearly identified by mass spectrometry. The direct interaction between RNAPα and RPL2 was further confirmed both *in vivo* and *in vitro* by co-immunoprecipitation and bacterial two-hybrid experiments. The functional role of this interaction was also investigated in the presence of a ribosomal promoter by using a β-galactosidase gene reporter assay. The results clearly demonstrated that RPL2 was able to increase β-galactosidase expression only in the presence of a specific ribosomal promoter whereas it was inactive when assayed with an unrelated promoter. Interestingly, other ribosomal proteins (L1, L3, L20, L27) did not show any influence on rRNA expression. Findings reported here strongly suggest that besides ribosome assembly the highly conserved RPL2 protein plays also a specific and direct role in transcription regulation.

Keywords: L2 ribosomal protein, RNA polymerase, transcription regulation, functional proteomics, mass spectrometric procedures.

Running title: Transcriptional role of ribosomal protein L2 in *Escherichia coli*.

* Corresponding Author:
Angela Duilio
Dipartimento di Chimica Organica e Biochimica,
Università di Napoli “Federico II”
Complesso Universitario di Monte Sant’Angelo, Via Cinthia 4, 80126 Napoli (Italy)
Tel +39-081674314
Fax +39-081674313
email: anduilio@unina.it

Introduction

Understanding the mechanism and regulation of transcription in bacteria requires dissection of the specific roles of the individual components of the multi-protein transcription complex including the multi-subunit enzyme RNA polymerase (RNAP). It is well known, in fact, that one of the major factors ensuring correct gene expression in microorganisms is the

efficiency with which RNAP recognises the specific promoters of different genes. The single form of RNAP in *Escherichia coli* consists of a tetrameric core enzyme ($\beta\beta'\alpha_2$) capable of RNA synthesis and factor-independent termination, and the σ subunit responsible for recognition of specific transcription initiation sites. However, although RNAP σ constitutes the main determinant for promoter recognition, α subunit also plays a key role in the stability of the transcription complex (3,10). The *E. coli* RNAP α subunit consists of two domains, the amino terminal domain (NTD) and the carboxy terminal domain (CTD) connected by a flexible linker that allows the CTD domain to interact with promoter elements located at different distances from the RNAP binding site (2). The α subunit fulfils at least two functions during genes transcription: it is involved in both the assembly of the core enzyme and the regulation of transcription initiation. The role of the two domains is clear: NTD is needed for dimerization and interaction with the β and β' subunits whereas CTD is important for both DNA binding and protein-protein interactions. Transcription initiation requires coordinated protein-protein and protein-DNA interactions often involving one or more transcription factors in addition to RNAP subunits (19). Transcription factors can interact with nearly all RNAP components, although αCTD constitutes the most frequent target. Class I activators bind to an upstream site and contact the αCTD thereby recruiting RNAP to the promoter (6, 20). The α subunit is then a key regulatory element in gene transcription which functions through direct interaction with other proteins to control all stages of the process. A clear description of the α subunit protein partners will greatly benefit the understanding of transcription modulation. Recently functional proteomics approaches essentially based on pull-down or immunoprecipitation experiments (17) have been introduced to investigate protein-protein interactions both *in vitro* and *in vivo*. The association of a specific protein with molecular partners belonging to protein complexes involved in particular mechanisms can greatly contribute to the description of the cellular processes at the molecular level (9, 12).

In this work, a 6-His tagged recombinant form of the α subunit from *E. coli* RNAP (RNAPα) was used as a bait to fish its molecular partners out from the entire cellular extract. Among other interactors, the 50S ribosomal protein L2 (RPL2) was clearly identified by mass spectrometric procedures. The direct interaction between RNAPα and RPL2 was further confirmed both *in vivo* and *in vitro* by co-immunoprecipitation experiments and two hybrid technique. The functional role of this interaction was investigated by *in vivo* transcription assays using *lacZ* as reporter gene. RPL2 was shown to increase β-galactosidase expression only when a specific ribosomal promoter was used in the assay, whereas it was inactive in the presence of an unrelated promoter. Interestingly, other ribosomal proteins (L1, L3, L20, L27) did not show any effect on rRNA expression. Findings reported here strongly suggest that besides ribosome assembly the highly conserved RPL2 protein plays also a specific and direct role in transcription regulation.

Material and methods

Bacterial strains and plasmids.

The bacterial strains and the plasmids used in this work are listed in Table 1.

Media and chemicals.

Luria-Bertani or nutrient broths (for

Table 1: Bacterial strains and plasmids.

Strains/plasmids	Description	Reference or Source
Strains		
E. coli K12	wild type	DSMZ
C41 (DE3)	Strain that derives from BL21 [F⁻*ompT hsdSB* (r_B-m_B-) *gal dcm* (DE3)]. This strain has at least one uncharacterized mutation that prevents cell death associated with expression of many toxic recombinant proteins	ref. 15
R721	71/18 glp O-$P_{434}/_{p22}$*lacZ*	ref. 5
Plasmids		
pET22b(+)	carries an N-terminal *pelB* signal sequence, plus an optional C-terminal His-tag sequence; ampicillin resistance	Novagen
pET28a(+)	carries an N-terminal His•Tag®/thrombin/T7•Tag® configuration, plus an optional C-terminal His•Tag sequence; kanamycin resistance	Novagen
pET22b-αpol	pET22b(+)Δ(*Nde*I-*Xho*I)Ω(*rpoA* gene)	This work
pET22b-c-Myc	pET22b(+)Δ(*Nde*I-*BamH*I)Ω(*c-myc* tag)	This work
pET22b-c-Myc-L2	pET22b-c-MycΔ(*BamH*I-*Xho*I)Ω(*rplB* gene)	This work
pET22b-c-Myc-Thioredoxin	pET22 b-c-MycΔ(*BamH*I-*Xho*I)Ω(*trxC* gene)	This work
pET28a-αpol	pET28aΔ(*BamH*I- *Xho*I)Ω(*rpoA* gene)	This work
pcI_{P22}	pC132 derivative carrying N-terminal end of p22 repressor; ampicillin resistance	ref. 5
pcI_{434}	pACYC177 derivative carrying N-terminal end of 434 repressor; chloramphenicol resistance	ref. 5
pcI_{P22}-RNAPα	pcI_{P22}Δ(*Sal*I- *BamH*I)Ω(*rpoA* gene)	This work
pcI_{434}-RNAPα	pcI_{434}Δ(*Sal*I- *BamH*I)Ω(*rpoA* gene)	This work
pcI_{434}-RPL2	pcI_{434}Δ(*Sal*I- *BamH*I)Ω(*rplB* gene)	This work
pcI_{434}-L2	pcI_{434}Δ(*Nde*I- *BamH*I)Ω(*rplB* gene)	This work
pcI_{434}-L1	pcI_{434}Δ(*Nde*I- *BamH*I)Ω(*rplA* gene)	This work
pcI_{434}-L3	pcI_{434}Δ(*Nde*I- *BamH*I)Ω(*rplC* gene)	This work
pcI_{434}-L20	pcI_{434}Δ(*Nde*I- *BamH*I)Ω(*rplT* gene)	This work
pcI_{434}-L27	pcI_{434}Δ(*Nde*I- *BamH*I)Ω(*rpmA* gene)	This work
pET22b-P*rrnD*	pET22bΔ(*Sph*I-*Hind*III)Ω(*rrsD* promoter)	This work
pET22b-P*aidB*	pET22bΔ(*Sph*I-*Hind*III)Ω(*aidB* promoter)	This work
pET22b-P*rrnD-lacZ*	pET22b-P*rrnD*Δ(*Hind*III-*Xho*I)Ω(*lacZ* gene)	This work
pET22b-P*aidB-lacZ*	pET22b-P*aidB*Δ(*Hind*III-*Xho*I)Ω(*lacZ* gene)	This work

bacterial cultures and plating) and suspension medium (for bacterial dilutions) were used as described by Miller (14). Ampicillin, kanamycin and chloramphenicol (Sigma) were used at 100, 50 and 34 μg/ml, respectively.

Construction of the expression plasmids.

The *rpoA*, *rplA*, *rplB*, *rplC*, *rplT*, *rpmA* and *trxC* genes were amplified from DNA genomic of *E. coli* by PCR using the specific primers listed in Table 2. The amplified fragments were digested with the restriction enzymes underlined in Table 1 and cloned into the commercial expression vectors which were linearized with the same restriction enzymes. To obtain pET22b-c-Myc, a *Nde*I/*BamH*I digested fragment corresponding to the c-Myc epitope was inserted into the

pET22b(+) expression plasmid (Novagen) linearized with *Nde*I and *BamH*I.
The resulting plasmids, designated as listed in Table 1, were verified by automated DNA sequencing.

Production and purification of recombinant proteins.
The constructs pET22b-αpol, pET22b-c-Myc-L2, pET22b-c-Myc-Thioredoxin and pET28a-αpol (Table 1) were individually transformed into the *E. coli* strain C41 (DE3). For αRNAP and Thioredoxin production, the recombinant cells were grown in LB at 37°C without induction until the OD_{600} reached 3.0. The *rplB* gene was expressed as follows: recombinant cells were grown at 37°C to an optical density at 600 nm of ~ 0.9, at which time 0.05 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) was added and the cultures were allowed to grow until the OD_{600} reached 3.0. Cells were harvested by centrifugation at 5000 x g for 15 min at 4°C, resuspended in buffer A (50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl, 10 mM imidazole, 1 mM PMSF), disrupted by passage through a French Press and centrifuged at 14000 x g for 30 min at 4°C. The recombinant proteins were purified by affinity chromatography on His-Select Nickel Affinity Gel (Sigma) according to slightly modified manufacturer's instructions. The lysate was loaded onto His-Select Nickel Affinity Gel equilibrated with equilibration buffer (50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl, 10 mM imidazole). After 1 min of incubation at 4°C, the matrix was collected by centrifugation at 11000 x g for 1 min and washed 3 times with washing buffer (50 mM Na_2HPO_4, pH 7.4, 0.8 M NaCl, 10 mM imidazole). The recombinant proteins were eluted with buffer containing 250 mM imidazole in 50 mM Na_2HPO_4, pH 7.4, 0.3 M NaCl. Protein concentration was estimated with Bradford assay (1) and protein content was checked by SDS-polyacrylamide gel electrophoresis (SDS-PAGE).

Fishing for partners.
100 µl of His-Select Nickel Affinity Gel (Sigma) were incubated twice with 10 mg of bacterial protein extract for 2 hours at 4°C as preclearing step. The unbound protein extract was then incubated with 6HisRNAPα linked onto agarose beads by His-Ni(2+) interactions for 2 hours at 4°C. The matrix was collected by centrifugation at 5000 x g for 10 min and washed 4 times with washing buffer. The retained proteins were eluted with 100µl of Laemmli buffer (62.5 mM Tris-HCl, pH 6.8, 10% glycerol, 2% sodium dodecyl sulfate (SDS), and 0.1% bromophenol blue) containing 0.1 M DTT (4). The eluate was fractionated by 10 % monodimensional SDS-PAGE and stained with Brilliant Blue G-colloidal Coomassie (Sigma).

In situ digestion.
Coomassie blue-stained protein bands were excised from SDS-PAGE gels and washed in deionised MilliQ grade water (three times, 10 min). The excised spots were then washed first with acetonitrile and then with 0.1 M ammonium bicarbonate (three times, 15 min). Protein bands were then in gel digested as previously described (16).

MALDI MS analyses.
MALDI-TOF mass spectra were recorded using an Applied Biosystem Voyager DE STR instrument. A mixture of analyte and matrix solution (alfa-cyano-hydroxycinnamic acid 10 mg/ml in 66% ACN, 10 mM citric acid, in MilliQ water) was applied to the metallic sample plate and dried down at room temperature. Mass calibration was performed using external peptide standards.

Table 2: Oligonucleotides.

rpoA-pET22b Fw	5' ATACATATGATGCAGGGTTCTGTGACAG 3'
rpoA-pET22b Rv	5' TAACCTGTGCTCGAGTAACTCGTCAGCG 3'
c-myc Fw	5'ATTCATATGGAACAAAAACTCATCTCAGAAGAGGATCTGAAT GGGGCCGCAGGATCCTAT3'
c-myc Rv	5'ATAGGATCCTGCGGCCCCATTCAGATCCTCTTCTGAGATGA GTTTTTGTTCCATATGAAT3'
rplB-pET22b Fw	5' TATGGATCCATGGCAGTTGTTAAATG 3'
rplB-pET22b Rv	5' ATACTCGAGTAATTTGCTACGGCGAC 3'
trxC-pET22b Fw	5' CCGGATCCATGAATACCGTTTGTACCCATTG 3'
trxC-pET22b Rv	5' CCCTCGAGTAAAAGAGATTCGTTCAGCCAG 3'
rpoA-pET28a Fw	5' GAGGGATCCATGCAGGGTTCTGTGACAG 3'
rpoA-pET28a Rv	5' GTGCTCGAGTTACTCGTCAGCGATGC 3'
rpoA-pcI Fw	5' TATCTAGAGCGTCGACCATGCAGGGTTCTGTGACAG 3'
rpoA-pcI Rv	5' ATACTCGAGCGGGATCCTTACTCGTCAGCGATGCTT 3'
rplB-pcl_{434} Fw	5' TATCTAGAGCGTCGACCATGGCAGTTGTTAAATGTA 3'
rplB-pcl_{434} Rv	5' ATACTCGAGCGGGATCCTTATTTGCTACGGCGACG 3'
L2-pcl_{434} Fw	5' TAACATATGATGGCAGTTGTTAAATGTAA 3'
L2-pcl_{434} Rv	5' TTAGGATCCTTATTTGCTACGGCGACG 3'
rplA-pcl_{434} Fw	5' GAGCCATATGCATGGCTAAACTGACCAAGCG 3'
rplA-pcl_{434} Rv	5' GCGGGATCCTTAGTTTACAGAAGCGCTCAGG 3'
rplC-pcl_{434} Fw	5' GAGCCATATGCATGATTGGTTTAGTCGGTAAA 3'
rplC-pcl_{434} Rv	5' GCGGGATCCTTACGCCTTCACAGCTGGTTT 3'
rplT-pcl_{434} Fw	5' GAGCCATATGCATGGCTCGCGTAAAACGTGG 3'
rplT-pcl_{434} Rv	5' GCGGGATCCTTATGCCAGAGCTGCTTTCGC 3'
rpmA-pcl_{434} Fw	5' GAGCCATATGCATGGCACATAAAAAGGCTGG 3'
rpmA-pcl_{434} Rv	5' GCGGGATCCTTATTCAGCTTCGATGCTGATAAA 3'
P*rrnD*-pET22b Fw	5' GTGCGCATGCACAGAAAAAAAGATC 3'
P*rrnD*-pET22b Rv	5 'TCGAAGCTTCGGAGGCGCATTATAG 3'
P*aidB*-pET22b Fw	5' GTGCGCATGCATAAGAATGTTTTAGC 3'
P*aidB*-pET22b Rv	5' TCGAAGCTTCACCATTAGTATGGTC 3'
lacZ-pET22b Fw	5' TGTAAGCTTATAACAATTTCACACAGGAA 3'
lacZ-pET22b Rv	5' CGGCTCGAGTTATTTTTGACACCAGAC 3'

Raw data were analysed using the computer software provided by the manufacturer and reported as monoisotopic masses. Peptide masses of tryptic fragments from each digested protein were used to search for protein databases using an in-house version of the Mascot software (Matrix Science).

Nano-Liquid Chromatography-ESI Tandem Mass Spectrometry (LC/MS/MS).

Tryptic peptide mixtures obtained as previously described were also analysed by nano-LC/ES/MS/MS on-line using a linear ion trap instrument (4000Q-trap Applied Biosystems). The proteolytic digest was fractionated on a HP 1100 nano HPLC apparatus (Hewlett-Packard, Palo Alto, CA) using a capillary C18 column (75 µm x 150 mm, 300 Å) (Torrance, CA) with 0.1% (v/v) formic acid, 2% (v/v) ACN in H_2O (solvent A) and 0.1% (v/v) formic acid, 2% (v/v) H_2O in ACN (solvent B) by means of a linear gradient from 5 to

70% solvent B for 60 min at a flow rate of 0.2 µl/min. The column was directly connected to the ion source through the nanospray probe and both ES/MS and ES/MS/MS spectra were acquired throughout the entire analysis by dependent data scanning, monitoring the five most intense ions.

Co-immunoprecipitation and Western blotting analysis.

For co-immunoprecipitations, *E. coli* strain C41 (DE3) was co-transformed with pET22b-c-Myc-L2/pET28a-αpol and with pET22b-c-Myc-Thioredoxin/pET28a-αpol as control of the experiment. 2 ml of cultures grown overnight were inoculated into 200 ml of LB medium containing 100 µg/ml ampicillin and 50 µg/ml kanamicin. After expression of the recombinant genes without induction, cells were harvested by centrifugation at 5000 x g for 15 min at 4°C, resuspended in 50 mM Na_2HPO_4 (pH 7.4), 1mM PMSF, disrupted by passage through a French Press and centrifuged at 14000 x g for 30 min at 4°C. The protein contents were checked by Western blot analysis. The supernatants were then used for the co-immunoprecipitation experiments. Cell lysates were incubated with agarose-linked T7 antibody (Bethyl) and with agarose beads alone (control of the experiment) at 4°C overnight. The beads were then collected by centrifugation and washed; the bound proteins were eluted with 1×SDS–PAGE sample buffer (2% SDS, 10% glycerol, 62.5 mM Tris-HCl [pH 6.8], 100 mM dithiothreitol, 0.01% bromophenol blue) and subjected to SDS-PAGE followed by Western blot analysis that was performed by using anti-T7 mouse antibody (Novagen) and anti-c-Myc mouse antibody (Calbiochem) as first antibodies and anti-mouse IgG conjugated to peroxidase as a secondary antibody (Calbiochem).

The two-hybrid system: growth conditions and dimerization assay.

The plasmids pcI_{P22}, pcI_{434} and their derivates needed for the two hybrid experiment (Table 1) were introduced into R721 competent cells. The recombinant cells were inoculated into 0.5 ml LB medium, diluted 1:1000 in 10 ml LB supplemented with $1x10^{-4}$ M IPTG and grown with aeration at 34 °C for about 5h. At this time the OD_{600} should range between 0.3 and 0.4 (5).

Assay of β-galactosidase activity was performed as described by Miller (14). The activity was represented in Miller units and was calculated as follows: activity (Miller units) = 1,000 × A_{420}/(time [min] × volume [culture] × optical density at 600 nm).

Construction of pET22b-PrrnD-lacZ and pET22b-PaidB-lacZ for transcription assays.

The *rrnD* P1 and the *aidB* promoter were amplified from DNA genomic of *E. coli* by PCR using the primers listed in Table 2. The amplified products were digested with *Sph*I and *Hind*III and cloned into the pET22b(+) vector which was linearized with the same restriction enzymes. Then the *lacZ* gene was positioned downstream the *rrnD* P1 and the *aidB* promoter using *Hind*III and *Xho*I. Plasmid constructions were verified by automated DNA sequencing.

In vivo transcription assays.

The C41 cells transformed with pET22b-P*rrnD-lacZ* and the C41 cells co-transformed with pET22b-P*rrnD-lacZ*/pcI_{434}-L2 were grown overnight in LB medium at 37°C and then were diluted 1:100 in fresh medium and selective antibiotics were added. As control of the experiment, the C41 cells transformed with pET22b-P*aidB-lacZ* and the C41 cells co-transformed with pET22b-P*aidB-lacZ*/pcI_{434}-L2 were used. At an optical density at 600 nm

of 0.9, IPTG was added (final concentration 0.05 mM) and cellular pellets were collected during the growth. The cells were resuspended in 50 mM Na_2HPO_4 (pH 7.4), disrupted by passage through a French Press and centrifuged at 14000 x g for 30 min at 4°C. The supernatant was collected and protein concentration was determined with the Bio-Rad protein assay (1), using bovine serum albumine as standard. β-galactosidase activity was determined by measuring ONPG-hydrolysis, as described by Miller (14). The putative effect of other ribosomal proteins (L1, L3, L20, L27) on rRNA transcription was also tested by *in vivo* transcription assays. To this aim, the C41 cells were co-transformed with the following constructs: pET22b-P*rrnD-lacZ*/pcI_{434}-L1, pET22b-P*rrnD-lacZ*/pcI_{434}-L3, pET22b-P*rrnD-lacZ*/pcI_{434}-L20, pET22b-P*rrnD-lacZ*/pcI_{434}-L27 and the β-galactosidase activity was monitored during the exponential growth phase. As control of the experiment, the C41 cells co-transformed with pET22b-P*aidB-lacZ* and pcI_{434}-L1/L3/L20/L27 were used.

Results and discussion

Identification of novel proteins associated with the RNAP alpha subunit.

Understanding protein functions as well as unravelling cellular mechanisms at the molecular level constitutes a major challenge in modern biology. Both targets can be addressed through the identification of interacting protein partners *in vivo*. The association of an unknown protein with components belonging to a specific protein complex involved in a particular mechanism would in fact be strongly suggestive of its biological function (11, 13, 17). Furthermore, a detailed description of the cellular signalling pathways might greatly benefit from the elucidation of protein-protein interactions in the cell (7, 8).

Identification of interacting proteins in stable complexes in a cellular system by functional proteomic approach is essentially accomplished by affinity-based procedures. The basic idea is to express the protein of interest with a suitable tag to be used as a bait to fish its specific partners out from the entire cellular extract. Individual components within the multi-protein complex are then fractionated by SDS-PAGE and identified by mass spectrometric methodologies. Investigation of the regulatory transcriptional network in *E. coli* was carried out by identifying transcriptional modulators interacting with the RNAP α subunit employing a "fishing for partners" strategy combined with mass spectrometric procedures. The recombinant plasmid pET22b-αpol was introduced into the *E. coli* cells and the resulting expressed 6x-His tagged RNAP α subunit (6HisRNAPα) was purified by affinity chromatography on a His-Select Nickel Affinity Gel. From 1 liter of bacterial growth about 20 mg of 6HisRNAPα pure protein was obtained. In a classical pull down experiment, the 6HisRNAPα was linked to agarose beads by His-Ni(2+) interactions and the clear bacterial extract was incubated with the bait. The immobilized protein formed stable non-covalent interactions with specific partners occurring in the *E. coli* extract that were retained on the insoluble matrix, while the unbound proteins were eluted by repetitive washings. After mild washes to eliminate non-specific interactions, the protein components specifically recognised by the bait were eluted in Laemli buffer and fractionated onto a 10% mono-dimensional SDS PAGE. An equal amount of bacterial extract was incubated with His-Select Nickel Affinity Gel beads lacking RNAPα

under the same experimental conditions as negative control. Figure 1 shows the resulting SDS-PAGE gel following colloidal Coomassie staining where lane 4 and lane 5 correspond to the negative control and the sample, respectively. Due to the complexity of the gel patterns and the low resolution of 1D electrophoresis, several proteins can occur in the same gel band.

Therefore, protein bands specifically present in the sample lane and absent in the control lane cannot be identified by simply comparing the two gel profiles. The entire lines from the sample and the control were then cut in slices (38 slices each) and each gel slice was submitted to the identification procedure. Proteins contained in the slices were reduced, alkylated and *in situ* digested with trypsin.

The resulting peptide mixtures were directly analysed by mass spectrometry methodologies using both the peptide mass fingerprinting procedure and LC-MS/MS analyses (12, 17). The comparison between the sample and the control was then performed on the basis of the proteins effectively identified in each gel slice. Common proteins identified in both the control and the sample slices were discarded and only those proteins solely identified in the sample slices and absent in the control were selected as putative RNAPα interactors. Following this procedure, the bait protein RNAPα was only identified in the sample slice C19 with an apparent molecular mass of about 40 kDa, whereas it was absent in the corresponding control slice 19 and it was not detected in any gel slice from the control lane.

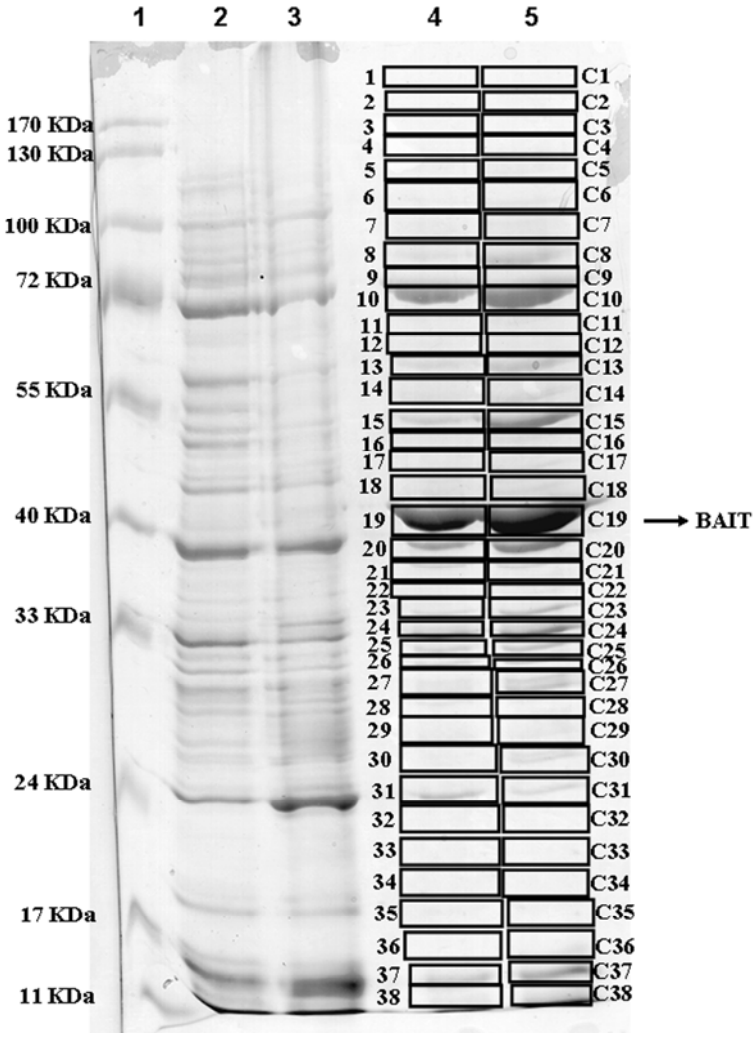
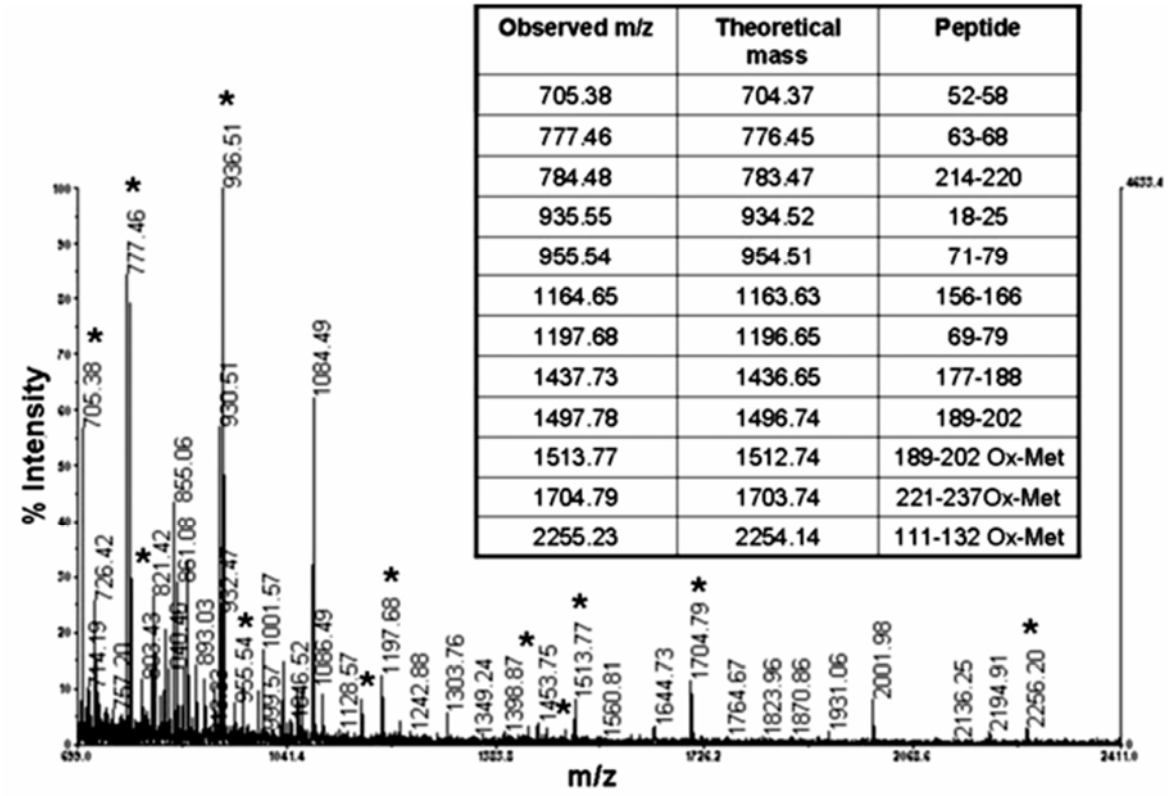


Figure 1. Fractionation of RNAPα interactors by mono-dimensional SDS PAGE.
After a preclearing step, the bacterial extract was incubated with the 6HisRNAPα linked to agarose beads and the protein components specifically recognised by the bait were eluted in Laemli buffer. Lane 1, molecular markers; lane 2, total bacterial extract; lane 3, unbound proteins not retained by the bait; lane 4, protein extract retained by His-Select Nickel Affinity Gel alone and eluted with Laemli buffer (control); lane 5, proteins specifically retained by the bait.

Among many others, we focused our attention on the protein occurring in band C25 at about 30 kDa (Figure 1, lane 5). The MALDI-MS spectrum obtained from an aliquot of the tryptic digest of band C25 is shown in Figure



Observed m/z	Theoretical mass	Peptide
705.38	704.37	52-58
777.46	776.45	63-68
784.48	783.47	214-220
935.55	934.52	18-25
955.54	954.51	71-79
1164.65	1163.63	156-166
1197.68	1196.65	69-79
1437.73	1436.65	177-188
1497.78	1496.74	189-202
1513.77	1512.74	189-202 Ox-Met
1704.79	1703.74	221-237Ox-Met
2255.23	2254.14	111-132 Ox-Met

Figure 2. MALDI MS analysis of the tryptic digest of band C25.
Band C25 was *in situ* digested with trypsin and the resulting peptide mixture extracted from the gel and directly analysed by MALDI MS. Peaks labelled with an asterisk were used to search for a non redundant sequence database using the MASCOT software, leading to the identification of RPL2.

2. Peaks labelled with an asterisk were used to search for a non redundant sequence database using a in-house version of the MASCOT software. The query returned a highly significant match (score 160) with ribosomal protein L2, whose predicted molecular mass of 30 kDa corresponded to the observed electrophoretic mobility of band C25. The mass spectral analyses led to a sequence coverage greater than 50% of RPL2. The remaining portion of the tryptic digest was analysed by LC/MS/MS. Peptides were fractionated by nano-HPLC directly inserted into the nanoelectrospray source and their mass values accurately determined. Peptide ions were simultaneously isolated and fragmented into the mass spectrometer producing daughter ion spectra from which sequence information could be inferred. Several sequence stretches were obtained that matched peptide fragments occurring within RPL2 sequence thus confirming previous MALDI data and indicating RPL2 as a putative RNAPα interactor. No trace of RPL2 could be observed in any gel slice from the control lane.

Verification of RNAPα–RPL2 interaction by co-immunoprecipitation. The interaction between RNAPα subunit and RPL2 proposed by the functional proteomic approach had to be confirmed by both *in vitro* and *in vivo* investigations. First, this interaction was verified by a totally independent approach based on co-immunoprecipitation experiments. Both T7-tagged RNAPα and c-Myc labelled RPL2 were co-expressed in *E. coli* C41 cells under the control of

an IPTG inducible promoter. However, since it is well known that pET vectors always express a small amount of the recombinant protein even in the absence of induction, stimulation of the recombinant cells with IPTG was omitted to avoid unnecessary overproduction of the two proteins. The protein extract from recombinant cells was fractionated by SDS-PAGE and stained by colloidal Coomassie as control. The protein bands corresponding to RNAPα and RPL2 could not be detected, thus ruling out overexpression of the two proteins. The total protein extract was immunoprecipitated by anti-T7 antibody linked to agarose beads and underivatised agarose beads as control. Figure 3A shows the Western blot of the T7-RNAPα immunoprecipitate developed with anti c-Myc antibodies to highlight the presence of RPL2. A protein band

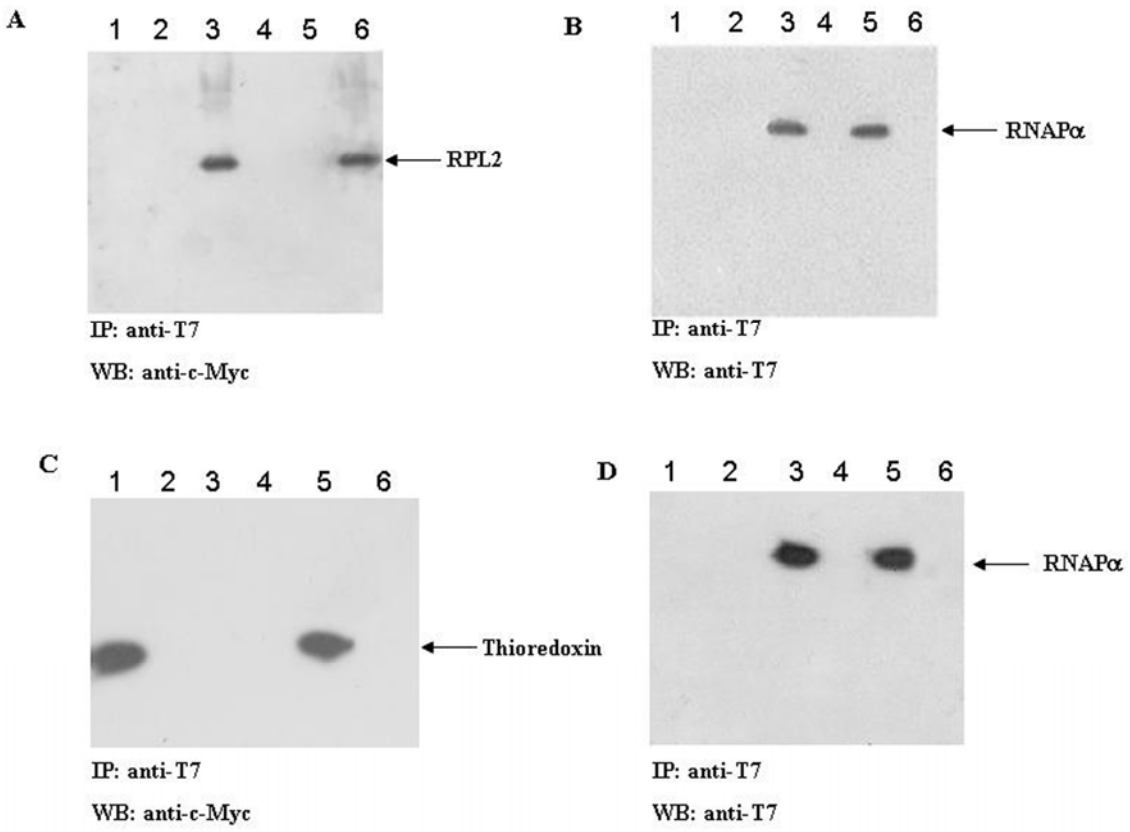


Figure 3. Co-immunoprecipitation experiments.
Total protein extracts from *E. coli* strain C41 co-transformed with pET28a-αpol/pET22b-c-Myc-L2 and with pET28a-αpol/pET22b-c-Myc-Tioredoxin were subjected to immunoprecipitation with agarose-linked T7 antibody and with agarose beads alone as control, followed by immunoblotting with anti-c-Myc antibody (left panel) and with anti-T7antibody (right panel). In this experiment the recombinant RNAPα contains a T7 tag. In all the experiments (A-D), lanes 1-3 corresponding to the unbound, the wash and the elution fraction when anti-T7 antibody beads were used while lane 4 corresponds to the fraction eluted from agarose beads alone. The panels A and B represent the T7-RNAPα:c-Myc-L2 coimmunoprecipitation. In A and B, lane 5 contains the purified T7-tagged form of RNAPα and lane 6 contains the purified c-Myc tagged form of RPL2. The panels C and D represent the absence of T7-RNAPα:c-Myc-Tioredoxin coimmunoprecipitation. In C, lane 5 contains the purified c-Myc tagged form of Thioredoxin and lane 6 contains the purified T7-tagged form of RNAPα. In D, lane 5 contains the purified T7-tagged form of RNAPα and lane 6 contains the purified c-Myc tagged form of Thioredoxin.

positive to anti c-Myc antibody was clearly detected in lane 3 corresponding to the anti T7 immunoprecipitate. A pure c-Myc tagged RPL2 sample was loaded in lane 6 as control and it was positively stained by the antibody whereas pure T7-RNAPα was not recognised as expected (lane 5).
The presence of T7-RNAPα in the immunoprecipitate was verified by developing the Western blot with anti-T7 antibodies (Figure 3B). A protein band of about 40 KDa was clearly detected by the antibody in lane 3, that perfectly corresponded to pure T7-RNAPα loaded in lane 5 as control. As expected, pure c-Myc RPL2 was not recognised by the anti T7 antibody (lane 6) thus ruling out any cross reactivity in the experiment. No signal was generated in the fraction eluted from agarose beads alone (lanes 4, Figure 3A-B).
A co-immunoprecipitation experiment using T7-tagged RNAPα and an unrelated protein, a c-Myc labelled Thioredoxin was performed as control to rule out any possible artefact in the co-IP between RNAPα and RPL2 and to confirm the physiological relevance of this interaction. Thioredoxin is a cytosolic oxidoreductase involved in the reduction of other proteins by cysteine thiol-disulfide exchange and then it is completely unrelated with the transcriptional mechanism. The recombinant proteins were produced in *E. coli* C41 cells using the same experimental conditions described above and the total protein extract was immunoprecipitated by anti-T7 agarose-conjugated antibody and underivatised agarose beads as control. As shown in Figure 3D, when the Western blot was developed with anti-T7 antibodies, a protein band corresponding to pure T7-RNAPα loaded in lane 5 as control was clearly visible in the lane 3 corresponding to fraction eluted from anti-T7 antibody beads. On the contrary, Figure 3C shows that, when the anti-c-Myc antibody was used, a protein band corresponding to pure thioredoxin (loaded in lane 5 as control) was only detected in lane 1 containing the unbound material. No trace of this band could be observed in the IP lane (Fig. 3C, lane 3). These results demonstrated that thioredoxin did not interact with RNAPα, indicating that, at least in the conditions used, no aspecific aggregation of the expressed proteins had occurred and ruling out any possible artefact in the co-IP experiments. A biologically significative interaction seems to take place between RNAPα and RPL2 confirming previous proteomic data.

Investigation of RNAPα-RPL2 interaction.

A further question to be addressed was whether RNAPα and RPL2 proteins gave origin to a direct interaction or they just belong to the same multi-protein complex. To address this point, we designed a biochemical experiment based on the bacterial two hybrid system (5).
The genes coding for RNAPα and RPL2 were cloned in the two plasmid vectors pcI_{P22} and pcI_{434} which contained the N-terminal domain of phage 434 and P22 repressor, respectively. The recombinant plasmids obtained were used to transform the *E. coli* strain R721 that harbours the 434-P22 chimeric operator located upstream the *lacZ* gene. If RNAPα and RPL2 had given origin to a direct interaction, a functional repressor would be produced and the β-galactosidase activity expressed by *lacZ* gene would be decreased.
The ability of the reconstituted repressor to bind to the chimeric operator and to affect enzymatic

Table 3. Two-hybrid assay.
(a) The pcI plasmids were inserted into *E. coli* strain R721 by transformation and the ability of the reconstituted repressors to bind to the chimeric operator was then tested by measuring the residual β-galactosidase activity. (b) Two-hybrid assay to investigate the interaction between RNAPα and RPL2. The plasmids pcI_{P22}-RNAPα and pcI_{434}-L2 were both expressed in *E. coli* R721 cells and the residual β-galactosidase activity was recorded. In the absence of any pcI plasmid, strain R721 produced 2500 Miller units of β-galactosidase activity (mean of 5 independent experiments). Other values are the mean of 9 independent experiments.

pcI plasmids	β-galactosidase (Miller units)	proportion of residual β-galactosidase activity (%)
(a)		
pcI_{434}	2350±2.5	95%
pcI_{P22}	2517±3.4	100%
pcI_{P22}+pcI_{434}	2373±1.5	94.1%
pcI_{434}-434	1886±2.6	75%
pcI_{P22}-434	2113±3.1	86%
pcI_{434}-434+pcI_{P22}-434	351±2.7	14%
(b)		
pcI_{P22}-RNAPα	1915±1.8	76.2%
pcI_{434}-RPL_2	1829±2.5	72.7%
pcI_{434}-RNAPα	2200±3.4	88%
pcI_{P22}-RNAPα+pcI_{434}-RPL_2	880±2.9	35%
pcI_{P22}-RNAPα+pcI_{434}-RNAPα	401±3.2	16%

activity was tested by measuring the residual β-galactosidase activity following induction of the bacterial cells with 0.1 mM IPTG. The results are reported in Table 3. When the plasmids pcI_{P22}-RNAPα and pcI_{434}-RPL2 were both expressed in *E. coli* R721 cells, a considerable decrease in the β-galactosidase activity was recorded with the residual enzymatic activity dropping down to 35%.
Since RNAPα is known to form a functional dimer, the plasmids pcI_{P22}-RNAPα and pcI_{434}-RNAPα were also co-expressed in R721 cells as control. As expected, when RNAPα gave origin to the dimeric structure, a functional repressor was produced and the β-galactosidase activity decreased to 16% of the initial value.
The results of the bacterial two hybrid assay demonstrated that a direct interaction occurs between RNAPα and RPL2 within the multi-protein complex.

RPL2 is a transcriptional modulator.
Finally we addressed the point whether the binding of RPL2 to RNAPα might have a functional role in modulating the transcription mechanism in *E. coli*. Since RPL2 is essential for growth and survival of *E. coli*, we could not design a knock-out gene experiment. The putative transcriptional role of the interaction between RNAPα and RPL2 was then investigated by *in vivo* transcription experiments using a classic gene reporter assay.
The *lacZ* gene coding for *E. coli* β-galactosidase was cloned into the pET22b(+) expression vector, under the control of the ribosomal promoter *rrnD* P1, originating the recombinant expression vector pET22b-P*rrnD-lacZ*. The *rrnD* promoter P1 used contains the DNA sequence extending from -61 to +1 including the core promoter plus

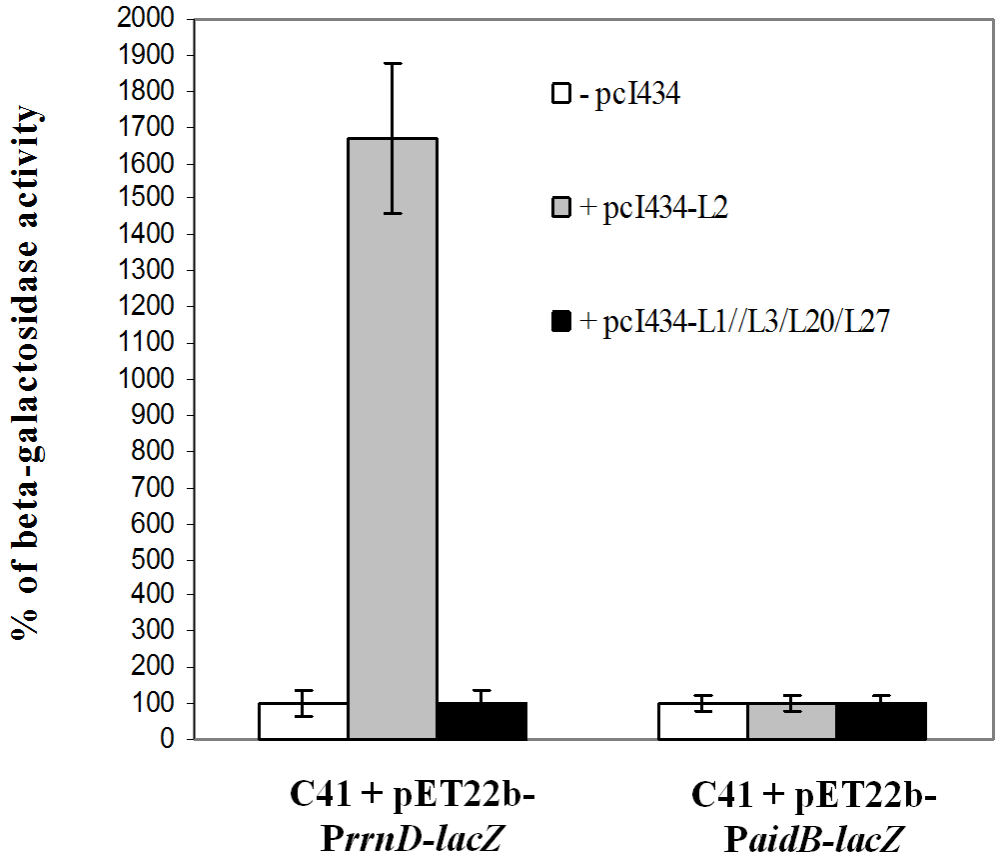


Figure 4. *In vivo* transcription assays.
The *E. coli* C41 cells were transformed with pET22b-P*rrnD-lacZ* or with pET22b-P*aidB-lacZ* vector and the β-galactosidase specific activity was monitored in the absence and in the presence of the recombinant proteins L1, L2, L3, L20 and L27. The β-galactosidase activity was evaluated at log phase. White bars represent the β-galactosidase activity in the absence of recombinant production; grey bars represent the β-galactosidase activity when recombinant L2 is overproduced; black bars represent the β-galactosidase activity measured in the presence of L1/L3/L20 or L27. Means and standard deviations have been calculated from three independent assays.

the upstream element (Table 2). The gene coding for RPL2 was cloned into the pcI_{434} plasmid, generating the recombinant expression vector pcI_{434}-L2. The two compatible expression vectors were then used to co-transform *E. coli* C41 cells. The same cells were transformed with pET22b-P*rrnD-lacZ* vector alone as control. *E. coli* cells were allowed to grow and the β-galactosidase activity was monitored in both cell cultures. Figure 4 shows the results obtained. A large increase in the β-galactosidase activity could clearly be observed in the *E. coli* cells producing recombinant RPL2 as compared to the control. These finding suggested that the ribosomal protein might act as a transcriptional activator of the ribosomal operon *rrnD.*

The specificity of RPL2 toward the *rrnD* promoter P1 was tested by an *in vivo* transcription assay using β-galactosidase as reporter gene under the control of an unrelated promoter, the *aidB* promoter. The P*aidB* used contains the DNA sequence extending from -61 to +1 including the core promoter plus the upstream element (Table 2). As shown in Fig 4, recombinant RPL2 did not affect β-galactosidase expression when the

aidB promoter was used in the assay, as the enzymatic activity in cell cultures transformed with the constructs pET22b-P*aidB-lacZ* and pcI_{434}-L2 remains unchanged. These data confirmed previous results, that RPL2 is a transcriptional regulator specific for the ribosomal promoter *rrnD* P1.

In addition, to ensure that the effect on *rrnD* P1 could be specifically attributed to RPL2, the putative role of other ribosomal proteins on rRNA transcription was tested. As appropriate controls, four proteins of the bacterial ribosome were chosen: L1 and L3 that are relatively large ribosomal proteins such as L2; L20, a ribosomal protein that participates in the early assembly steps of the 50 S ribosomal subunit (18) and L27 that is a relatively late assembly protein (22). As clearly shown in Fig. 4, L1, L3, L20 and L27 were unable to affect *lacZ* expression driven by both *rrnD* P1 and P*aidB*, thus demonstrating the specific and direct role of RPL2 on rRNA expression.

Besides their role in ribosome assembling, several ribosome proteins have been shown to fulfil other biological functions in the cell. One of the best characterised functions is the ability of some *E. coli* ribosomal proteins to regulate the translation of their own multicistronic mRNAs (23). Other *E. coli* ribosomal proteins have been demonstrated to play roles in transcriptional antitermination (S4 and S10) or to be involved in DNA repair and replication (S9 and L14) (21). Data reported in this paper showed that the highly conserved RPL2 protein is involved in transcription regulation by acting as an activator of the ribosomal operon transcription. Moreover, the functional proteomic approach indicated that RPL2 belongs to a large multi-protein transcription complex that gathers at the RNAPα subunit. Finally, we demonstrated that the transcriptional activator function exerted by RPL2 is carried out through the direct and specific binding of the ribosomal protein to the α subunit of the RNA polymerase.

ACKNOWLEDGEMENTS

We are indebted to Prof. Paolozzi and Prof. Di Lallo from the University Tor Vergata di Roma for the bacterial two-hybrid system. This work was supported by grants of Ministero dell'Università e della Ricerca Scientifica (Progetti di Rilevante Interesse Nazionale 2004, 2005, 2006; FIRB 2001), of Programma Nazionale di Ricerche in Antartide 2004, and of Regione Campania L.R. 05/03. Support from the National Center of Excellence in Molecular Medicine (MIUR - Rome) and from the Regional Center of Competence (CRdC ATIBB, Regione Campania – Naples), Istituto Nazionale Biomolecole e Biosistemi (INBB) is gratefully acknowledged.

References

1. Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**:248-254.
2. Browning, D. F., and S. J. Busby. (2004). The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol*. **2**:57-65.
3. Busby, S., and R. H. Ebright. (1999). Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol*. **293:**199–213.
4. Chen, J., C. Röcken, C. Lofton-Day, H.U. Schulz, O. Müller, N. Kutzner, P. Malfertheiner, and M.P. Ebert. (2005). Molecular analysis of APC promoter methylation and protein expression in colorectal cancer metastasis. *Carcinogenesis.* **26**:37-43.
5. Di Lallo, G., L. Castagnoli, P. Gherardini, and L. Paolozzi.

(2001). A two-hybrid system based on chimeric operator recognition for studying protein homo/heterodimerization in *Escherichia coli*. *Microbiology* **147**:1651-1656.

6. Ebright, R. H. (1993). Transcription activation at class I CAP-dependent promoters. *Mol. Microbiol.* **8**:797-802.
7. Fraldi, A., E. Zito, F. Annunziata, A. Lombardi, M. Cozzolino, M. Monti, C. Spampanato, A. Ballabio, P. Pucci, R. Sitia, and M. P. Cosma. (2008). Multistep, sequential control of the trafficking and function of the Multiple Sulfatase Deficiency gene product, SUMF1 by PDI, ERGIC-53 and ERp44. *Hum. Mol. Genet.* **17**:2610-2621.
8. Gavin, A.C., M. Bösche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J. M. Rick, A. M. Michon, C. M. Cruciat, M. Remor, C. Höfert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M. A. Heurtier, R. R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, and G. Superti-Furga. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**:141-147.
9. Gingras, A.C., M. Gstaiger, B. Raught, and R. Aebersold. (2007). Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* **8**:645-654.
10. Gourse, R.L., W. Ross, and T. Gaal**.** (2000). Ups and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Mol. Microbiol.* **37:**687–695.
11. Ho, Y., A. Gruhler, A. Heilbut, G. D. Bader, L. Moore, S. L. Adams, A. Millar, P. Taylor P, K. Bennett, K. Boutilier, L. Yang, C. Wolting, I. Donaldson, S. Schandorff, J. Shewnarane, M. Vo, J. Taggart, M. Goudreault, B. Muskat, C. Alfarano, D. Dewar, Z. Lin, K. Michalickova, A. R. Willems, H. Sassi, P. A. Nielsen, K. J. Rasmussen, J. R. Andersen, L. E. Johansen, L. H. Hansen, H. Jespersen, A. Podtelejnikov, E. Nielsen, J. Crawford, V. Poulsen, B. D. Sørensen, J. Matthiesen, R. C. Hendrickson RC, F. Gleeson, T. Pawson, M. F. Moran, D. Durocher, M. Mann, C. W. Hogue, D. Figeys, and M. Tyers. (2002). Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**:180-183.
12. Köcher, T., and G. Superti-Furga. (2007). Mass spectrometry-based functional proteomics: from molecular machines to protein networks. *Nat. Methods.* **4**:807-815.
13. Lewis, T.S., J. B. Hunt, L. D. Aveline, K. R. Jonscher, D. F. Louie, J. M. Yeh, T. S. Nahreini, K. A. Resing, and N. G. Ahn. (2000). Identification of novel MAP kinase pathway signaling targets by functional proteomics and mass spectrometry. *Mol. Cell.* **6**:1343-1354.
14. Miller, J.H. (1972). Experiments in molecular genetics. *Cold Spring Harbor Laboratory Press*, Cold Spring Harbor, N.Y.
15. Miroux, B., and J. E. Walker. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**:289-298.
16. Monti, G., L. De Napoli, P. Mainolfi, R. Barone, M. Guida, G. Marino, and A. Amoresano. (2005)**.** Monitoring Food quality by microfluidic electrophoresis, gas chromatography, and mass spectrometry techniques: effects of aquaculture on the sea bass

(*Dicentrarchus labrax*). *Anal. Chem.* **77:**2587–2594.
17. Monti, M., S. Orrù, D. Pagnozzi, and P. Pucci. (2005). Interaction proteomics. *Biosci. Rep*. **25**:45-56.
18. Raibaud, S., I. Lebars, M. Guillier, C. Chiaruttini, F. Bontems, A. Rak, M. Garber, F. Allemand, M. Springer and F. Dardel. (2002)**.** NMR structure of bacterial ribosomal protein L20: implications for ribosome assembly and translational control. *J. Mol. Biol.* **323**:143-151.
19. Rojo, F. (1999) Repression of transcription in bacteria. *J. Bacteriol.* 181:2987-2991
20. Savery, N., V. Rhodius, and S. Busby. (1996). Protein-protein interactions during transcription activation: the case of *Escherichia coli* cyclic AMP receptor protein. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **351**:543-550.
21. Torres, M., C. Condon, J. M. Balada, C. Squires, and C. L. Squires**.** (2001). Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. *EMBO J.* **14**:3811-3820.
22. Wower, I.K., J. Wower, and R.A. Zimmermann. (1998). Ribosomal protein L27 participates in both 50 S subunit assembly and the peptidyl transferase reaction. *J. Biol. Chem*. **273**: 19847-19852.
23. Zengel, J.M., and L. Lindahl**.** (1994). Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli*. *Prog. Nucleic Acid Res. Mol. Biol.* **47**:331-370.

Available online at www.sciencedirect.com
ScienceDirect
Journal of Biotechnology 127 (2007) 199–210
Journal of BIOTECHNOLOGY
www.elsevier.com/locate/jbiotec

An effective cold inducible expression system developed in *Pseudoalteromonas haloplanktis* TAC125

Rosanna Papa [a], Valentina Rippa [a], Giovanni Sannia [a], Gennaro Marino [a,b], Angela Duilio [a,*]

[a] *Department of Organic Chemistry and Biochemistry, Federico II University of Naples, Napoli, Italy*
[b] *School of Biotechnological Sciences, Federico II University of Naples, Napoli, Italy*

Received 30 November 2005; received in revised form 18 May 2006; accepted 2 July 2006

Abstract

A regulative two-component system previously identified in *Pseudoalteromonas haloplanktis* TAC125 was used to construct a cold inducible expression system that is under the control of L-malate. Performances of the inducible system were tested for both psychrophilic and mesophilic protein production using two "difficult" proteins as control. The results obtained demonstrated that both psychrophilic β-galactosidase and yeast α-glucosidase are produced in a fully soluble and catalytically competent form. Optimal conditions for protein production, including growth temperature, growth medium and L-malate concentration were also investigated. Under optimized conditions yields of 620 and 27 mg/l were obtained for β-galactosidase and α-glucosidase, respectively.

Keywords: Psychrophiles; Protein production; Cold inducible system; β-Galactosidase; α-Glucosidase

1. Introduction

Incorrect folding of the nascent polypeptide chains is one of the main problems occurring during heterologous protein production in bacteria. A key role in this process is played by intermolecular hydrophobic interactions taking place among partly folded intermediates that cause protein molecules to stick together thus driving them away from the productive folding pathway (Georgiou and Valax, 1996). Since formation of inclusion bodies often impairs the recombinant production of valuable proteins, many experimental approaches have been explored to minimize this undesirable effect, including expression of chimerical proteins (Mitra et al., 2005) and co-expression with chaperonines (Luo and Hua, 1998). Expression of "difficult" proteins has also been carried out by lowering the temperature at the physiological limit allowed for the growth of

* Corresponding author at: Dipartimento di Chimica Organica e Biochimica, Complesso Universitario di Monte Sant'Angelo, Via Cinthia, 80126 Napoli, Italy. Tel.: +39 081674314; fax: +39 081674313.
E-mail address: anduilio@unina.it (A. Duilio).

0168-1656/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2006.07.003

mesophilic host organisms (between 15 and 18 °C for *Escherichia coli*). Lowering the temperature, in fact, has a pleiotropic effect on the folding process, destabilising the hydrophobic interactions needed for intermediates aggregation (Jeon et al., 1995). Although in some cases this approach has been reported to increase yields of soluble and active recombinant protein products, the exploitation of an industrial process performed at suboptimal growth condition of the expression host might hardly be considered.

A rational alternative to mesophilic organisms is the use of naturally cold-adapted bacteria as hosts for protein production at low temperature (even at around 0 °C). The development of a shuttle genetic system for the transformation of the cold adapted Gram-negative bacterium *Pseudoalteromonas haloplanktis* TAC125 (P*h*TAC125) (Birolo et al., 2000; Tutino et al., 2001) has already been reported. This system has made it possible the isolation of constitutive psychrophilic promoters and the construction of cold expression systems for the homologous/heterologous protein production at low temperatures (Duilio et al., 2004a,b). The described expression system represented the first example of heterologous protein production based on a true cold-adapted replicon (Duilio et al., 2003). However, the development of an effective cold expression system needs to be finely tuned possibly using *ad hoc* promoters. Physical separation between bacterial growth phase and expression of the desired proteins, in fact, cannot only improve the productivity of the entire system but can also play an important role in the production of proteins toxic for the host cells. These goals can only be achieved by using regulated promoters and efficient induction strategies. Recently, using a proteomic approach and taking advantage from the genome sequence of *Ph*TAC125 (Medigue et al., 2005) we isolated and characterized a functionally active two-component system. The regulatory system (PSHAb0361–PSHAb0362) is involved in the transcriptional regulation of the gene coding for an outer membrane porin (PSHAb0363), and it is strongly induced by the presence of L-malate in the medium (Papa et al., in press).

In this paper we used the regulative region comprising the two-component system located upstream the PSHAb0363 gene to construct an inducible expression vector, named pUCRP that is under the control of L-malate. Performances of the inducible system were tested for both psychrophilic and mesophilic protein production using two "difficult" proteins as model systems. Moreover, an evaluation of optimal induction conditions for protein production was also carried out. Data presented in this paper demonstrated that both psychrophilic β-galactosidase from *Ph*TAE19 (Hoyoux et al., 2001) and mesophilic α-glucosidase from *Saccharomyces cerevisiae* (Kopetzki et al., 1989a,b) are produced in *Ph*TAC125 in good yields and in a completely soluble and catalytically competent form.

2. Materials and methods

2.1. Bacterial strains, DNA constructs and media

Pseudoalteromonas haloplanktis TAC125 (*Ph*TAC125) (Birolo et al., 2000), a Gram-negative bacterium was grown in aerobic conditions at 4 and 15 °C in minimal medium (1 g/l KH_2PO_4, 1 g/l NH_4NO_3, 10 g/l NaCl, 0.2 g/l $MgSO_4 \times 7H_2O$, 10 mg/l $FeSO_4$, 10 mg/l $CaCl_2 \times 2H_2O$, supplemented with casamino acid 0.5% w/v and different concentration of L-malate when requested) and in TYP broth (16 g/l yeast extract, 16 g/l bacto tryptone, 10 g/l marine mix) at pH 7.5, supplemented with 100 μg/ml ampicillin if transformed. The plasmids were mobilized into *P. haloplanktis* TAC125 by interspecific conjugation using the transformed *E. coli* S17-1 (λ_{pir}) cells as donor (Tascon et al., 1993). Aliquots (100 μl) of logarithmic cultures of the donor and the recipient strains were mixed and spotted as a drop onto TYP plate. After 16 h of mating at 15 °C, the cells were resuspended in 200 μl of TYP medium. Psychrophilic transconjugants were selected by plating serial dilutions at 4 °C on TYP plates containing 100 μg/ml ampicillin (Tutino et al., 2001).

2.2. Construction of the expression vectors

Plasmids are all reported in Table 1.

Genomic DNA preparation from *Ph*TAC125 was carried out as previously described (Tosco et al., 2003).

The pUCRP inducible cold-adapted gene-expression vector was constructed by cloning DNA fragment corresponding to the cold-active promoter of P(PSHAb0363) vector (Papa et al., in press; http://

Table 1
Plasmids and oligonucleotides

Plasmid	Description	References
P(PSHAb0363)	pPLB containing the PSHAb0363 promoter region (*Ph*TAC125 genome chromosome II from 423068 to 419631 bp)	Papa et al. (in press)
pPLB	Promoter-trap cold-adapted vector containing the promoter-less *Ph*TAE79 *lacZ* gene	Duilio et al. (2004a,b)
pUCLT/Rterm	Vector deriving from the pUC18 plasmid, containing the T/R box and the transcription termination signal from *Ph*TAC125 *aspC* gene	Tutino et al. (2002)
pUCRP	pUCLT/Rterm containing PSHAb0363 promoter region	This work
pPM4	pUCLT/Rterm containing P4 *Ph*TAC125 promoter region (422 bp, GenBank accession number AJ557253)	Duilio et al. (2004a,b), Siani et al. (in press)
pPM4GLUCPl	Expression vector containing *Sc*GLUCPl gene	Duilio et al. (2003)
pKK-177GLUCPl	Recombinant vector for the cytoplasmic, IPTG-inducible production of *Sc*GLUCPl in *Escherichia coli*	Kopetzki et al. (1989a)
pUCRPGLUCPl	pUCRP containing *Sc*GLUCPl gene	This work
Oligonucleotide		
PSHAb0363a Fw	5′-CAAAGCTAGGCAAAGCTTAATTATAC-3′	
PSHAb0363a Rv	5′-CCTGGATCCAATATCGATAGTTTTACG-3′	
PSHAb0363b Fw	5′-GATGGACGTCTAGAACTATCGATATTAG-3′	
PSHAb0363b Rv	5′-CCTTCAATCTAGATATCTGCAGGAGTATC-3′	
PSHAb0363c Fw	5′-GATACTCCTGCAGATATATTAATTG-3′	
PSHAb0363c Rv	5′-GTTTAACGTGTCTGCAGTTTATCATATGGTGTCC-3′	
AG1	5′-CAAGCTTACACAGGAAGAGAAATCATATGACTATTTC-3′	
AG2	5′-GGGGAATCTGGTAAACCAGGACGTTTCGAATAC-3′	

www.bioinfo.hku.hk/GenoList/index.pl?database=psychrolist) into pUCLT/Rterm vector (Fig. 1, Table 1).

P(PSHAb0363) promoter region was inserted into pUCLT/Rterm vector by a three steps procedure as described in Fig. 1. The first step consisted in the double *Hind*III/*Cla*I digestion of a 1700 bp fragment (PCR amplified by using the primers PSHAb0363a Fw and PSHAb0363a Rv and the *Ph*TAC125 DNA genomic as template). In the second step a 750 bp fragment was amplified using PSHAb0363bFw and PSHAb0363b Rv as primers and double *Cla*I/P*st*I digested. Finally, the third fragment containing the 3′ region of P(PSHAb0363) promoter region was amplified using PSHAb0363c Fw and PSHAb0363c Rv as primers and *Pst*I digested. The amplified fragments were cloned into pUCLT/Rterm corresponding sites by using two consecutive ligation reactions, and their nucleotide sequence checked to rule out the occurrence of any mutation during synthesis. The resulting vector was called pUCRP (Table 1).

The amplifications were performed using in a mixture containing 60 ng of template, 50 pmol of each oligonucleotide primer, 1.8 mM, $MgCl_2$, 50 mM KCl, 20 mM Tris–HCl pH 8.3, 0.1% gelatine, 200 μM dNTP in a final volume of 50 μl. The mixtures were incubate at 95 °C for 10 min, after which 1.25 U of Taq DNA polymerase were added. Twenty cycles of amplification (consisting of 1 min at 95 °C, 1.5 min at 58 °C and 1 min at 72 °C) were carried out and were followed by a cycle in which the extension reaction at 72 °C was prolonged for 10 min in order to complete DNA synthesis.

2.3. Construction of expression vectors for the production of Saccharomyces cerevisiae α-glucosidase in PhTAC125

The *S. cerevisiae MAL12* gene (*ScMAL12*, GenBank accession no. E.C.3.2.1.20), encoding the α-glucosidase (Kopetzki et al., 1989a) was inserted into the pUCRP expression vector by a single step procedure, consisting in the double *Nde*I/*Bam*HI digestion using pPM4GLUCPl plasmid as template (Duilio et al., 2003). The resulting vector was called pUCRPGLUCPl (Table 1).

pPM4GLUCPl expression vector was generated by a two-step procedure, the first one consisting in the

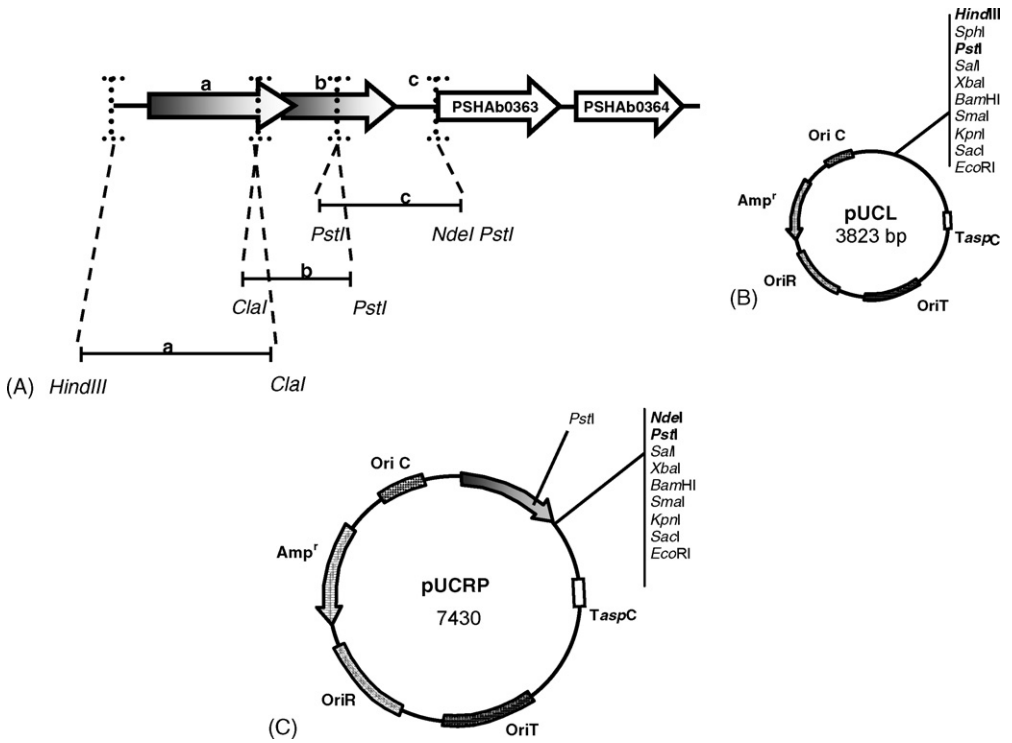


Fig. 1. Construction of pUCRP expression vector. (A) *Ph*TAC125 DNA genomic fragment corresponding to the cold-active promoter of P(PSHAb0363) vector (Papa et al., in press). (B) pUCLT/Rterm promoter-less vector. (C) pUCRP expression vector.

double *Nde*I/*Pst*I digestion of a 580 bp fragment (PCR amplified by using the AG1 and AG2 primers and the expression vector pKK-177GLUCPl as template). In the second step, a 1180 bp fragment, containing the 3′ region of *MAL12* gene, was recovered from pKK-177GLUCPl (Kopetzki et al., 1989a) by a double *Pst*I/*Bam*HI digestion. Finally, the two fragments were ligated into the pPM4 vector (Siani et al., in press), previously digested with *Nde*I and *Bam*HI, generating the plasmid pPM4GLUCPl (Table 1). A DNA sequencing reaction was performed on the *ScMAL12* amplified region to rule out the occurrence of mutations introduced by the PCR synthesis.

2.4. Enzymatic assays and analytical procedures

Duplicate cultures of recombinant *Ph*TAC125 cells were aerobically grown in liquid culture and cellular pellets (corresponding to 25 $OD_{600\,nm}$) were collected during the growth. The cells were resuspended in 1 ml of lysis buffer (100 mM sodium phosphate, 5 mM EDTA, 2% Triton X-100, 1 mM PMSF, 1 mM DTT, and 5 mg/ml lisozyme, pH 7.8), and incubated for 20 min at 15 °C. After a centrifugation step of 20 min at 4 °C 9500 × *g*, the supernatant was collected.

Protein concentration was determined with the Bio-Rad protein assay (Bradford, 1976), using bovine serum albumine as standard.

Recombinant cold-active β-galactosidase was assayed spectrophotometrically at 25 °C as previously reported (Hoyoux et al., 2001), and calculation were performed on the basis of an extinction coefficient for *o*-nitrophenol of 3.5 $mM^{-1}\,cm^{-1}$ at 410 nm and specific activity of 138.2 U/mg purified enzyme (Hoyoux et al., 2001). Catalytic parameters were determined as reported in Hoyoux et al., 2001.

α-Glucosidase production was assayed using 5 mM PNPG (*p*-nitrophenyl-α-D-glucopyranoside) as chromogenic substrate at 25 °C in 100 mM potassium phosphate, 1 mM $MgCl_2$, 100 mM 2-mercaptoethanol, pH 6.8, and calculated on the basis of the *p*-nitrophenol $\varepsilon_{410\,nm} = 1.8\,mM^{-1}\,cm^{-1}$ and specific activity of

130 U/mg purified enzyme (Kopetzki et al., 1989a). Catalytic parameters were determined as reported in Kopetzki et al. (1989a,b).

2.5. In situ digestion

The protein bands stained by Colloidal Comassie (Pierce) and containing the proteins to be identified were excised, cut in small pieces and washed in 50 mM ammonium bicarbonate pH 8.0 in 50% acetonitrile to a complete destaining. The gel pieces were re-suspended in 50 mM ammonium bicarbonate pH 8.0, reduced with 10 mM DTT at 56 °C for 45 min and alkylated with a 55 mM solution of iodoacetamide in the same buffer for 30 min at room temperature in the dark. The excess of reagent was descarded and the gel pieces were washed several times with the buffer, resuspended in 50 mM ammonium bicarbonate and incubated with 100 ng of trypsin for 2 h at 4 °C and overnight at 37 °C. The supernatant containing peptides was removed and the remaining gel pieces were washed with acetonitrile in order to extract the peptides still present in the gel. These two fractions were then collected and freeze-dried.

2.6. MALDI MS analysis and protein identification

MALDI mass spectra were recorded on an Applied Biosystem Voyager DE-PRO mass spectrometer equipped with a reflectron analyser and used in delayed extraction mode. A 1 μl of peptide sample was mixed with an equal volume of α-cyano-4-hydroxycynnamic acid as matrix (10 mg/ml in 0.2% TFA in 70% acetonitrile), applied to the metallic sample plate and air dried. Mass calibration was performed by using the standard mixture provided by manufacturer. Raw data, reported as monoisotopic masses, were then introduced into MASCOT peptide fingerprinting search program (by Matrix Science, Boston, USA) available on the net and used for protein identification.

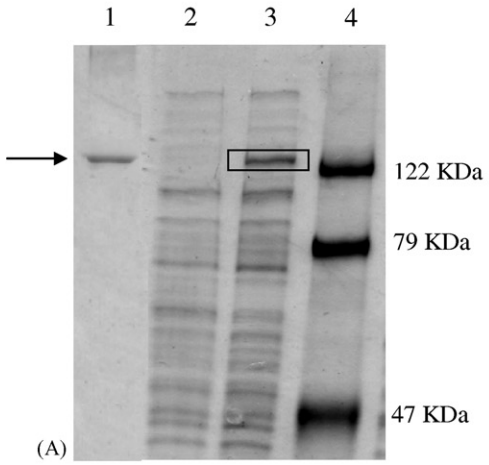
3. Results and discussion

The isolation and the structural and functional characterization of a L-malate inducible promoter from *Ph*TAC125 has already been reported (Papa et al., in press). This promoter, named P(PSHAb0363), is able to direct the expression of a promoter-less psychrophilic *lacZ* gene, exhibiting a 13-fold increase in the β-galactosidase activity when L-malate was added to the medium (Papa et al., in press). Performances of this inducible system were tested both for psychrophilic and mesophilic protein production using the psychrophilic β-galactosidase from *P. haloplanktis* TAE79 and the *S. cerevisiae* α-glucosidase. These proteins were chosen because they can hardly be expressed in recombinant form in mesophilic hosts even at sub-optimal temperature conditions. When the β-galactosidase was produced in *E. coli* cells at 18 °C, 20 mg of catalytically active enzyme was produced per liter of culture (Feller, personal communication). Analogously, recombinant yeast α-glucosidase produced in *E. coli* aggregates in insoluble form, the active soluble amount of protein being less than 1% of the total production (Le Thanh and Hoffmann, 2005).

3.1. Psychrophilic protein production in *PhTAC125*

The cold-active promoter P(PSHAb0363) was used to construct an inducible cold-adapted gene-expression vector named pUCRP containing the gene coding for the β-galactosidase from *P. haloplanktis* TAE79. *Ph*TAC125 cells harbouring β-galactosidase were grown in liquid culture in minimal medium in the presence and in the absence of L-malate at 15 °C up to the stationary phase. An aliquot of each culture was collected and the soluble protein content was analyzed by SDS gel electrophoresis. Fig. 2A shows the Coomassie blue-stained gels corresponding to the protein extracts from *Ph*TAC125 recombinant cells grown in the absence and in the presence of L-malate. The presence of a strong 118 kDa extra-band could clearly be detected in the cellular extract grown in the presence of the inducer (panel A, lane 3). In the absence of L-malate, only a tiny band with the same electrophoretic mobility was observed (panel A, lane 2). These data suggest that under the control of the inducible expression system β-galactosidase is produced in *Ph*TAC125 only in the presence of L-malate, indicating an undetectable basal activity of the inducible P(PSHAb0363) promoter.

The unambiguous identification and a detailed structure characterization of the expressed β-galactosidase were obtained by mass spectrometric analysis of the



(B)

Mass	Peptide
926.53	439 - 445
962.57	41 - 48
970.60	26 - 34
1094.63	49 - 57
1126.67	611 - 620
1135.67	511 - 520
1222.74	48 - 57
1237.83	2 - 11
1299.76	300 - 309
1420.86	621 - 631
1427.86	299 - 309
1511.98	211 - 223
1521.88	425 - 438
1548.95	282 - 295
1563.93	126 - 138
1584.88	224 - 236
1669.99	12 - 25
1861.15	282 - 298
1889.16	167 - 183
1905.13	908 - 923
2020.22	278 - 295
2063.21	244 - 262
2182.37	333 - 351
2224.36	58 - 76
2298.26	106 - 125
2363.38	224 - 243
2510.39	799 - 821
2571.32	577 - 598
2598.45	310 - 332
3190.75	476 - 504

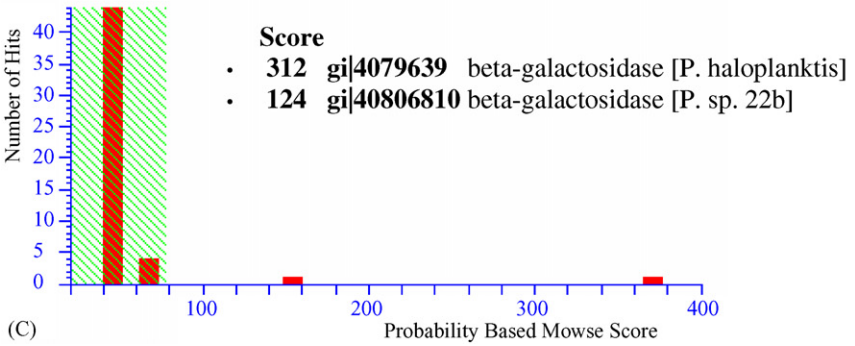


Fig. 2. Recombinant production of the thermally labile β-galactosidase from *Ph*TAE19 in *Ph*TAC125 cells. Panel A: 7.5% SDS PAGE gel electrophoresis of protein extracts from *Ph*TAC125 cells harboring P(PSHAb0363) and grown in minimal medium in the absence (lane 2) and in the presence (lane 3) of L-malate, in comparison with β-galactosidase from *Ph*TAE79 used as control (lane 1); molecular weight marker (lane 4). The recombinant protein is indicated by an open box. Panel B: MALDI-TOF analysis of the β-galactosidase peptides generated by tryptic digestion. Panel C: Mascot Search Results.

corresponding Coomassie blue stained protein band (Fig. 2A, lane 3) following *in situ* digestion with trypsin. The resulting peptide mixture was directly analysed by MALDI mass spectrometry and the accurate mass values recorded in the spectra were used to search for the entire NCBI protein databank using the MASCOT software. Fig. 2C shows the outcome of the search procedure that unambiguously identified the protein band as the *P. haloplanktis* β-galactosidase. Mass spectral data allowed us to verify a large portion of the β-galactosidase (Fig. 2B), ruling out possible modification and/or mutation of the protein product. This structural characterization demonstrated that the recombinant product was indeed the expected β-galactosidase from *Ph*TAE79.

To test the catalytic activity of recombinant β-galactosidase, *Ph*TAC125 recombinant cells were grown in liquid culture at 4 and 15 °C both in the absence and in the presence of L-malate up to the stationary phase. β-Galactosidase catalytic activity was assayed in the cell lysates as previously described by Hoyoux et al. (2001) using β-ONPG as substrate. The K_M for the synthetic substrate was determined and found to be almost identical to that obtained for the purified enzyme (Feller, personal communication; Table 2), thus indicating that the protein had assumed the correct native conformation.

Since wild type *Ph*TAC125 does not show any detectable β-galactosidase activity, the lysate from recombinant cells harbouring the void vector was used as negative control. The specific activity of the recombinant enzyme is reported in Table 3. Both at 4 and 15 °C enzymatic activity greatly increases in the presence of L-malate, being the largest increase detected at 15 °C. These data demonstrated that a soluble psychrophilic β-galactosidase is efficiently produced in *Ph*TAC125 in a catalytically competent form. The transcriptional activity of the inducible system was further investigated by monitoring β-galactosidase activity during the exponential and the late-stationary growth phase. These experiments demonstrated that β-galactosidase activity remained unchanged during cellular growth from the exponential to the late stationary phase (data not shown).

Table 2
Catalytic parameters of psychrophilic β-galactosidase and mesophilic α-glucosidase produced in *Ph*TAC125 in comparison with those determined for the purified enzymes

Proteins	K_M (mM)
β-Galactosidase produced in *Ph*TAC125	0.25 ± 0.04
Purified β-galactosidase (Feller, personal communication)	0.21
α-Glucosidase produced in *Ph*TAC125	0.18 ± 0.02
Purified α-glucosidase (Kopetzki et al., 1989a)	0.2

K_M value for β-galactosidase was determined according to the assay conditions reported in Hoyoux et al. (2001) using β-*o*NPG as substrate. K_M value for α-glucosidase was determined according to the assay conditions reported in Kopetzki et al. (1989a) using α-*p*NPG as substrate.

3.2. Induction conditions

Once the production of a recombinant biologically active β-galactosidase in *Ph*TAC125 was defined, different induction conditions were tested to optimize the productivity of the P(PSHAb0363) inducible system. This procedure consisted in two consecutive steps: (i) definition of the optimal cellular growth phase in which expression of the heterologous gene has to be induced (definition of t_0); (ii) definition of the optimal inducer concentration to obtain higher levels of expression of the psychrophilic gene (definition of C_0).

*Ph*TAC125 recombinant cells were grown in minimal medium in the absence of L-malate at 15 °C.

Table 3
β-Galactosidase and α-glucosidase specific activities (U/mg) from *Ph*TAC125 cells harbouring P(PSHA0363) and pUCRPGLUCPl at 4 and 15 °C grown in minimal medium (MM) in the presence and in the absence of L-malate, respectively

	Specific activity (U/mg) ± S.D.			
	β-Galactosidase		α-Glucosidase	
	4 °C	15 °C	4 °C	15 °C
MM	0.21 ± 0.05	0.86 ± 0.02	0.03 ± 0.00	0.03 ± 0.00
MM + L-malate	2.12 ± 0.05	11.42 ± 1.05	0.31 ± 0.03	0.75 ± 0.05

Data shown are the average of 12 measurements. S.D.: standard deviation.

Induction of psychrophilic gene expression was tested by adding L-malate 0.2% at four different phases during cellular growth, i.e. early ($t_0 = 5$ h; OD 600 nm = 0.32 ± 0.03) and mid (7 h; OD 600 nm = 0.53 ± 0.04) exponential phase, early (10 h; OD 600 nm = 1.31 ± 0.08) and mid ($t_0 = 23$ h; OD 600 nm = 1.92 ± 0.10) stationary phase. Recombinant β-galactosidase activity was monitored before the addition of the inducer and 2, 8 and 24 h after the induction (t_1). Fig. 3A shows the yield of β-galactosidase, expressed as mg of protein/l (Hoyoux et al., 2001), and the induction ratios (IR) obtained at the four different phases calculated as a ratio of β-galactosidase yield at a specific time after the induction (t_1) with respect to that before induction (t_0). The best induction ratios (t_1/t_0) were observed when L-malate was added during the exponential phase, with the highest ratio detected in the mid-exponential phase. High levels of recombinant enzyme can be already detected on cell lysates after 8 h from induction, but maximal β-galactosidase production was recorded at $t_1 = 24$ h from induction (716 mg/l).

The second step in the optimization procedure consisted in the definition of the optimal concentration of inductor (C_0). Four different concentrations of L-malate 0.2%, 0.4%, 1% and 5%, were added to recombinant *Ph*TAC125 cells, at mid-exponential phase. Cell lysates were collected at 2, 8 and 24 h from induction and β-galactosidase activity was assayed. Yield of β-galactosidase and calculated induction ratios are reported in Fig. 3B. Lower concentrations of inducer yielded the highest IR values with the maximum effect obtained using 0.4% L-malate. In these experiments, the results obtained at 0.2% L-malate, were slightly different from those reported in Fig. 3A possibly due to a slightly higher basal β-galactosidase activity, although the final enzyme production at 24 h was very similar (561 mg/l). Higher concentration of L-malate resulted in a strong reduction in the IR value, probably due to an effect of the dicarboxylic acid on the osmotic equilibrium of the cell; in particular when 5% L-malate was added to recombinant cells, a decrease of cellular growth was observed (data not shown). Also in this case, the highest production was obtained when cell lysates were collected 24 h after the induction (620 mg/l).

Finally, effect of the growth medium on production of the recombinant enzyme was investigated. *Ph*TAC125 recombinant cells were grown at 15 °C

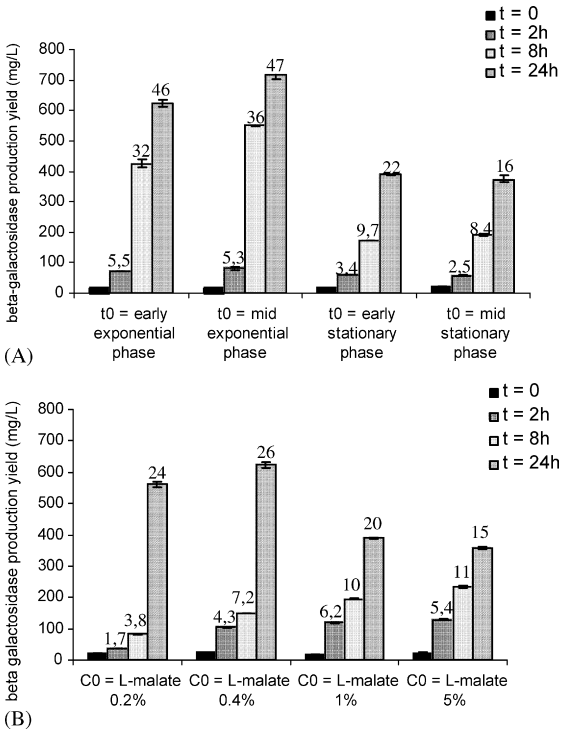


Fig. 3. Optimal conditions of induction for the β-galactosidase production. (A) Time of induction (t_0). β-Galactosidase production yield in *Ph*TAC125 cells harbouring P(PSHAb0363) grown in minimal medium (MM) obtained by adding L-malate 0.2% at different time (t_0) during cellular growth curve. The corresponding optical density before L-malate addiction were: early exponential phase $t_0 = 5$ h (OD 600 nm = 0.32 ± 0.03); mid exponential phase $t_0 = 7$ h (OD 600 nm = 0.53 ± 0.04); early stationary phase $t_0 = 10$ h (OD 600 nm = 1.31 ± 0.08); mid stationary phase $t_0 = 23$ h (OD 600 nm = 1.92 ± 0.10). (B) Concentration of inductor (C_0). β-Galactosidase production yield in *Ph*TAC125 cells harbouring P(PSHAb0363) grown in minimal medium (MM) obtained by adding L-malate at mid-exponential phase at four different concentrations (C_0). β-Galactosidase activity was monitored before the addition of inducer (as negative control of the experiment), and 2–8–24 h after the induction (t_1). Numbers above bars refer to induction ratios (t_1/t_0). Means and standard deviations have been calculated from four independent assays.

in both minimal (MM) and complex medium (Typ). According to the conditions previously determined, cells were induced with 0.4% L-malate at mid exponential phase and cell lysates were collected after 24 h from the induction. Table 4 summarises the results showing that recombinant β-galactosidase production in induced cells is always higher than that observed for cells not treated with L-malate. The inducer has a much

Table 4
β-Galactosidase and α-glucosidase production yields from *Ph*TAC125 cells harbouring P(PSHA0363) and pUCRPGLUCPl grown at 15 °C in minimal medium (MM) and in rich medium (Typ) in the presence and in the absence of L-malate, calculated accordingly to Hoyoux et al. (2001) and Kopetzki et al. (1989a), respectively

	Production yield (mg/l) ± S.D.	
	β-Galactosidase	α-Glucosidase
MM	12.8 ± 0.3	0.64 ± 0.00
MM + L-malate	620 ± 24	26.7 ± 1.8
TYP	5.6 ± 1.0	0.43 ± 0.04
TYP + L-malate	25.0 ± 0.9	9.22 ± 0.04

The data shown are the average of 12 measurements. S.D.: standard deviation.

higher effect (almost one order of magnitude) in minimal than in rich medium, showing a 48-fold increase in β-galactosidase production in MM compared to the 4.5-fold observed in Typ medium. Many reasons can account for this observation including competition for entry and interference with regulation networks due to the presence of many related metabolites in the rich medium.

3.3. Production of a mesophilic protein in PhTAC125

For the expression of the *S. cerevisiae* α-glucosidase, the *ScMAL12* gene encoding the yeast enzyme was inserted into the inducible expression vector pUCRP, as described in Section 2. The recombinant vector was mobilized into *Ph*TAC125 cells and the psychrophilic transconjugants were grown in minimal medium both in the absence and in the presence of L-malate at 15 °C up to the late exponential phase. Cell pellets were collected, disrupted and the soluble protein fractions were analysed by SDS gel electrophoresis. As shown in Fig. 4A, the presence of a protein band showing an electrophoretic mobility similar to that of commercial yeast α-glucosidase (lane 1) was clearly detectable (lane 3). This band is not present in the extract from recombinant *Ph*TAC125 cells grown in the absence of L-malate (lane 2), suggesting that a soluble α-glucosidase is produced by the psychrophilic inducible system only upon treatment with L-malate.

Cellular insoluble fractions were also analysed to detect any recombinant α-glucosidase aggregated in inclusion bodies. However, the SDS-PAGE analysis ruled out the occurrence of any protein band corresponding to the yeast enzyme (data not shown) thus demonstrating that the recombinant α-glucosidase is only produced in soluble form in the psychrophilic system. The unambiguous identification and a detailed structure characterization of the expressed α-glucosidase were obtained by mass spectrometric analysis as described above. The results are shown in Fig. 4B and C. Mass spectral data were used by the MASCOT software to search for the NCBI protein databank leading to a clear identification of the yeast enzyme (Fig. 4B). The peptide mass fingerprinting procedure showed that the identified peptides were scattered along the entire α-glucosidase sequence, ruling out the possible occurrence of genetic rearrangement. This structure characterization demonstrated that the recombinant product was indeed the expected α-glucosidase from *Saccharomyces cerevisiae*.

The biological activity of the recombinant enzyme was tested by enzymatic assay following the release of *p*-nitrophenol from *p*-nitrophenyl-α-D-glucopyranoside (PNPG) at 410 nm (Kopetzki et al., 1989a). As for β-galactosidase, the calculated K_M for the synthetic substrate was nearly identical to the reported by Kopetzki et al. (1989a) for the purified enzyme, indicating a correct folding process of the recombinant protein (Table 2).

*Ph*TAC125 recombinant cells were grown in liquid culture at 4 and 15 °C up to the mid exponential phase and then treated or not treated with 0.4% L-malate. Cellular lysates were collected 24 h after induction and the amount of recombinant α-glucosidase assessed by its catalytic activity as reported in Kopetzki et al. (1989a).

Table 3 reports the specific activity of recombinant α-glucosidase obtained. The highest α-glucosidase production was observed in lysates from induced cells grown at 15 °C (0.75 U/mg), demonstrating that the yeast α-glucosidase is efficiently expressed in the inducible psychrophilic host system in a totally soluble and catalytically competent form.

Effect of the growth medium in the production of the recombinant α-glucosidase was also investigated. *Ph*TAC125 recombinant cells were grown in duplicate in liquid culture at 15 °C in minimal and in complex medium; induction was performed with 0.4% L-malate at the mid exponential phase and cell lysates were collected 24 h after induction. Recombinant α-glucosidase was produced in induced *Ph*TAC125 cells in a totally

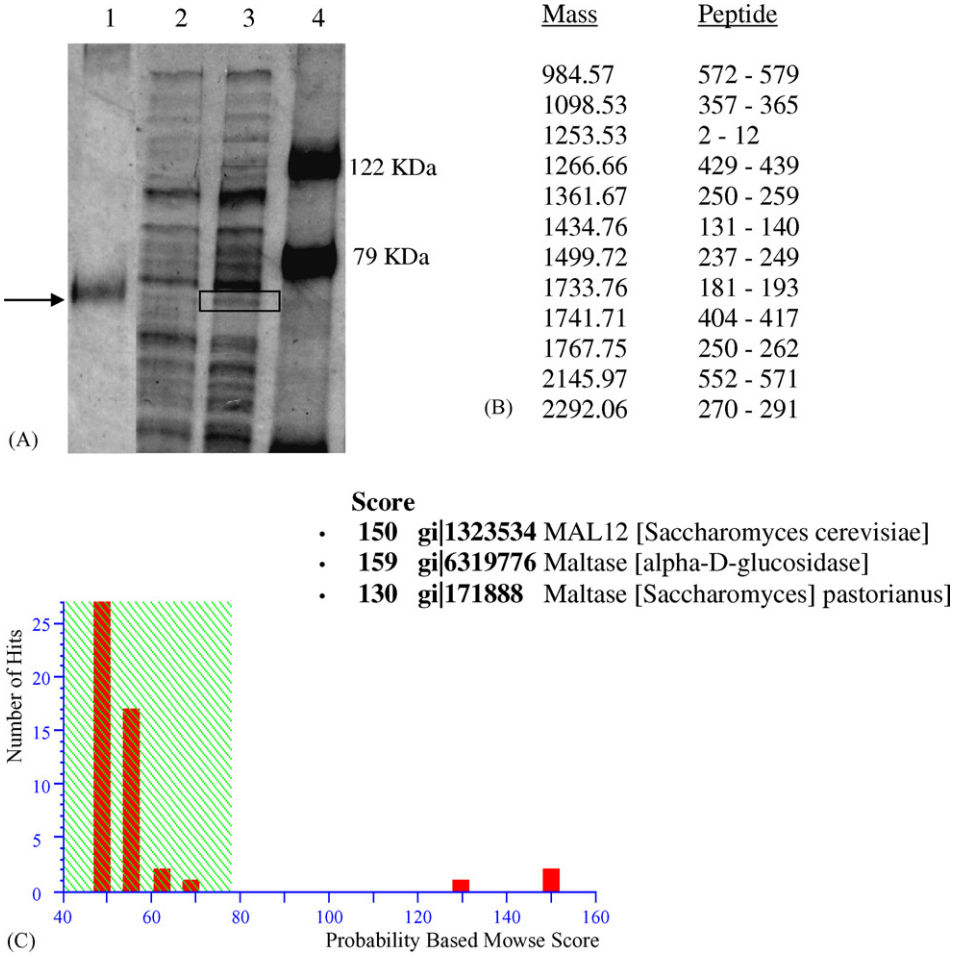


Fig. 4. Recombinant production of the mesophilic α-glucosidase from *Saccharomyces cerevisiae* in *Ph*TAC125 cells. Panel A: 7.5% SDS PAGE gel electrophoresis of protein extract from *Ph*TAC125 cells harboring pUCRPGLUCPl grown in minimal medium in the absence (lane 2) and in the presence (lane 3) of L-malate, in comparison with commercial α-glucosidase from yeast used as control (lane 1); molecular weight marker (lane 4). The recombinant protein is indicated by an open box. Panel B: MALDI-TOF analysis of the α-glucosidase peptides generated by tryptic digestion. Panel C: Mascot Search Results.

soluble and active form in both growth media. According to previous results, the effect of L-malate induction is observed in both conditions (Table 4) but is much higher in minimal than in rich medium (induction ratio 41 versus 21, total yield 26.7 mg/l).

4. Conclusions

An inducible cold expression system, which is effective in the production of both psychrophilic and mesophilic proteins, was developed in the Antarctic psychrophilic bacterium *Ph*TAC125. Performances of this inducible system were tested using two "arduous" proteins. Both recombinant psychrophilic β-galactosidase and yeast α-glucosidase were produced in *Ph*TAC125 as soluble and catalytically active enzymes. Structural and kinetic analyses of the recombinant proteins showed that both enzymes were nearly identical to their native counterparts. The absence of aggregated protein material might certainly be due to the lower expression temperature that destabilises hydrophobic interactions (Jeon et al., 1995). However, since the optimal expression temperature determined

for *Ph*TAC125 is only marginally lower than that used for proteins production in *E. coli* (15 °C compared to 18 °C), other factors must play an effect in preventing aggregation.

Experimental conditions for optimal protein production in the cold inducible expression system were also defined. Low concentrations of L-malate and long induction time are effective for maximal protein production. Higher L-malate concentration was shown either to be less effective or to impair cellular growth. The unusually long time of induction might possibly be ascribed to the long time of duplication displayed by *Ph*TAC125 when grown in minimal medium. The inducible system exhibits better performances in minimal medium when compared with rich medium.

At optimal expression conditions, recombinant β-galactosidase is produced in an high yields (620–720 mg/l), indicating that the inducible system can be very effective in the expression of psychrophilic proteins that are usually poorly produced in mesophilic hosts. A significant lower production yield is observed for yeast α-glucosidase putatively due to the different codon usage between the eukaryotic and bacterial organisms. Nevertheless, the cold expression system yielded a satisfactory amount of this protein in a soluble and active form.

Both recombinant proteins were structurally characterized in details by mass spectrometric analyses and their K_M values were determined showing to be almost coincident with those defined for the native enzymes.

The cold inducible expression system described in this paper has a number of interesting features, including the natural low temperature of cell growth and the possibility of fine tuning the recombinant expression *via* the inducible promoter. These aspects might be useful for biotechnological applications, although no attempts were performed to define how the expression system works upon up-scaling which really is the crucial step to disclose industrial perspectives. In this respect, the better performances of the system in minimal medium must be underlined both for economical reasons and for the utilization in bioreactors.

Acknowledgments

This work was supported by grants of Ministero dell'Universita e della Ricerca Scientifica (Progetti di Rilevante Interesse Nazionale 2003 to GM; FIRB 2001 to GM), of Programma Nazionale di Ricerche in Antartide 2004, and of Regione Campania L.R. 05/03. Support from the National Center of Excellence in Molecular Medicine (MIUR, Rome) and from the Regional Center of Competence (CRdC ATIBB, Regione Campania, Naples) is gratefully acknowledged. We also thank Prof. G. Feller (University of Liege, Belgium) for kindly providing the DNA coding for β-galactosidase from *Ph*TAE79 and Dr. E. Kopetzki (Boehringer Mannheim GmbH, Department of Genetics, Penzberg, F.R.G.) for kindly providing the DNA coding for α-glucosidase from *Saccharomyces cerevisiae*.

References

Birolo, L., Tutino, M.L., Fontanella, B., Gerday, C., Mainolfi, K., Pascarella, S., Sannia, G., Vinci, F., Marino, G., 2000. Aspartate aminotransferase from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. Cloning, expression, properties, and molecular modelling. Eur. J. Biochem. 267, 2790–2802.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Duilio, A., Marino, G., Mele, A., Sannia, G., Tutino, M.L., 2003. Sistema di espressione di proteine ricombinanti a basse temperature. Uff. It. Brev. Marchi RM2003/A000155.

Duilio, A., Madonna, S., Tutino, M.L., Pirozzi, M., Sannia, G., Marino, G., 2004a. Promoters from a cold-adapted bacterium: definition of a consensus motif and molecular characterization of UP regulative elements. Extremophiles 8, 125–132.

Duilio, A., Tutino, M.L., Marino, G., 2004b. Recombinant protein production in antarctic Gram-negative bacteria in recombinant gene expression. Methods Mol. Biol. 267, 225–238.

Georgiou, G., Valax, P., 1996. Expression of correctly folded proteins in *Escherichia coli*. Curr. Opin. Biotechnol. 7, 190–197.

Hoyoux, A., Jennes, I., Dubois, P., Genicot, S., Dubail, F., Francois, J.M., Baise, E., Feller, G., Gerday, C., 2001. Cold-adapted beta-galactosidase from the Antarctic psychrophile *Pseudoalteromonas haloplanktis*. Appl. Environ. Microbiol. 67, 1529–1535.

Jeon, Y.H., Negishi, T., Shirakawa, M., Yamazaki, T., Fujita, N., Ishihama, A., Kyogoku, Y., 1995. Solution structure of the activator contact domain of the RNA polymerase alpha subunit. Science 270, 1495–1497.

Kopetzki, E., Buckel, P., Schumacher, G., 1989a. Cloning and characterization of baker's yeast alpha-glucosidase: over-expression in a yeast strain devoid of vacuolar proteinases. Yeast 5, 11–24.

Kopetzki, E., Schumacher, G., Buckel, P., 1989b. Control of formation of active soluble or inactive insoluble baker's yeast alpha-glucosidase PI in *Escherichia coli* by induction and growth conditions. Mol. Gen. Genet. 216, 149–155.

Le Thanh, H., Hoffmann, F., 2005. Optimized production of active alpha-glucosidase by recombinant *Escherichia coli*. Evaluation of processes using in vivo reactivation from inclusion bodies. Biotechnol. Prog. 21, 1053–1061.

Luo, Z.H., Hua, Z.C., 1998. Increased solubility of glutathione S-transferase-P16 (GST-pl6) fusion protein by co-expression of chaperones groes and groel in *Escherichia coli*. Biochem. Mol. Biol. Int. 46, 471–477.

Medigue, C., Krin, E., Pascal, G., Barbe, V., Bernsel, A., Bertin, P.N., Cheung, F., Cruveiller, S., D'Amico, S., Duilio, A., Fang, G., Feller, G., Ho, C., Mangenot, S., Marino, G., Nilsson, J., Parrilli, E., Rocha, E.P.C., Rouy, Z., Sekowska, A., Tutino, M.L., Vallenet, D., von Heijne, G., Danchin, A., 2005. Coping with cold: the genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. Genome Res. 10, 1325–1335.

Mitra, A., Chakrabarti, K.S., Shahul Hameed, M.S., Srinivas, K.V., Senthil Kumar, G., Sarma, S.P., 2005. High level expression of peptides and proteins using cytochrome b5 as a fusion host. Protein Exp. Purif. 41, 84–97.

Papa, R., Glagla, S., Krin, E., Danchin, A., Schweder, T., Marino, G., Duilio, A., in press. Identification of a two-component regulatory system in *Pseudoalteromonas haloplanktis* TAC125 by using a proteomic approach. Extremophiles.

Siani, L., Papa, R., Di Donato, A., Sannia, G., in press. Recombinant expression of toluene *o*-xylene monooxygenase (ToMO) from *Pseudomonas stutzeri* OX1 in the marine Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. J. Biotechnol.

Tascon, R.I., Rodriguez-Ferri, E.F., Gutierrez-Martin, C.B., Rodriguez-Barbosa, I., Berche, P., Vazquez-Boland, J.A., 1993. Transposon mutagenesis in Actinobacillus *pleuropneumoniae* with a Tn10 derivative. J. Bacteriol. 175, 5717–5722.

Tosco, A., Birolo, L., Madonna, S., Lolli, G., Sannia, G., Marino, G., 2003. GroEL from the psychrophilic bacterium *Pseudoalteromonas haloplanktis* TAC125: molecular characterization and gene cloning. Extremophiles 7, 17–28.

Tutino, M.L., Duilio, A., Parrilli, E., Remaut, E., Sannia, G., Marino, G., 2001. A novel replication element from an Antarctic plasmid as tool for the expression of proteins at low temperatures. Extremophiles 5, 257–264.

Tutino, M.L., Parrilli, E., Giaquinto, L., Duilio, A., Sannia, G., Feller, G., Marino, G., 2002. Secretion of a-Amylase from *Pseudoalteromonas haloplanktis* TAB23: two different pathways in different hosts. J. Bacteriol. 184, 5814–5817.

RESEARCH LETTER

Identification of the transcription factor responsible for L-malate-dependent regulation in the marine Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125

Rosanna Papa[1], Valentina Rippa[2] & Angela Duilio[2,3]

[1]Department of Public Health Sciences, La Sapienza University, Rome, Italy; [2]Department of Organic Chemistry and Biochemistry, Università Federico II, Naples, Italy; and [3]School of Biotechnological Sciences, Università Federico II, Naples, Italy

Correspondence: Angela Duilio, Dipartimento di Chimica Organica e Biochimica, Complesso Universitario Monte Sant'Angelo, Università Federico II, Via Cinthia 4, 80126 Naples, Italy. Tel.: +39 081 674 314; fax: +39 081 674 313; e-mail: anduilio@unina.it

Received 15 January 2009; revised 10 March 2009; accepted 12 March 2009.
First published online 15 April 2009.

DOI:10.1111/j.1574-6968.2009.01589.x

Editor: Ezio Ricca

Keywords
two-component regulatory system; DNA–protein interaction; transcriptional regulation; σ^{54}; psychrophile.

Abstract

Two-component systems are widespread in nature and constitute the most common mechanism of transmembrane signal transduction in bacteria. Recently, a functionally active two-component system consisting of *malS* and *malR* genes possibly involved in the expression of a C4-dicarboxylate transporter system (*dctAB* operon) was identified in the marine Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. In this paper, we performed a functional analysis of the two-component system and demonstrated its involvement in the regulation of the expression of C4-dicarboxylate transporter genes. The expression of the C4-dicarboxylate transporter genes was induced by L-malate with the promoter element located upstream of the *dctA* gene being active only in the presence of the inducer. A σ^{54} promoter responsible for the L-malate dependent transcription regulation was identified and functionally characterized. The molecular mechanism involves an inverted repeat sequence located upstream the σ^{54} promoter that was shown to bind regulatory proteins only in the presence of L-malate. The protein factor responsible for the induction of the *dctAB* operon expression was eventually identified as the transcriptional regulatory protein MalR. MalR is the first transcriptional factor identified in *P. haloplanktis* TAC125 and one of the few transcriptional modulators reported so far in cold adapted bacteria.

Introduction

Two-component regulatory systems are widespread in nature and exist not only in nearly all prokaryotes and many Archaea but also in eukaryotes such as plants, fungi and yeasts (West & Stock, 2001). A typical two-component regulatory system consists of a signalling histidine kinase (sensor kinase) and a cytoplasmic response regulator, usually a transcription factor acting as activator or repressor. Binding of an environmental signal ligand to the sensory domain of the histidine kinase induces autophosphorylation of the transmitter domain that in turn serves as phosphodonor for its cognate response regulator, thus activating the control protein and leading to transcriptional regulation (Wolanin *et al.*, 2002). In bacteria, these regulatory systems are involved in many signal-processing mechanisms, from chemotaxis and flagellar movement to internal nitrogen availability or pathogenicity, from the control of gene expression for nutrient acquisition and virulence to antibiotic resistance (Galperin, 2004). Two-component regulatory systems are often used by free-living bacteria that have to adapt to frequent changes in nutritional availability and, more generally, in environmental conditions in which they need to finely tune gene expression. By contrast, bacteria that live in stable niches (e.g. symbionts of aphids, rickettsias, extremophiles or some marine hydrocarbon-degrading bacteria) seem to have less-regulated promoters and less regulatory factors (Cases *et al.*, 2003; Cases & de Lorenzo, 2005).

The Antarctic *Pseudoalteromonas haloplanktis* TAC125 is the cold-adapted bacterium, so far characterized, endowed with the highest specific growth rate at low temperature. It was considered a likely candidate to study the molecular basis of physiological adaptation with the added value of having interesting biotechnological features. This microorganism was, in fact, suggested as a promising novel host system for the recombinant protein production at low

temperatures (Duilio *et al.*, 2004a, b). Using genome sequencing, corroborated by *in silico* and *in vivo* analyses, exceptional genomic and metabolic features have been uncovered (Medigue *et al.*, 2005). Annotation of the genome highlighted the presence of a large number of regulatory mechanisms including typical two-component systems, although *P. haloplanktis* TAC125 usually lives in stable environmental niches. Using a proteomic approach, we identified a functionally active two-component system consisting of two Coding DNA Sequences (CDSs) *PSHAb0361* and *PSHAb0362* (Papa *et al.*, 2006). This regulatory element was suggested to be involved in the expression of a C4-dicarboxylate transporter system comprising CDSs *PSHAb0363* and *PSHAb0364*, coding for an outer membrane porin and a putative periplasmic transporter of dicarboxylic acids, respectively (Papa *et al.*, 2006). *Pseudoalteromonas haloplanktis* TAC125 metabolizes L-malate very efficiently. *In silico* analyses, revealed the presence of a gene presumably coding for malic enzyme (PSHAa2725, EC No. 1.1.1.40). These enzymes play a key role in allowing cellular growth on L-malate, a key intermediate of Kreb's cycle (Iwakura *et al.*, 1979). In this paper, we performed a functional analysis of the two-component system with the aim to investigate its involvement in the regulation of the expression of C4-dicarboxylate transporter.

Materials and methods

Methods

Bacterial strains, DNA constructs and media

Pseudoalteromonas haloplanktis TAC125 (Birolo *et al.*, 2000) was collected in 1992 from seawater near the French Antarctic Station Dumont d'Urville (60°40′; 40°01′E). It was routinely grown in aerobic conditions at 150 r.p.m. at 15 °C in minimal medium, containing 1 g L^{-1} KH_2PO_4, 1 g L^{-1} NH_4NO_3, 10 g L^{-1} NaCl, 0.2 g L^{-1} $MgSO_4 \times 7H_2O$, 10 mg L^{-1} $FeSO_4$, 10 mg L^{-1} $CaCl_2 \times 2H_2O$, supplemented with 0.5% casamino acid and 0.2% L-malate as carbon source (Papa *et al.*, 2006) and 100 μg mL^{-1} ampicillin when transformed.

Reverse transcriptase (RT)-PCR

Total RNA was extracted from *P. haloplanktis* TAC125 cells grown up to 3 $OD_{600\,nm}$ as described previously (Tosco *et al.*, 2003) in minimal medium in the presence and in the absence of 0.2% L-malate. RNA was reverse transcribed using SuperScript II RNAse H^- Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions. cDNA was amplified using specific oligonucleotides designed on *malRS* and *dctAB* respectively as primers (Table 1), and *Taq* polymerase (Promega, Madison, WI) according to the manufacturer's instructions. The reaction mixture was amplified (95 °C for 45″; 58 °C for 45″; 72 °C for 60″, 35 cycles), and the products were detected by gel electrophoresis. As a control of the experiment, PCR amplifications on genomic DNA with the same oligonucleotide pairs used for RT-PCR, were performed. Densitometric analysis of the bands for the quantification of transcription levels in the presence and in the absence of L-malate was performed using QUANTITY ONE Biorad software.

Primer extension analysis

Total RNA for primer extension analysis was extracted from recombinant P(PSHA0363) *P. haloplanktis* TAC125 cells (Papa *et al.*, 2006) grown in minimal medium in the presence of L-malate. Primer extension experiments were performed as described previously (Tosco *et al.*, 2003), using the specific PSHAb0363PE Rev oligonucleotide as primer (Table 1).

Preparation of protein extracts

Protein extracts were performed as described previously (Papa *et al.*, 2006).

Construction of promoter and UP (upstream) deletion mutants

Two mutants P(PSHAb0363σ^{70*}) and P(PSHAb0363σ^{54*}) were constructed from P(PSHAb0363) vector applying the one-step overlap extension PCR strategy (Urban *et al.*, 1997). Either the fw/rev σ^{70*} or fw/rev σ^{54*} primers were designed to introduce two single-point mutations, thereby destroying the respective putative σ^{70} −10 box and σ^{54} −24 box (Table 1). Further, a vector containing a mutated version of the upstream sequence containing only the proximal portion of the upstream region was generated and named P(PSHAb0363)UP*. The P(PSHAb0363) vector was hydrolysed with ClaI and EcoRV enzymes. ClaI is located about 1600 bp upstream of the inverted repeat sequence, while EcoRV is located within the inverted repeat sequence (Fig. 2a and c). This enzymatic digestion separates the distal upstream region from the proximal region. A PCR-amplified fragment was obtained using an oligonucleotide containing ClaI in its original position and a mutated oligonucleotide containing EcoRV site immediately upstream the inverted repeat sequence (Table 1). This fragment, opportunely controlled by sequencing, was then inserted into the P(PSHAb0363) hydrolysed vector to generate the P(PSHAb0363)UP* vector.

Table 1. Plasmids and oligonucleotides

	Description	References
Plasmids		
P(PSHAb0363)	pPLB containing the PSHAb0363 promoter region (3600 bp)	Papa *et al*. (2006)
Oligonucleotides		
PSHAbMalS RT-PCR Fw	5′-GGTTTGGTCGAATACATGACTTAGAGC-3′	This work
PSHAbMalS RT-PCR Rv	5′-CATCAACCATAACAACCGAGGTAACTGC-3′	This work
PSHAbMalR RT-PCR Fw	5′-GATGAGGCGATGATCCGCGATTCG-3′	This work
PSHAbMalR RT-PCR Fw	5′-CTAAACCCACCTTCATTCACCAATGCC-3′	This work
PSHAbdctA RT-PCR Fw	5′-CCGATACAGGCGTTATGCGCTACGC-3′	This work
PSHAbdctA RT-PCR Rv	5′-CACTCTGAGTGTTGATACAAAGCACC-3′	This work
PSHAbdctB RT-PCR Fw	5′-GTAGCCGCGCCTTATGCTCATCC-3′	This work
PSHAbdctB RT-PCR Rv	5′-CCACTAAACTCGCCTAGCAAAGGT-3′	This work
PSHAbMalSR RT-PCR Fw	5′-CTAAGGCCGCACCCTTTGTCGGC-3′	This work
PSHAbMalSR RT-PCR Rv	5′-GTAACGTGCAACCAGCTCTTTGC-3′	This work
P(PSHAb0363) PE Rev	5′-CGTTAAGTGGGCTATGTGC-3′	This work
PSHAbdctA bandshift Fw	5′-GGGCATGTAGTAGTGTGCG-3′	This work
PSHAbdctA bandshift Rv	5′-GATGGCCAACTACTAAATAAAC-3′	This work
dctA bandshift control Fw	5′-CTTACCTAATATGGTATTAGC-3′	This work
dctA bandshift control Rv	5′-GGCGCTTCTATTTGTGTGCCTTC-3′	This work
PSHAb0363σ^{70}* Fw	5′-GGGTGTACAACAAAGAGTAGATTC-3′	This work
PSHAb0363σ^{70}* Rv	5′-GAATCTACTCTTTGTTGTACACCC-3′	This work
PSHAb0363σ^{54}* Fw	5′-GTAGTTAACCATCATTTTGCTTAATGG-3′	This work
PSHAb0363σ^{54}* Rev	5′-CCATTAAGCAAAATGATGGTTAACTAC-3′	This work
PSHAb0363UP* Fw	5′-CGTAAAACTATCGATATTAGATACAGG-3′	This work
PSHAb0363UP* Rv	5′-CACTTTTATTTAAAGATATCACCTTTATCG-3′	This work
PSHAb0363C Fw	5′-GATACTTCCTGCAGATATATTAATTG-3′	Papa *et al*. (2006)
PSHAb0363C Rev	5′-CCTGTGTCCCGGGTATCATCATGTGTCC-3′	Papa *et al*. (2006)
PSHAb0363 PE Rv	5′-GCTCTAAGTCATGTATTCGACC-3′	This work

Underlined sequences represent restriction sites inserted.

Transcriptional fusion assays

Recombinant strains were grown either in the presence or in the absence of L-malate up to midstationary phase (24 h). Reporter assays of transcriptional fusions were performed by measuring β-galactosidase activity as described by Duilio *et al*. (2004a, b).

Electromobility shift assay (EMSA)

A 120-bp DNA fragment containing the inverted repeat sequence located upstream *dctA* gene (UP*dctA*) was obtained by PCR amplification using the specific oligonucleotide pairs indicated in Fig. 2a and Table 1. The obtained fragment was ^{32}P labelled with the random primed DNA labelling kit (Roche) according to the manufacturer's instructions. EMSAs were performed in 20-μL reaction volume, in binding buffer (10 mM Tris, pH 7.5, 10 mM EDTA, 0.1 mM PMSF). Total protein extract (10 μg) was incubated with the ^{32}P-labelled DNA (2 ng, 50 000–100 000 c.p.m.) in binding buffer in the presence of 8 μg of poly(dI–dC) as a nonspecific competitor for the binding reaction, for 20 min at room temperature. A 120-bp PCR-amplified fragment of the intragenic region of *dctA* was used as control (for oligonucleotide pairs, see Table 1). Mixtures were then analysed by electrophoresis on 6% native polyacrylamide gel (29 : 1 cross-linking ratio) in TBE buffer (45 mM Tris-HCl, pH 8.0, 45 mM boric acid, 1 mM EDTA). Electrophoreses were performed at room temperature at 200 V. The gels were dried and analysed by autoradiography. In competition experiments, incubations were performed after the addition of five- to 50-fold molar excesses of unlabelled competitor DNA to the reaction mixture, containing proteins and poly(dI-dC).

Western blot analysis

An antiserum against MalR was raised in rabbit (Primm, Milan, Italy), using the synthetic peptide comprising the region 122–136 of the protein (NH_2-CDKRSLVMENRALKR-COOH) as antigen. Protein extracted from *P. haloplanktis* TAC125 cells grown both in the presence and in the absence of L-malate, were resolved by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions using standard procedures (Sambrook & Russell, 2001). The proteins were transferred on to a polyvinylidene difluoride (PVDF) membrane using an electroblotting transfer apparatus (Trans-Blot Semi-Dry Transfer Cell, Bio-Rad). MalR protein was detected using the

anti-rabbit polyclonal antibody produced on the 122–136 target peptide of the protein (1 : 500) and peroxidase-conjugated anti-rabbit secondary antisera (1 : 20 000) (A9169, Sigma). The membrane was developed using SuperSignal West Femto Maximum Sensitivity Substrate detection kit (Pierce) according to the manufacturer's instructions.

Southwestern blot and cross-interaction analyses

Proteins extracted from *P. haloplanktis* TAC125 cells grown both in the presence and in the absence of L-malate, were resolved by SDS-PAGE and transferred on to a PVDF membrane as described previously. The UP*dctA* ^{32}P-labelled fragment used for EMSA was also used for the Southwestern experiment. The PVDF membrane was first washed three times in 10 mM Tris-HCl, pH 7.5, containing 5% nonfat dry skimmed milk, 10% glycerol, 2.5% Triton X-100, 0.1 mM dithiothreitol and 150 mM NaCl at 25 °C. The membrane was then soaked in binding buffer (10 mM Tris-HCl, pH 7.5, 0.125% milk, 8% glycerol, 1 mM dithiothreitol, 1 mM EDTA and 40 mM NaCl). Incubation with UP*dctA* fragment (500 000 c.p.m. mL^{-1}) was performed in 3 mL of binding buffer containing 5 mM $MgCl_2$ at 25 °C for 16 h with gentle agitation. The membrane was then washed three times in 10 mM Tris-HCl, pH 7.5, and 50 mM NaCl for 15 min. The membrane was then exposed for autoradiography. The PVDF membrane preincubated with the anti-MalR sera was then incubated in the presence of the radiolabelled UP*dctA* fragment in binding buffer for 16 h with gentle agitation at 25 °C. The membrane was then washed three times in 10 mM Tris-HCl, pH 7.5, and 50 mM NaCl for 15 min and exposed for autoradiography.

Results

Functional analysis of the two-component regulatory system

We identified four genes possibly involved in the uptake of C4-dicarboxylic acids (Papa *et al.*, 2006). These genes, schematically shown in Fig. 1c, are located on *P. haloplanktis* TAC125 chromosome II. *malS* and *malR* encode for a putative C4-dicarboxylate sensor kinase and a putative C4-dicarboxylate response regulator, respectively. Downstream the *malR* gene, *dctA* and *dctB* are coding for an outer membrane porin and a putative transporter of tri-dicarboxylic acids, respectively. The transcriptional regulation of *malRS* and *dctAB loci* under different conditions was

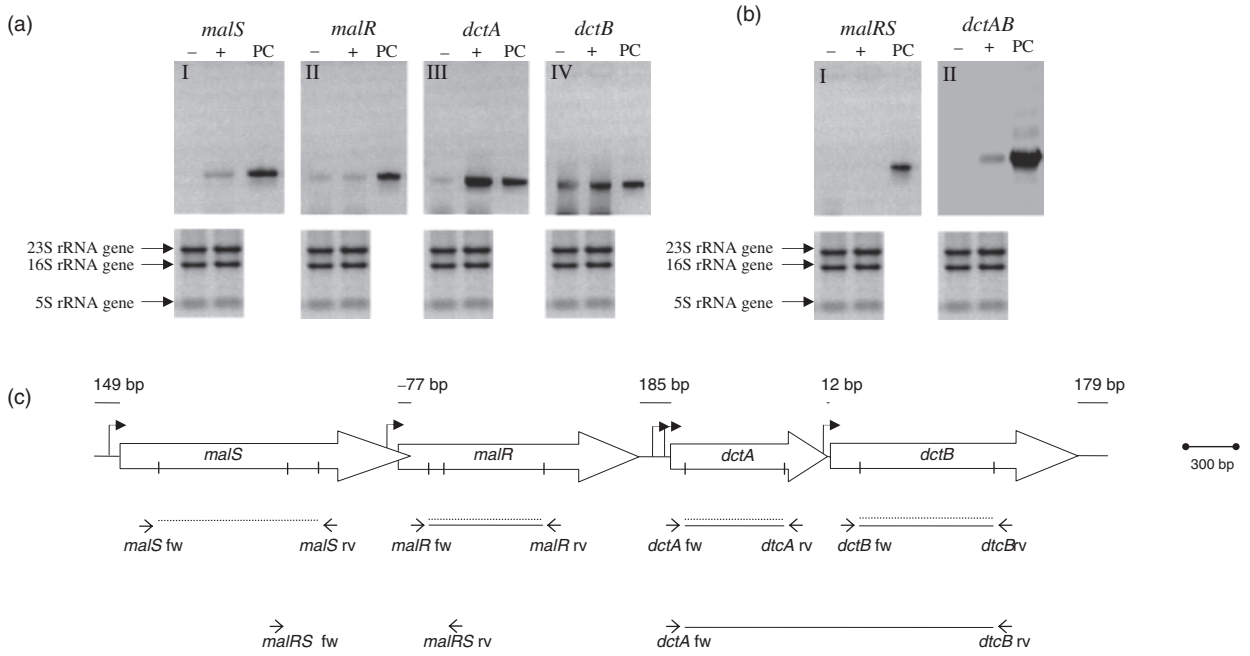


Fig. 1. Schematic organization and transcriptional analysis of *malRS* two-component regulatory system. (a) L-Malate-dependent transcriptional induction of the *malRS* and *dctAB* genes, monitored by RT-PCR (top panels) and visualization of the same samples in an agarose–formaldehyde gel of total RNA (bottom panels). Total RNA samples were extracted from cells grown in minimal medium in the presence (+) and in the absence (−) of L-malate. *Pseudoalteromonas haloplanktis* TAC125 genomic DNA was used as positive control (PC). (b)Transcriptional analysis of *malRS* and *dctAB loci* both in the presence and in the absence of L-malate (top panels) and visualization of the same samples in an agarose–formaldehyde gel of total RNA (bottom panels). *Pseudoalteromonas haloplanktis* TAC125 genomic DNA was used as positive control (PC). (c) A schematic representation of the *malRS* and *dctAB locus* and their transcriptional organization. The position of the oligonucleotides used for the amplifications (←) and the corresponding transcripts are also indicated. ↱, putative promoter element; ——, transcript detected in the absence of L-malate; ·····, transcript detected in the presence of L-malate.

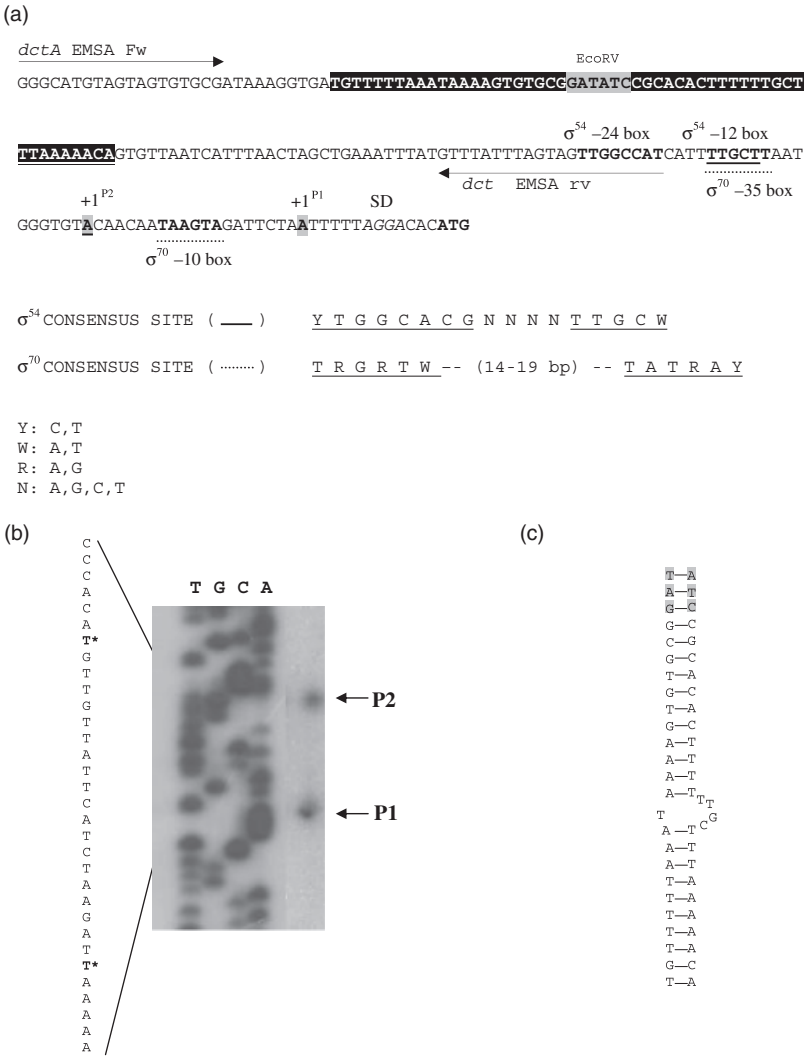


Fig. 2. Transcriptional analysis of *dctAB* operon by primer extension analysis. (a) The nucleotide sequence upstream of the *dctA* gene. The putative P1 and P2 promoters are underlined and highlighted in bold. The corresponding –10 and –35 regions of σ^{70} *core* promoter (indicated with broken lines) and the corresponding –12 and –24 regions of σ^{54} *core* promoter (indicated with continuous lines) are in bold. The transcriptional start sites are underlined and highlighted in grey. The ribosome-binding site (SD) is indicated in italics, while the dctA start codon is indicated in bold. Specific oligonucleotide pairs used to PCR amplify the region are shown. The putative binding site recognized by MalR is highlighted in black (EcoRV endonuclease site is shown). (b) Primer extension analysis of the *dctA* transcript. A 19-bp oligonucleotide was annealed to recombinant P(PSHAb0363) *Pseudoalteromonas haloplanktis* TAC125 total RNA and extended using reverse transcriptase. The nucleotide sequence of the upstream region was determined using the same oligonucleotide as primer (lane T, G, C and A). (c) Putative hairpin structure within the 53-bp-long inverted repeat sequence (EcoRV endonuclease site is shown in grey).

investigated by RT-PCR. RT-PCR analyses were performed on the RNA extracted from *P. haloplanktis* TAC125 cultured in minimal medium both in the presence and in the absence of L-malate. The results are shown in Fig. 1a. *malS* was not detected in the absence of L-malate, whereas a transcription product was weakly detected in the presence of the inducer (Fig. 1a). The light bands in panel II seem to show that transcription of the *malR* gene occurred essentially at the same level under both conditions. Analogously, a *dctA* transcription product was detectable in both conditions although it appeared strongly induced by the presence of L-malate (Fig. 1a). By contrast, transcription of the *dctB* gene did not seem to be strongly affected by the presence of L-malate, showing a band under both conditions (Fig. 1a). However, in the absence of malate, the band in panel IV of Fig. 1a appeared weaker than the band detected when malate was added. A densitometric analysis of the transcripts was then performed using the QUANTITY ONE Biorad software and the data are reported in Table 2. A slight but clear increase in the intensity of the transcript band was observed following malate addition.

We also analysed the possible presence of bicistronic transcripts for both *malRS* and *dctAB loci* (Fig. 1b, Table 1). A cotranscriptional regulation for *malS* and *malR* genes could not be detected, although previous sequence analysis had revealed that these two genes are partially overlapped (7 bp). This result could be probably due to the very low expression level of this locus. As shown in Fig. 1b, a very small amount of the bicistronic transcription product encompassing the *dctAB* locus could only be detected when L-malate was added to the minimal medium. A schematic view of the transcriptional regulation of *malRS* and *dctAB* loci is reported in Fig. 1c.

The occurrence of both the single *dctA* and the *dctA-dctB* bicistronic messengers was further investigated by primer extension experiments. Figure 2b shows the occurrence of

Table 2. Densitometric analysis of *malRS* and *dctAB* transcripts and ribosomal RNA, in the presence and in the absence of L-malate, as reported in Fig. 1

malS		*malR*		*dctA*		*dctB*		*dctAB*	
− malate	+ malate	− malate	+ malate	− malate	+ malate	− malate	+ malate	− malate	+ malate
13.45	308.97	141.36	190.12	208.31	2004.22	523.93	888.59	14.23	381.94

23S rRNA gene		16S rRNA gene		5S rRNA gene	
− malate	+ malate	− malate	+ malate	− malate	+ malate
1667.66	1735.55	1579.79	1590.87	699.20	673.78

To calculate the intensity of the band, the same area was used for each transcript. The value reported for each condition corresponds to the total intensity of all the pixels in the volume divided by the area of the volume (density), as reported in user guide of the QUANTITY ONE software. Estimated errors are c. 5% of the value.

two different transcription initiation sites. A proximal start site, corresponding to an adenine base, was identified only 13 bp upstream of the *dctA* start codon (P1), while a distal start site (adenine) was identified 33 bp upstream of the translational start site (P2). The nucleotide sequence upstream of the *dctA* gene was then carefully examined. A putative σ^{70} promoter element was identified upstream of the proximal start site, according to the *P. haloplanktis* TAC125 consensus sequence previously defined (Duilio *et al.*, 2004a, b). Moreover, a putative σ^{54} promoter element was detected upstream of the distal transcriptional start site. The consensus sequence required to define the σ^{54} -dependent core promoter was derived from a collection of 186 promoters from 47 different bacterial species (Barrios *et al.*, 1999). In contrast to σ^{70}, it is well known that σ^{54} can originate a transcriptionally competent open complex only in the presence of a transcriptional activator, known as bacterial enhancer-binding protein (EBP), usually bound at least 100 bp upstream of the promoter site (Studholme & Dixon, 2003). The DNA region located upstream of the σ^{54} *core* promoter was then examined revealing the presence of a 53-bp long sequence corresponding to a perfect inverted repeat element positioned 71-bp upstream of the P2 transcription initiation site (Fig. 2a–c).

To investigate which promoter is involved in the L-malate-dependent transcriptional regulation, the -10 box of the σ^{70} and the -24 box of the σ^{54} core promoters were individually mutated to destroy the respective promoter consensus sequence. P(PSHAb0363σ^{54}*) and P(PSHAb0363σ^{70}*) were obtained by mutating the nucleotide residues mainly conserved within the corresponding consensus sequences (Barrios *et al.*, 1999; Paget & Helmann, 2003). The nucleotide sequences of the two mutated promoters are shown in Fig. 3a. The activity of the promoters was investigated by transcriptional fusion experiments. The two mutated DNA sequences were individually fused to a promoter-less *lacZ* gene contained in the pPLB plasmid (Duilio *et al.*, 2004a, b). Recombinant cells harbouring the mutated and wild-type promoters were grown in minimal medium either in the absence or in the presence of L-malate and recombinant cells were collected in midstationary growth phase, as the maximum production of β-galactosidase was observed during this growth phase (Papa *et al.*, 2007). As expected, mutation of each core promoter sequence resulted in a clear decrease of β-galactosidase activity in comparison with the corresponding value obtained with the wild-type promoter (Fig. 3b). This result very likely indicates a cooperative effect of two promoters that could be disturbed by the introduction of a mutation in one of two core promoters. However, the transcriptional efficiency of the two mutated promoters was very differently affected by the presence of L-malate. The β-galactosidase activity remained unchanged when L-malate was supplied to cells containing the σ^{54}-mutated vector (PSHAb0363σ^{54}*). On the contrary, cells harbouring the σ^{70} mutated promoter (PSHAb0363σ^{70}*) showed an increase in the β-galactosidase activity when L-malate was added. These findings clearly indicated that only the σ^{54} promoter is involved in the transcriptional regulation mediated by L-malate. Moreover, a transcriptional *lacZ* fusion construct containing a deletion of the distal upstream region (UP) was also prepared. The deletion destroys the inverted and repeat sequence, thus preventing the formation of the hairpin structure (Fig. 2c). Cells harbouring (PSHAb0363)UP* were grown in the presence and in the absence of L-malate. As shown in Fig. 3b, the β-galactosidase activity remained unchanged when L-malate was supplied, suggesting that an intact UP element is needed for transcription regulation.

Identification of a *cis*-regulatory element

Functional analysis of the inverted repeat element located upstream of the P2 transcriptional start site, was carried out by EMSA to investigate whether this region might be able to interact with *P. haloplanktis* TAC125 proteins. A 120-bp DNA fragment encompassing the inverted repeat sequence and the -24 box of the σ^{54} promoter (UP*dctA*) (Fig. 2a), was incubated with the entire protein extracts from cells grown in minimal medium either in the absence or in the presence of L-malate. A clear gel mobility shift band could

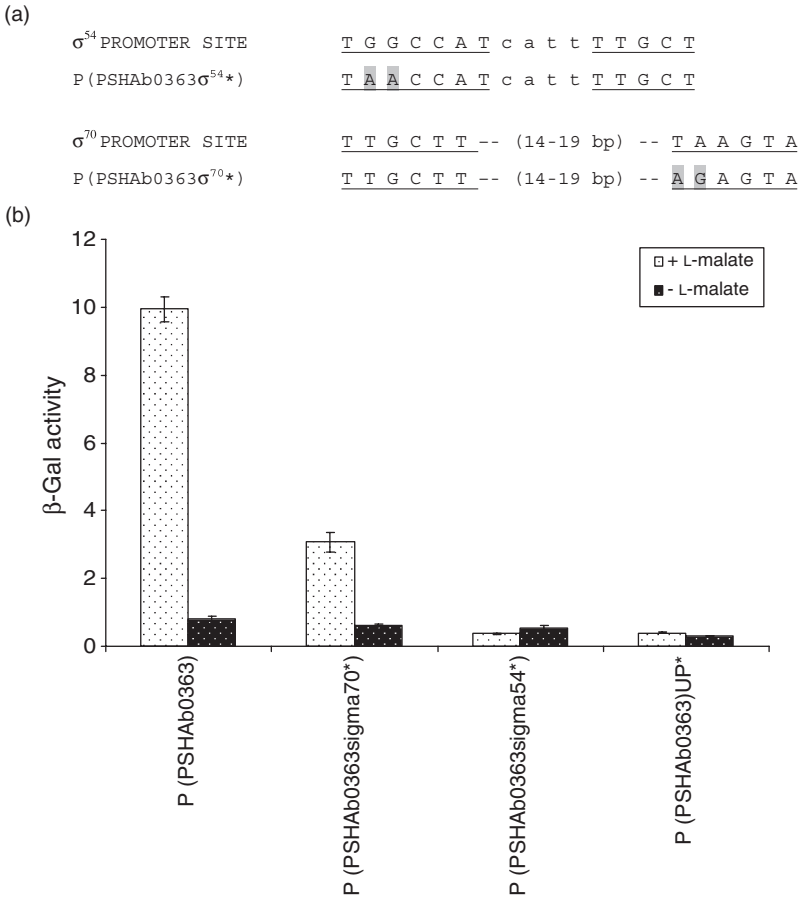


Fig. 3. Functional characterization of σ^{54} and σ^{70} *core* promoters by two single-point mutations. (a) The two *core* promoter sequences are underlined. Single-point mutations for each promoter are highlighted in grey. (b) Profiles of β-galactosidase activity of P(PSHAb0363), P(PSHAb0363σ^{70}*), P(PSHAb0363σ^{54}*) and P(PSHAb0363)UP* transcriptional fusion constructs measured in midstationary phase in *Pseudoalteromonas haloplanktis* TAC125 cells grown in minimal medium in the presence and in the absence of L-malate.

only be detected when the inverted repeat fragment was incubated with the protein extract from cells grown in the presence of L-malate, whereas no binding was observed in the absence of the inducer. A 120-bp DNA fragment of the intragenic region of *dctA* was incubated with the entire protein extracts from cells grown in minimal medium either in the absence or in the presence of L-malate, as control. The free probe is shown for comparison (Fig. 4a). The specificity of the binding was further demonstrated by competition EMSA experiments (Fig. 4b). The specific labelled probe was incubated with total protein extracts in the presence of increasing amounts (five-, 20- and 50-fold) of unlabelled specific and nonspecific DNA competitors.

Identification of the transcriptional regulatory factor

Bacterial EBPs typically consist of three domains, the N-terminal regulatory domain acts either positively or negatively on ring formation and ATPase activity while the central domain with ATPase activity contains the signature GAFTGA motif, which mediates interactions with σ^{54} (Bose *et al.*, 2008). Finally, the C-terminal DNA-binding domain enables specific promoter recognition by allowing EBPs to bind to enhancer-like sequences located *c.* 100–150 bp upstream of the transcription start site. An *in silico* INTERPROSCAN analysis of MalR revealed domain architecture similar to the nitrogen assimilation regulatory protein NtrC, with an N-terminal response regulator receiver domain that contains the phosphorylation site, a central output domain directly responsible for the interaction with the σ^{54} holoenzyme form of RNAP and a C-terminal DNA-binding domain. Mal R also presents the motif that mediates interactions with σ^{54} at residues 223–228, but with a tyrosine replacing the phenylalanine (GAYTGA). For this reason, MalR was suggested to be the likely candidate for the regulation of *dctAB* via σ^{54} activation.

Southwestern blot analysis was then designed to investigate the ability of MalR to specifically recognize the *cis*-regulatory element. The radiolabelled UP*dctA* was incubated with protein extracts from cells grown in the presence and in the absence of L-malate. Figure 5 shows that a single hybridized band with an apparent molecular mass of about 50 kDa was clearly detected in the presence of the inducer. The tiny band observed in the absence of L-malate indicated that recruitment of the regulatory factor also occurred in the

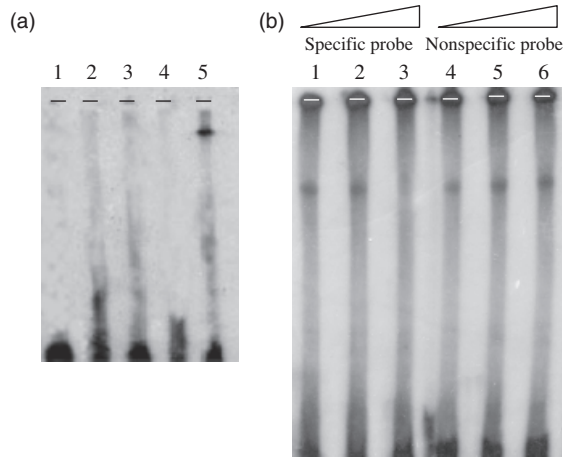


Fig. 4. EMSAs performed using UP*dctA* region and total protein extract of *Pseudoalteromonas haloplanktis* TAC125 cells grown in the presence and in the absence of L-malate. (a) Lane 1, free probe; lane 2, nonspecific probe incubated with proteins extracted from cells grown in the absence of L-malate; lane 3, specific probe incubated with proteins extracted from cells grown in the absence of L-malate; lane 4, nonspecific probe incubated with proteins extracted from cells grown in the presence of L-malate; lane 5, specific probe incubated with proteins extracted from cells grown in the presence of L-malate. (b) Competition assays performed by incubating specific radiolabelled probe with total protein extract of *P. haloplanktis* TAC125 grown in the presence of L-malate, and increasing quantity (five-, 20- and 50-fold) of unlabelled nonspecific and specific DNA competitors.

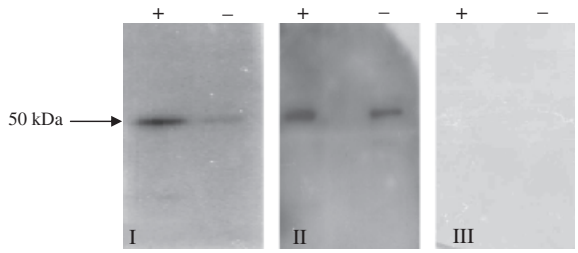


Fig. 5. Southwestern blot analysis of UP*dctA* region and total protein extract of *Pseudoalteromonas haloplanktis* TAC125 cells grown in the presence and in the absence of L-malate. Panel I, Southwestern assay showing the interaction between the UP*dctA* used as probe and a 50-kDa protein in the cellular extract of cells grown both in the presence (+) and in the absence (−) of L-malate. Panel II, Western immunoblot with anti-MalR antibodies showing the presence of MalR protein in the cellular extract of cells grown both in the presence (+) and in the absence (−) of L-malate. Panel III, Cross-competition assay. The PVDF membrane containing the protein extracts from cells grown in the presence (+) and in the absence (−) of L-malate was first incubated with anti-MalR serum and then hybridized with the radiolabelled UP*dctA* fragment.

absence of the inducer, although it was strongly induced when L-malate was added to the medium. The identity of the 50-kDa protein involved in the binding to the regulatory element was univocally assessed by two further experiments. First, the same membrane was developed with antibodies raised against the MalR protein. A clear immunostained band perfectly matching the radiolabelled signal was detected by the anti-MalR antibody when the nucleotide probe was incubated with the protein extract from cells grown in the presence of L-malate. According to the transcriptional data, an identical immunostained band could also be detected when L-malate was omitted from the medium. Second, the specificity of the interaction between the MalR and the DNA *cis*-regulatory element was confirmed by a specific cross-competition experiment. The membrane containing the protein extracts from cells grown in the presence and in the absence of L-malate was first incubated with anti-MalR serum and then hybridized with the radiolabelled DNA fragment. No hybridization bands could be detected following interaction between the radiolabelled DNA fragment and the cell extract, as it was clearly observed in the previous experiment (Fig. 5), suggesting that anti-MalR antibodies recruited MalR preventing its interaction with the DNA fragment. As a whole, these findings demonstrated that MalR is able to specifically recognize the *cis*-regulatory element located upstream of *dctA* gene and to regulate the L-malate dependent expression of the downstream genes.

Discussion

A functionally active two-component system was recently identified in the Antarctic bacterium *P. haloplanktis* TAC125 by differential proteomic experiments (Papa *et al.*, 2006). This regulatory element consists of *PSHAb0361* and *PSHAb0362* and was suggested to be involved in the expression of a C4-dicarboxylate transporter system comprising *PSHAb0363* and *PSHAb0364* (Papa *et al.*, 2006). The proteins encoded by *PSHAb0361* and *PSHAb0362*, hereby indicated as *malS* and *malR*, exhibited 31.7% and 58.3% identity with a C4-dicarboxylate transport sensor protein (Q87R82) and a C4-dicarboxylate transport transcriptional regulatory protein (Q8ECK1) from *Vibrio parahaemolyticus* and *Shewanella oneidensis*, respectively. *PSHAb0363* encoded a putative protein sharing 33.1% identity with an outer membrane porin from *S. oneidensis* (Q8EGP2). *PSHAb0364* coding for a putative protein sharing 46.3% identity with a di-tricarboxylate transporter from *Corynebacterium glutamicum* (Q8NTS7) was found 12 bp downstream of the outer membrane porin gene, thus suggesting the involvement of these two CDSs in the transport and the uptake of C4-dicarboxylic acids. Accordingly, these CDSs were renamed *dctA* and *dctB* (C4-dicarboxylate transport), respectively. Functional analysis of the *malRS* and *dctAB* loci seems to indicate that each gene is independently transcribed by its own promoter in both loci. However, functional differences exist among the various promoters. The promoter elements located upstream of the *malS* gene and, particularly, the *dctA* gene were strongly affected by L-malate with the former

being active only in the presence of the inducer. On the contrary, the promoter located upstream of the *malR* and *dctB* genes seem to be insensitive or slightly responsive to the presence of L-malate, as in the case of *dctB*. The malate dependence of the promoter located upstream of the *dctA* gene confirmed the involvement of the corresponding porin in the L-malate uptake, as reported previously (Papa *et al.*, 2006). The occurrence of two different transcription initiation sites upstream of the *dctAB* locus was demonstrated. A putative σ^{70} promoter element was identified upstream of the proximal start site. Moreover, a putative σ^{54} promoter element was inferred upstream of the distal transcriptional start site (Barrios *et al.*, 1999). The σ^{54} core promoters seem to be remarkably highly conserved both in sequence and in structure as compared with the more variable σ^{70} promoters. This conservation very likely reflects the strict requirements for promoter recognition and function needed for a highly controlled regulation. Functional analysis of the two promoters demonstrated that only the activity of the σ^{54} promoter was affected by L-malate. However, the *dctA* gene was found to be expressed from both promoters, the constitutive σ^{70} promoter active also in the absence of L-malate and responsible for the basal level of the *dctA* transcript, and the inducible σ^{54} promoter, which assures high levels of transcript in the presence of the inducer.

In contrast to σ^{70} RNAP bound at its cognate promoter sites, σ^{54} RNAP is unable to spontaneously isomerize from a closed complex to a transcriptionally competent open complex (Studholme & Dixon, 2003). To proceed with initiation of transcription, the closed complex must participate in an interaction with transcriptional activators known as bacterial EBPs, involving nucleotide hydrolysis. Transcriptional activators usually bind at least 100 bp upstream of the promoter site, and DNA looping is required for the activator to contact the closed complex and catalyse the formation of the open promoter complex. Accordingly, the molecular mechanism exerted by L-malate on the expression of the *dctAB* operon involves the inverted repeat sequence located upstream of the σ^{54} promoter. This DNA sequence represents a *cis*-acting region able to bind *P. haloplanktis* TAC125 regulatory proteins only when the cells were grown in the presence of L-malate. The protein factor responsible for the L-malate dependent induction was identified as the transcriptional regulatory protein MalR.

In silico analysis revealed a close similarity of MalR sequence with NtrC, a typical bacterial EBP. Two-component systems have evolved to allow bacteria to sense and respond to a wide range of stresses and environmental cues, using specialized EBPs. Examples include the nitrogen assimilation regulatory protein NtrC and the dicarboxylic acid transport regulator DctD (Studholme & Dixon, 2003). Bacterial EBPs show a classical architecture consisting of three domains, the N-terminal regulatory domain acting either positively or negatively on ring formation and ATPase activity, the central AAA+domain endowed with ATPase activity and the C-terminal DNA-binding domain enabling specific promoter recognition.

Binding of MalR to the inverted repeat sequence produced a very highly retarded shift band in the EMSA experiment, suggesting something very large binding to it, either RNAP or a multimer of MalR. Indeed, upon phosphorylation, NtrC forms oligomers (possibly an octamer) that can hydrolyse ATP and couple the energy available from ATP hydrolysis to the formation of an open complex with σ^{54} holoenzyme (Lee *et al.*, 2001). Interaction of MalR with the DNA inverted repeat sequence very likely stabilizes the interaction of the RNAP on the σ^{54} core promoter thus activating the transcription of downstream genes. However, transcription of *malR* gene and production of MalR protein seem to occur at the same level both in the absence and in the presence of L-malate. A possible explanation for this apparent discrepancy is that L-malate induces the expression of the MalS histidine kinase that can act as phospho-donor for its cognate MalR response regulator increasing the amount of the active form of the control protein and leading to transcriptional activation.

Acknowledgement

This work was supported by grants of MURST (PRIN 2003, and FIRB 2001 to Gennaro Marino), of PNR in Antartide 2004 to Gennaro Marino, and of Regione Campania L.R. 05/03.

References

Barrios H, Valderrama B & Morett E (1999) Compilation and analysis of σ^{54}-dependent promoter sequences. *Nucleic Acids Res* **27**: 4305–4313.

Birolo L, Tutino ML, Fontanella B, Gerday C, Mainolfi K, Pascarella S, Sannia G, Vinci F & Marino G (2000) Aspartate aminotransferase from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC 125. Cloning, expression, properties, and molecular modelling. *Eur J Biochem* **267**: 2790–2802.

Bose D, Joly N, Pape T, Rappas M, Schumacher J, Buck M & Zhang X (2008) Dissecting the ATP hydrolysis pathway of bacterial enhancer-binding proteins. *Biochem Soc T* **36**: 83–88.

Cases I & de Lorenzo V (2005) Promoters in the environment: transcriptional regulation in its natural context. *Nature Rev Genet Microbiol* **3**: 105–118.

Cases I, de Lorenzo V & Ouzounis CA (2003) Transcription regulation and environmental adaptation in bacteria. *Trends Microbiol* **11**: 248–253.

Duilio A, Madonna S, Tutino ML, Pirozzi M, Sannia G & Marino G (2004a) Promoters from a cold-adapted bacterium:

definition of a consensus motif and molecular characterization of UP regulative elements. *Extremophiles* **8**: 125–132.

Duilio A, Tutino ML & Marino G (2004b) Recombinant protein production in Antarctic Gram-negative bacteria. *Meth Mol Biol* **267**: 225–237.

Galperin MY (2004) Bacterial signal transduction network in a genomic perspective. *Environ Microbiol* **6**: 552–567.

Iwakura M, Hattori J, Arita Y, Tokushige M & Katsuki HW (1979) Studies on regulatory functions of malic enzymes. VI. Purification and molecular properties of NADP-linked malic enzyme from *Escherichia coli* W. *J Biochem (Tokyo)* **85**: 1355–1365.

Lee YS, Han JS, Jeon Y & Hwang DS (2001) The arc two-component signal transduction system inhibits *in vitro Escherichia coli* chromosomal initiation. *J Biol Chem* **276**: 9917–9923.

Medigue C, Krin E, Pascal G *et al.* (2005) Coping with cold: the genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. *Genome Res* **15**: 1325–1335.

Paget MSB & Helmann JD (2003) The σ70 family of sigma factors. *Genome Biol* **4**: 203–208.

Papa R, Glagla S, Danchin A, Schweder T, Marino G & Duilio A (2006) Proteomic identification of a two-component regulatory system in *Pseudoalteromonas haloplanktis* TAC125. *Extremophiles* **10**: 483–491.

Papa R, Rippa V, Sannia G, Marino G & Duilio A (2007) An effective cold inducible expression system developed in *Pseudoalteromonas haloplanktis* TAC125. *J Biotechnol* **127**: 199–210.

Sambrook J & Russell DW (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Studholme DJ & Dixon R (2003) Domain architectures of σ54-dependent transcriptional activators. *J Bacteriol* **185**: 1757–1767.

Tosco A, Birolo L, Madonna S, Lolli G, Sannia G & Marino G (2003) GroEL from the psychrophilic bacterium *Pseudoalteromonas haloplanktis* TAC125: molecular characterization and gene cloning. *Extremophiles* **7**: 17–28.

Urban A, Neukirchen S & Jaeger KE (1997) A rapid and efficient method for site-directed mutagenesis using one-step overlap extension PCR. *Nucleic Acids Res* **25**: 2227–2228.

West AH & Stock AM (2001) Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem Sci* **26**: 369–377.

Wolanin PM, Thomason PA & Stock JB (2002) Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biol* **3**: 3013.1–3013.8.