Protein Engineering vol.13 no.5 pp.329–337, 2000

Linker insertion mutagenesis based on IS*21* transposition: isolation of an AMP-insensitive variant of catabolic ornithine carbamoyltransferase from *Pseudomonas aeruginosa*

Thomas Seitz[1], Bernard Berger[1], Van Thanh Nguyen[1,2], Catherine Tricot[2], Vincent Villeret[2], Sergio Schmid[3], Victor Stalon[2,4] and Dieter Haas[1,5]

[1]Laboratoire de Biologie Microbienne, Université de Lausanne, CH-1015 Lausanne, Switzerland, [2]Institut de Recherches Microbiologiques Jean-Marie Wiame, B-1070 Brussels, Belgium, [3]Ecole d'Ingénieurs du Valais, CH-1950 Sion, Switzerland and [4]Laboratoire de Microbiologie, Université Libre de Bruxelles, B-1070 Brussels, Belgium

[5]To whom correspondence should be addressed. E-mail: dieter.haas@lbm.unil.ch

The bacterial insertion sequence IS*21* when repeated in tandem efficiently promotes non-replicative cointegrate formation in *Escherichia coli*. An IS*21*–IS*21* junction region which had been engineered to contain unique *Sal*I and *Bgl*II sites close to the IS*21* termini was not affected in the ability to form cointegrates with target plasmids. Based on this finding, a novel procedure of random linker insertion mutagenesis was devised. Suicide plasmids containing the engineered junction region (pME5 and pME6) formed cointegrates with target plasmids in an *E.coli* host strain expressing the IS*21* transposition proteins *in trans*. Cointegrates were resolved *in vitro* by restriction with *Sal*I or *Bgl*II and ligation; thus, insertions of four or 11 codons, respectively, were created in the target DNA, practically at random. The cloned *Pseudomonas aeruginosa arcB* gene encoding catabolic ornithine carbamoyltransferase was used as a target. Of 20 different four-codon insertions in *arcB*, 11 inactivated the enzyme. Among the remaining nine insertion mutants which retained enzyme activity, three enzyme variants had reduced affinity for the substrate ornithine and one had lost recognition of the allosteric activator AMP. The linker insertions obtained illustrate the usefulness of the method in the analysis of structure–function relationships of proteins.

Keywords: catabolic ornithine carbamoyltransferase/IS*21*/linker insertion mutagenesis/transposition/*Pseudomonas aeruginosa*

Introduction

Insertion of oligopeptides into proteins at various sites provides a convenient means to investigate structure–function relationships. Some insertions made at 'permissive' sites may not affect the function of a protein, whereas insertions at other sites may abolish protein function altogether. However, the most interesting insertions may be those that change protein properties by modifying substrate or effector recognition (Shortle and Sondek, 1995; Hayes *et al.*, 1997; Manoil and Bailey, 1997).

Short, in-frame linker insertions can be generated by different mutagenesis methods. *In vitro*, target genes can be opened at specific sites with restriction endonucleases or at random sites by limited digestion with DNase I. Short oligonucleotides of 2–4 codons are then inserted (Barany, 1985; Goff and Prasad, 1991; Kegler-Ebo *et al.*, 1994). Drawbacks of these methods are that, on the one hand, restriction sites are not distributed randomly and, on the other, DNase I rarely produces flush double-stranded breaks without adjacent deletions (Graf and Schachman, 1996), although improvements of the method using DNase I have alleviated the deletion problem to some extent (Dykxhoorn *et al.*, 1997). Another type of approach uses transposable elements containing artificially introduced restriction sites close to both ends. After insertion *in vivo*, the bulk of the transposable element lying between the introduced restriction sites is cut out *in vitro* and ligation then creates a residual small insertion, which is formed by the ends of the transposable element and the target duplication. This principle has been used with derivatives of Tn*3* (Hoekstra *et al.*, 1991), Tn*5* (Manoil and Bailey, 1997; Nelson *et al.*, 1997), Tn*4430* (Hallet *et al.*, 1997; Hayes *et al.*, 1997) and Moloney leukemia virus (MoLV) (Singh *et al.*, 1997).

Here we describe the development and application of a novel insertion mutagenesis procedure that produces insertions of four or 11 codons with little or no target specificity. The method is based on replicon fusion mediated by the tandemly repeated insertion sequence IS*21*, in *Escherichia coli* as the host. Plasmids carrying an IS*21* tandem readily form cointegrates with other replicons in a non-replicative reaction that resembles integration of retroviruses (Reimmann and Haas, 1990). IS*21*-dependent cointegrate formation requires two IS*21*-encoded proteins, cointegrase (the 45 kDa *istA* gene product) and the IstB accessory protein (Schmid *et al.*, 1998). The cointegrates formed *in vivo* can then be freed of the donor plasmid and the major part of both flanking IS*21* elements (engineered to contain unique restriction sites at their ends), by restriction and ligation *in vitro*. In this way, small in-frame insertions can be generated. We have chosen catabolic ornithine carbamoyltransferase (cOTC) of *Pseudomonas aeruginosa* as the target of our linker insertion mutagenesis method. This dodecameric enzyme catalyzes the reaction citrulline + $P_i \rightarrow$ ornithine + carbamoylphosphate, the second step of the arginine deiminase pathway (Baur *et al.*, 1987; Marcq *et al.*, 1991; Villeret *et al.*, 1995). The deiminase pathway provides the cells with ATP in the absence of respiration (Vander Wauven *et al.*, 1984). AMP, a signal of low energy, allosterically activates cOTC (Tricot *et al.*, 1993, 1998; Sainz *et al.*, 1998). A linker insertion has now allowed us to obtain for the first time an AMP-insensitive cOTC. Some preliminary results have been included in a previous review (Haas *et al.*, 1996).

Materials and methods

Growth conditions, bacterial strains and plasmids

Media, selective antibiotic concentrations and conditions of incubation have been described before (Jeenes *et al.*, 1986; Reimmann and Haas, 1990). *E.coli* strains and plasmids are listed in Table I.

Assay of cointegrate formation in a three-plasmid system

The detailed procedure has been described by Schmid *et al.* (1998). Cointegration frequencies are calculated as cointegrates

Table I. Bacterial strains and plasmids

E.coli strain	Genotype	Reference or construction
CC118(λ*pir*)	*araD139 Δ(ara leu)7697 ΔlacX74 phoAΔ20 galE thi rpsE rpoB* $argE_{am}$ *recA1 (λpir)*	Herrero *et al.* (1990)
CM236	*Δ(lac proAB argF) ΔargI thi-1 supE recA1 srl::Tn10,* r_K^-, m_K^-	Jeenes *et al.* (1986)
ED8767	*metB hsdS supE supF recA56*	Sambrook *et al.* (1989)
ECOLIST = RR28*istA*(P45)*istB*	*pheS12 hsdS20 recA proA2 leu-6 ara-14 galK2 lacY1 xyl-5 mtl-1 supE44 rpsL20 endA thi-1 [*$lacI^Q$ *istA*(P45), *istB rrnB] in attB*	This work, Figure 3
HB101	*hsdS recA proA2 leu-6 ara-14 galK2 lacY1 xyl-5 mtl-1 rpsL20 thi-1 supE44*	Sambrook *et al.* (1989)
RR28	*pheS12 hsdS20 recA proA2 leu-6 ara-14 galK2 lacY1 xyl-5 mtl-1 supE44 rpsL20 endA thi-1*	Hennecke et al. (1982)
Plasmid	**Phenotype or relevant property**	**Reference or construction**
pACYC184	Tc Cm, p15A replicon	Chang and Cohen (1978)
pBluescript KS(+)	Ap, ColE1 replicon	Stratagene
pGP704	Ap, R6K suicide replicon	Herrero *et al.* (1990)
pHP45Ω	Ap Sm/Sp, ColE1 replicon	Prentki and Krisch (1984)
pJP5603	Km, R6K suicide replicon	Penfold and Pemberton (1992)
pLDR8	Km, pSC101 temperature-sensitive replicon carrying *int*(λ)	Diederich *et al.* (1992)
pLDR10	Cm Ap, ColE1 replicon carrying *attP*(λ)	Diederich *et al.* (1992)
pME5	Km, R6K suicide replicon carrying IS*21*–IS*21* junction region with *Sal*I and *Bgl*II sites	This work, Figure 2b
pME6	Cm, R6K suicide replicon carrying IS*21*–IS*21* junction region with *Sal*I and *Bgl*II sites	This work, Figure 2b
pME183–7	Ap Sm, IncQ replicon carrying *arcDAB* genes	Nguyen *et al.* (1994)
pME3659	Ap, ColE1 replicon carrying *arcB*	This work, Figure 1b
pME3913	Ap, ColE1 replicon carrying *istA*(P45) *istB*	Schmid *et al.* (1998)
pME3915	Cm, p15A replicon carrying IS*21*-IS*21* junction	Schmid *et al.* (1998)
pME3920	Cm Sm/Sp, ColE1 replicon carrying *attP* $lacI^Q$ *istA*(P45)*istB*	This work, Figure 1a
pME3923	Cm, p15A replicon carrying IS*21*–IS*21* junction region with *Bgl*II sites	This work, Figure 2a
pME3924	Cm, p15A replicon carrying IS*21*–IS*21* junction region with *Sal*I and *Bgl*II sites	This work, Figure 2a
R751	Tp Tra, IncPβ replicon	Hirsch and Beringer (1984)

Antibiotic resistance markers are abbreviated as explained in the text; Tp, trimethoprim.

(R751::pME3915, R751::pME3923 or R751::pME3924) transferred per target plasmid (R751) transferred, in conjugation with *E.coli* HB101 as the recipient.

Recombinant DNA methods and plasmid construction

Standard DNA methods were utilized (Sambrook *et al.*, 1989) with some refinements quoted by Schmid *et al.* (1998). To introduce the *istA*(P45) and *istB* genes into the *E.coli* chromosome, the delivery system described by Diederich *et al.* (1992) was used, pLDR10 being the vehicle. The $lacI^Q$ gene, the *tac* promoter, the *istA*(P45) and *istB* genes and the *rrnB* terminator were recruited, on a 4.5 kb *Nru*I–*Sca*I fragment, from pME3913, extended with linkers containing *Bgl*II and *Not*I sites and inserted into the 2.7 kb backbone of pLDR10 cut with *Bam*HI and *Not*I (partially). The resulting 7.2 kb plasmid was cut partially with *Eco*RI such that the 2.0 kb *Eco*RI ΩSp/Sm fragment (Prentki and Krisch, 1984) could be inserted next to the *attP* site. Thus, pME3920 (Figure 1a) was obtained. Its *Not*I fragments were circularized by ligation and introduced by transformation into *E.coli* RR28/pLDR8, with selection for streptomycin (20 μg/ml) and spectinomycin (20 μg/ml) resistance. The temperature-sensitive helper plasmid pLDR8 carries the λ *int* gene (Diederich *et al.*, 1992). Of 80 transformants obtained, 70 were Cm-sensitive and presumably had integrated the 7.1 kb *Not*I fragment (Figure 1a) into the chromosome, via *int*-driven recombination between the *attP* site on the 7.1 kb *Not*I fragment and the *attB* site in the chromosome. The correct integration of the 7.1 kb fragment was verified in four isolates by Southern hybridization, using an IS*21* probe (the 2.0 kb *Sal*I–*Hin*dIII fragment of pME3913). All four isolates gave the expected pattern (data not shown) and one representative isolate designated ECOLIST was kept.

The *arcB* gene was subcloned from pME183-7 (Nguyen *et al.*, 1994) on a minimal 1.1 kb *Xho*I–*Eco*RI fragment

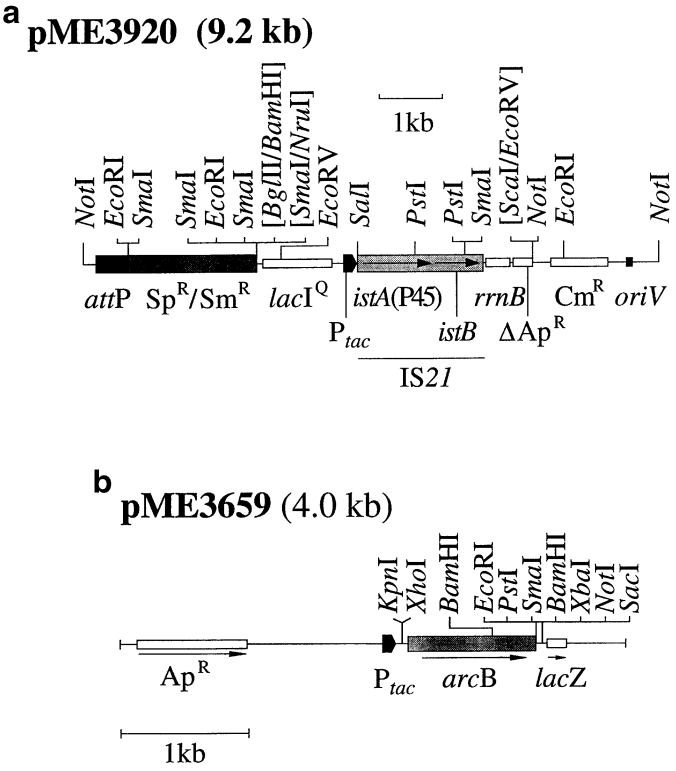

Fig. 1. (**a**) Construction of pME3920. The *istA*(P45) and *istB* genes of IS*21* are expressed from the *tac* promoter (Ptac) and controlled by the $lacI^Q$ repressor gene. The 7.1 kb *Not*I fragment was circularized by ligation and integrated into the *E.coli* chromosome by recombination between the *attP* site of pME3920 and the chromosomal *attB* site. Δ, deletion; *oriV*, origin of replication. (**b**) Construction of pME3659. This is a pBluescript derivative carrying the *arcB* gene of *P.aeruginosa*.

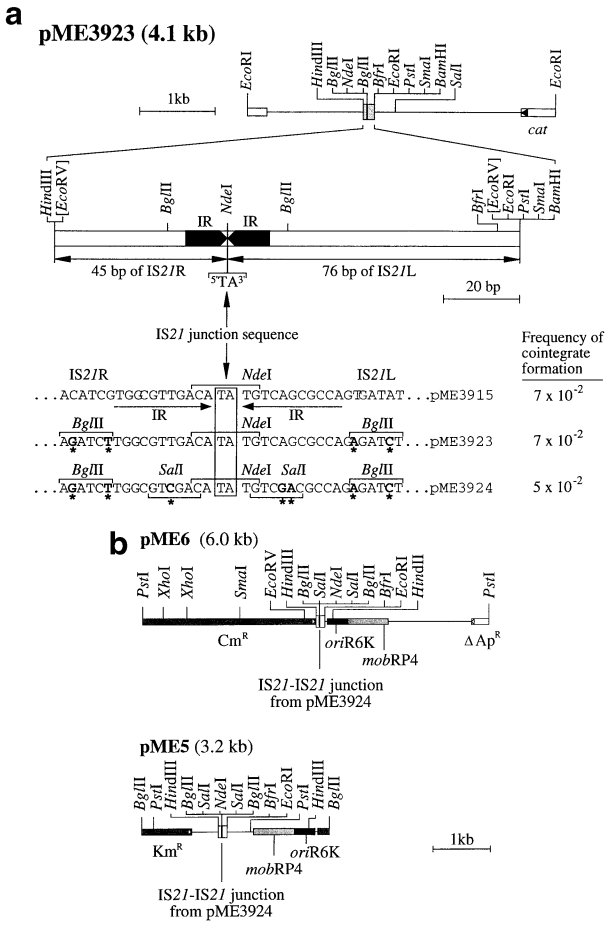

Fig. 2. (**a**) Modification of the IS*21*–IS*21* junction region by introduction of restriction sites: effect on non-replicative cointegrate formation in a three-plasmid system. Plasmid pME3915, a pACYC184 derivative containing wild-type IS*21* termini, has been described (Schmid *et al.*, 1998). Plasmids pME3923 and pME3924 were constructed sequentially by insertion of synthetic oligonucleotides into pACYC184 (Materials and methods). The three IS*21*–IS*21* junction constructs formed cointegrates with R751 in the presence of the helper pME3913; cointegration frequencies are calculated as previously described (Schmid *et al.*, 1998). (**b**) Construction of R6K-derived suicide plasmids containing the modified IS*21*–IS*21* junction region from pME3924. See Materials and methods for details. *mob*, Site required for conjugal mobilization by IncP plasmids.

into pBluescript KS(+) such that the vector's *lac* promoter transcribed *arcB* (Figure 1b).

For the construction of pME3923 (Figure 2a), a synthetic 123 bp linker was constructed from seven overlapping phosphorylated oligonucleotides. The 123 bp region contains 76 bp of IS*21*R, a TA spacer and 45 bp of IS*21*L. This region is identical with the wild-type IS*21* sequence, except for four base changes introducing two *Bgl*II restriction sites (Figure 2a). The assembled linker was inserted into the *Eco*RV site of pBluescript KS(+), checked by nucleotide sequencing and inserted, as a *Bam*HI–*Hin*dIII fragment, into pACYC184 cut with *Bam*HI and *Hin*dIII, producing pME3923. The additional three point mutations giving two *Sal*I restriction sites in pME3924 (Figure 2a) were introduced on a 31 bp phosphorylated oligonucleotide containing *Bgl*II-compatible ends. This oligonucleotide replaced the corresponding fragment in pME3923.

Plasmid pME5 was constructed by subcloning the 0.14 kb *Eco*RI–*Hin*dIII fragment of pME3924, which carries the IS*21*–IS*21* junction region including *Sal*I and *Bgl*II sites (Figure 2a), into pJP5603, cut with *Hin*dIII (partially) and *Eco*RI. Plasmid pME6 is based on the 3.0 kb fragment of pGP704, obtained by digestion with *Hin*dIII (partial) and *Pst*I. This fragment was combined with a 2.9 kb *Pst*I–*Hin*dIII Cm resistance cassette from Tn*1725 (*Ubben and Schmitt, 1986) and the 0.14 kb IS*21*–IS*21* junction region of pME3924 (delimited by *Eco*RI and *Hin*dIII sites). Both pME5 and pME6 (Figure 2b) were replicated in *E.coli* CC118/λ*pir* and prepared from this strain by Qiagen column purification (Qiagen Inc.).

Linker insertion mutagenesis protocol

E.coli ECOLIST containing pME3659 was grown with shaking in nutrient yeast broth (NYB) containing ampicillin (Ap) (100 μg/ml). An overnight culture was diluted 1:100 in the same medium and isopropyl-β-D-thiogalactoside (IPTG) was added to a final concentration of 1 mM, to induce the *istA*(P45) and *istB* genes. When cells had reached an optical density corresponding to about 4×10^8 cells/ml, they were harvested by centrifugation and made competent by treatment with 100 mM $CaCl_2$ (Sambrook *et al.*, 1989). Competent cells were transformed in 100 μl with about 0.2 μg of pME5 or pME6 DNA. Several transformation experiments were conducted in parallel, to ensure formation of a sufficient number of independent transformants. After incubation on ice for 30 min and a 2 min heat shock (42°C), cells were incubated in NYB for 90 min, concentrated 20-fold and plated on nutrient agar (NA) containing 50 μg/ml kanamycin (Km) (for pME5) or 50 μg/ml chloramphenicol (Cm) (for pME6). This procedure typically resulted in ⩾20 Km- or Cm-resistant colonies, respectively. From each individual transformation tube one clone was picked and grown in selective NYB. Plasmid mini-preparations (Del Sal *et al.*, 1988) were analyzed for the presence of pME3659::pME5 or pME3659::pME6 cointegrates among unreacted pME3659. Digestion with *Kpn*I plus *Xba*I (enzymes cutting the polylinker flanking the *arcB* gene, but not cutting the suicide plasmids) was used to map the pME5 or pME6 insertions to the target gene or the vector moiety (Figure 3). Cointegrates containing an insertion in the *arcB* gene were separated from pME3659 by transformation of *E.coli* RR28 with 100–1000-fold diluted mini-preps. The insertion sites were sequenced by the method of Tsang and Bentley (1988) using two IS*21*-specific primers: 5′-TGTT-GGGTGGAGCGG-3′, positions 60–46 at the 5′ end of IS*21*, and 5′-GGGCATGAAAATGGC-3′, positions 2087–2101 at the 3′ end of IS*21* (Reimmann *et al.*, 1989). These primers read from the IS*21* termini into the target gene.

Cointegrates of interest were restricted with *Sal*I, ligated and introduced into *E.coli*, with Ap selection. Alternatively, *Bgl*II was used for resolution of cointegrates. Loss of Km or Cm resistance confirmed that the bulk of the suicide plasmid had been excised. Expression of mutated cOTCs was tested in the OTC-negative *argF argI* mutant CM236 of *E.coli*.

Assay of cOTC

The assay conditions have been detailed by Nguyen *et al.* (1996).

Results

Some point mutations in the inverted repeats of IS21 do not affect cointegrate formation

The method of IS*21*-based linker insertion mutagenesis described below relies on the observation that two IS*21* termini, joined together by a short junction sequence (typically two

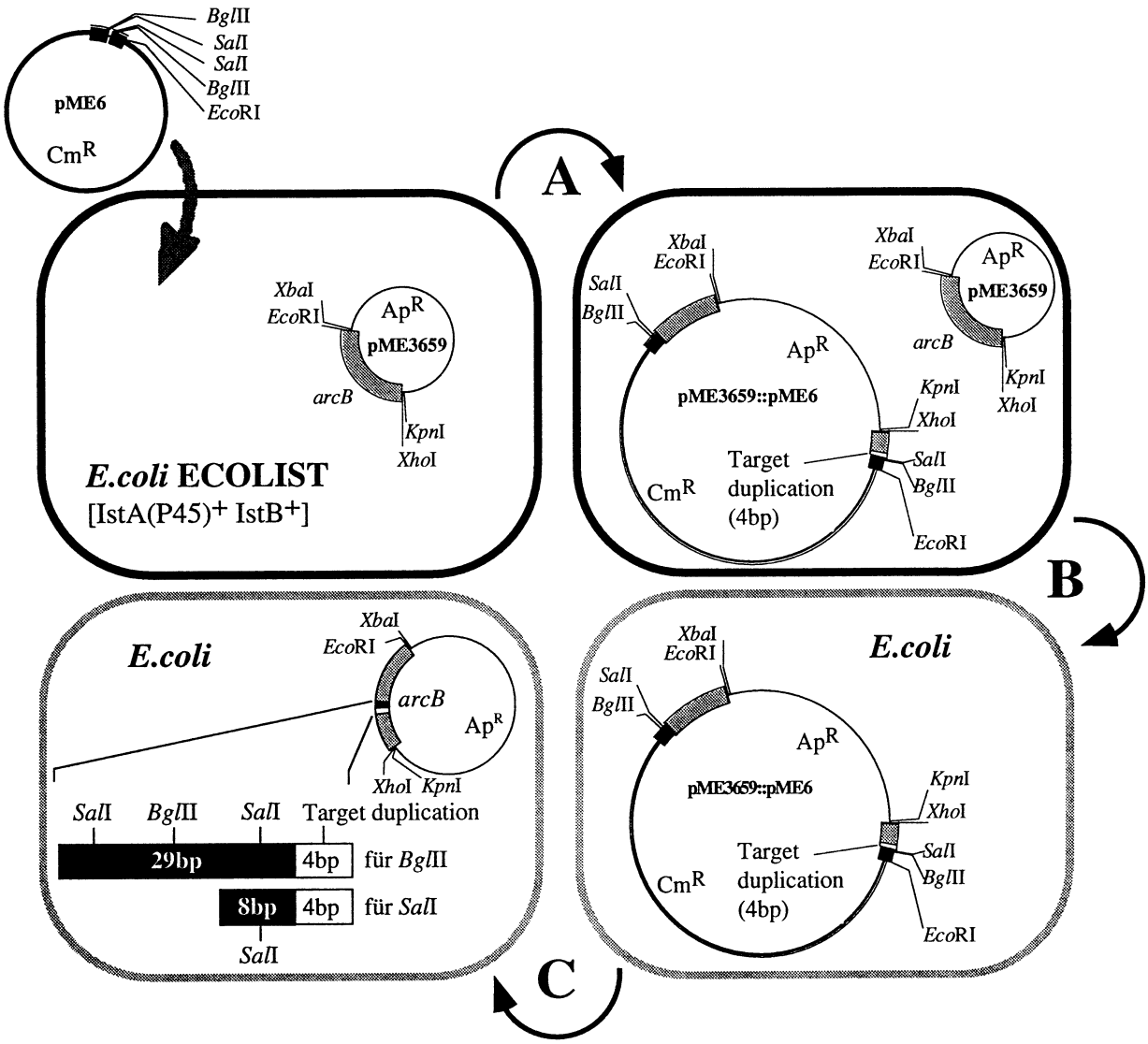

Fig. 3. Linker insertion mutagenesis procedure. (A) *E.coli* ECOLIST containing the target plasmid pME3659 is transformed with pME6, with selection for Cm resistance. After cointegrate formation, plasmid minipreparations are digested with *Kpn*I and *Xba*I, to map the pME3659::pME6 cointegrates. (B) Uncut mini-preparations of cointegrates containing pME6 in the *arcB* gene are diluted 100–1000-fold and introduced into *E.coli* RR28 by transformation, with selection for Cm resistance. (C) Purified cointegrates are cut with *Sal*I or *Bgl*II, ligated and introduced into *E.coli* RR28 by transformation, with selection for Ap resistance. An aliquot of the cointegrate preparation is used for sequencing the IS*21*–*arcB* borders. Instead of pME6, pME5 can be used, with Km selection. The 12 bp insertions obtained with *Sal*I, which consist of an 8 bp invariant core and a 4 bp target duplication, are shown in Figure 5. The 33 bp linker insertion sequence obtained with *Bgl*II (bold face) is TGT CGA CGC CAA **GAT CT**C TGG CGT CGA CAN NNN; the four variable nucleotides forming the target duplication are designated by N.

nucleotides), are an excellent substrate for the IS*21* transposition proteins IstA(P45) and IstB (Schmid *et al.*, 1998). The reactive IS*21*–IS*21* junction region is relatively small – 123 bp are sufficient (Figure 2a) – and can be genetically modified to contain unique restriction sites which are essential features of the mutagenesis method. However, the newly introduced restriction sites must not obstruct the transposition machinery. Therefore, it was important to measure the formation of cointegrates by an *in vivo* mating-out assay previously described (Schmid *et al.*, 1998). An *E.coli* host strain was used that contains three plasmids: (i) pME3913, a helper plasmid expressing the *istA*(P45) cointegrase and *istB* genes of IS*21* but not containing the inverted repeats of IS*21*; (ii) pME3915, a plasmid containing a wild-type IS*21*–IS*21* junction in which the inverted repeats are connected by a 2 bp (TA) sequence (Figure 2a); and (iii) R751, a conjugative plasmid serving as the target for pME3915. Cointegrase and the IstB protein, supplied *in trans* by pME3913, connect the IS*21*–IS*21* junction to the target. The R751::pME3915 cointegrates thus formed are recovered, after conjugation, in an *E.coli* recipient strain (Schmid *et al.*, 1998). In order to introduce point mutations into the inverted repeats, we reconstructed the IS*21*–IS*21* junction region using synthetic oligonucleotides. In this way, two artificial *Bgl*II sites were created flanking the inverted repeats in pME3923 (Figure 2a). Additionally, two *Sal*I sites were introduced into the inverted repeats in pME3924 (Figure 2a). Although pME3923 and pME3924 contained four and seven point mutations, respectively, in comparison with the original IS*21*–IS*21* junction region in pME3915, the two mutant plasmids formed cointegrates with R751 at the same frequencies (about 5×10^{-2}) as did pME3915 (Figure 2a). Thus, the IS*21*–IS*21* junction region of pME3924 could be exploited

to develop a system of transpositional linker insertion mutagenesis.

Elements of IS21-based linker insertion mutagenesis

First, the modified IS*21*–IS*21* junction of pME3924 was introduced into two suicide plasmids derived from plasmid R6K, producing pME5 (equipped with a Km resistance) and pME6 (carrying a Cm resistance) (Figure 2b). These plasmids replicate in *E.coli* strains that express the R6K π protein, the product of the *pir* replication control gene (Miller and Mekalanos, 1988). In the absence of *pir*, pME5 and pME6 fail to replicate, unless they have formed a cointegrate with another replicon. Second, the *istA*(P45) and *istB* transposition genes of IS*21* were fused to the *tac* promoter and inserted, together with the *lacI*^{Q} gene, into the chromosomal λ attachment site *attB* of *E.coli* RR28, producing strain ECOLIST (see Materials and methods for details). Thus, expression of the *istA*(P45) and *istB* genes could be induced by addition of IPTG. Third, the insertion target, the *P.aeruginosa arcB* gene encoding cOTC (Baur *et al.*, 1987), was cloned into a high-copy-number plasmid (pBluescript) which gave pME3659 (Figure 1b). This plasmid was introduced into *E.coli* ECOLIST by transformation (Figure 3). Note that the *arcB* gene and the vector do not contain any *Sal*I or *Bgl*II sites. This fact facilitates the resolution of cointegrates (see below).

Procedure of IS21-based linker insertion mutagenesis

E.coli ECOLIST/pME3659 was grown with IPTG and transformed with pME6; selection was made for the Cm resistance marker of pME6. (Details of the experimental protocol are described in Materials and methods.) Cm-resistant colonies arose at a frequency of 10^{-6}–10^{-7} per *E.coli* cell (Figure 3, A). An analogous procedure was used with pME5, except that selection was made for Km resistance. Although the suicide plasmids can insert into the chromosome, a majority of the antibiotic-resistant transformants revealed an insertion in the multi-copy plasmid: 38 out of a total of 44 pME5 insertions and 71 out of 75 pME6 insertions were found in pME3659. It may be that the recovery of plasmid cointegrates was favored by the selective antibiotic concentrations used, which were relatively high (50 μg/ml).

The pME3659::pME6 or pME3659::pME5 cointegrates were separated from the co-existing unchanged target plasmid pME3659 by a second transformation step (Figure 3, B) and analyzed for the site of insertion. The *arcB* gene (1.1 kb) was the target for an insertion of pME5 or pME6 in 35% of the cointegrates analyzed, whereas the vector moiety being larger (2.9 kb) carried an insertion in the remaining 65% of the cointegrates. The pME6 inserts in the *arcB* gene were mapped by sequencing using IS*21*-specific primers and found to be located throughout *arcB* (Figure 4a). Although some clustering of insertions was observed in *arcB*, no target sequence specificity could be detected (data not shown). Insertions were flanked by 4 bp target duplications, as expected from previous work (Reimmann *et al.*, 1989; Schmid *et al.*, 1998), except for two insertions, which were surrounded by 5 and 6 bp direct repeats, respectively.

Twenty pME3659::pME6 cointegrates having a 4 bp target duplication were restricted *in vitro* with *Sal*I, ligated and established in *E.coli* RR28 by transformation. This created tetrapeptide insertions in cOTC (Figure 3, C). Had *Bgl*II been used for resolution, inserts of 11 amino acids would have been formed (Figure 3, C).

Three important properties of the IS*21*-based linker insertion mutagenesis deserve to be pointed out. First, stop codons are not created (Figure 5), hence truncated proteins will not arise. Second, owing to the 4 bp direct repeats, no amino acid residue of the target protein is altered or deleted by an insertion. Third, the procedure carried out with *Sal*I restriction can give rise to 205 different tetrapeptide insertions (Figure 5). The 8 bp core sequence, which is derived from the IS*21* termini and carries the *Sal*I site, is palindromic. Hence the orientation of the suicide plasmid pME6 or pME5 in the cointegrate will not influence the amino acid composition of the insertion. Depending on the reading frame, three different patterns of invariant dipeptides (Cys–Arg, Ser–Thr, Val–Asp) are produced within the tetrapeptide insertions (Figure 5). An illustration of this variation is given by the insertions 2S (nucleotide position 802 in the *arcB* gene) and 6S (nucleotide position 803 in the *arcB* gene); both insertions follow Ile247 but differ in amino acid composition (Figure 4b). When cointegrates are resolved with *Bgl*II, the number of different undecapeptides is $2\times205 = 410$. The reason for this is the non-symmetrical sequence of the 29 bp core element (Figure 3), implying that the amino acid sequence of the insertion depends on the orientation of the core element.

Characterization of mutant cOTCs

The mutant *arcB* genes on pME3659 were transcribed from the vector's *lac* promoter and expressed in the OTC-negative *E.coli* mutant CM236 (*argF argI*). The recombinant cOTCs, which represented about 5% of the total soluble proteins and could all be detected by Western blotting (data not shown), were assayed for enzymatic activity. The enzymes were tested in the thermodynamically favored reverse reaction (citrulline formation from ornithine and carbamoylphosphate). The allosteric activator AMP has a similar effect on both the forward (citrulline cleavage) and the reverse reaction (Sainz *et al.*, 1998). Eleven tetrapeptide insertions in cOTC resulted in loss of activity (Figure 4b). In general, these insertions were located either close to amino acid residues which are highly conserved in OTCs (Lys31, Glu88, Asp92, Asn167, Gly180, Arg320) or in a sequence involved in the binding of the substrate carbamoylphosphate (Lys54 to Arg60) (Villeret *et al.*, 1995; Valentini *et al.*, 1996; Ha *et al.*, 1997).

Nine tetrapeptide insertions in cOTC did not abolish enzyme activity. Of these, three lying in helix H9 (designated 2S, 4S and 6S; Figure 4b) strongly reduced the affinity for ornithine in that the K_m for ornithine was increased 22–65-fold (Table II). One insertion (designated 24S), which was at the end of helix H5 (Figure 4b), was of particular interest to this study because the mutant enzyme was insensitive to the allosteric activator AMP at 10 mM. In contrast, the wild-type enzyme showed increased affinity for carbamoylphosphate in the presence of 10 mM AMP (Figure 6; Table II). In the absence of effectors, the 24S mutant enzyme had a lower affinity for carbamoylphosphate, in comparison with the wild-type cOTC (Table II). However, the mutant enzyme retained its sensitivity towards the activator P_i, although the P_i concentration required to bring about 50% of maximal activation was increased 4-fold, from 0.5 to 2 mM. These properties suggest that amino acid residues in the vicinity of the insertion site 24S may be involved in AMP recognition.

Discussion

The fact that the IS*21*–IS*21* junction region, in a stretch of 35 nucleotides, tolerated seven point mutations without any effect

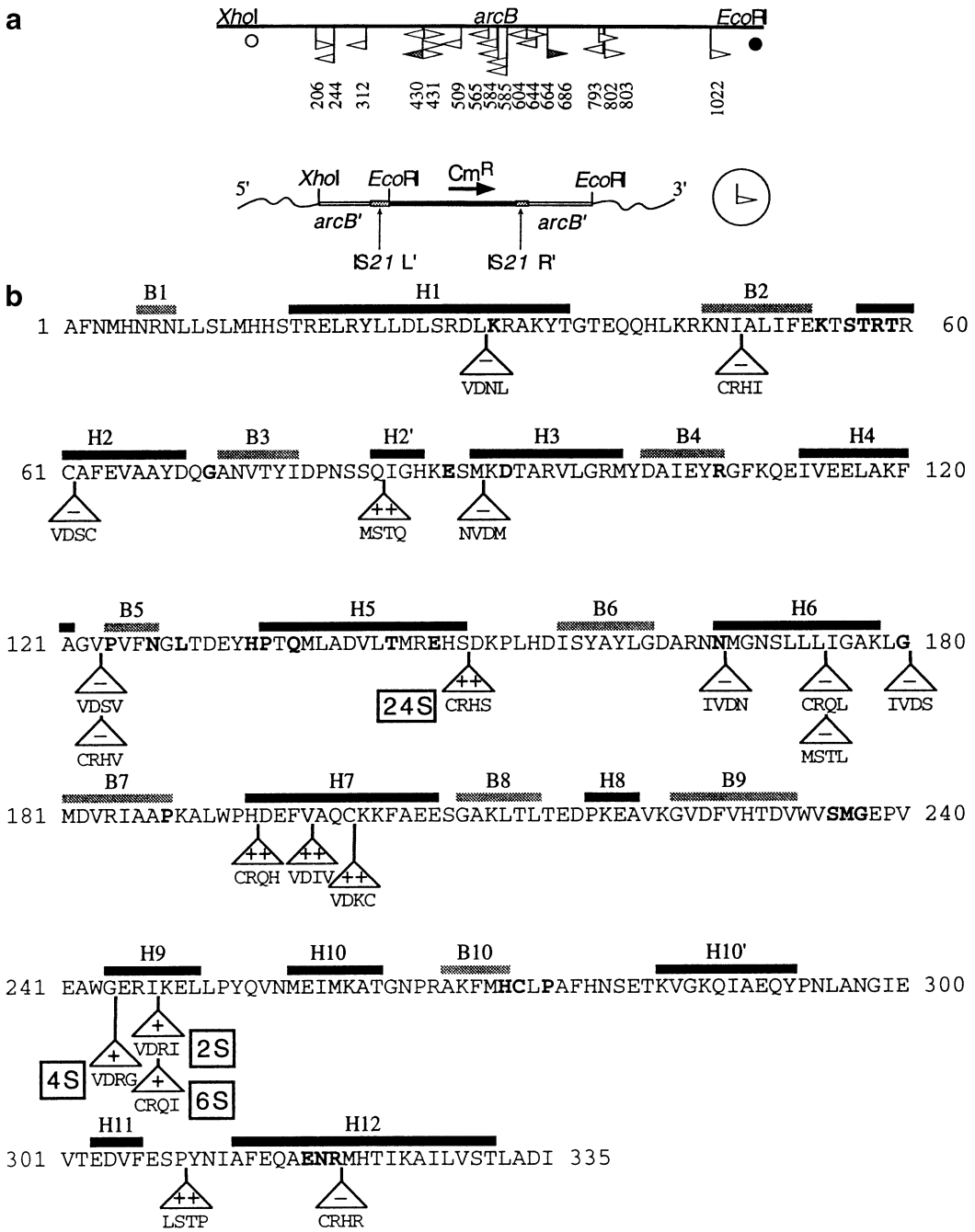

Fig. 4. Insertions of pME6 in the *arcB* gene of pME3659. (**a**) Sites of insertion in *arcB* are indicated by flags whose orientation is defined below; numbers refer to *arcB* nucleotides downstream of the unique *Xho*I site (Baur *et al.*, 1987); ○, translation start; ●, translation stop of *arcB*; open flags, 4 bp target duplication; shaded flags, 5 or 6 bp target duplication. The orientation of pME6 insertions is defined below. (**b**) Translated *arcB* sequence with α-helices H1 to H12 and β-sheets B1–B10 (Villeret *et al.*, 1995). Amino acid residues which are conserved in 25 OTCs are indicated by bold face; – in triangle, insertion causing an cOTC-negative phenotype; + in triangle, insertion causing reduced cOTC activity because of low affinity for ornithine; ++ in triangle, insertion not affecting cOTC activity. Tetramer peptides below triangles specify the sequence of the insertions.

on the frequency of cointegrate formation (Figure 2a) was surprising at first. This tolerance allowed us to devise a versatile system of linker insertion mutagenesis, by introducing into the IS*21*–IS*21* junction region two strategically placed restriction sites for each *Sal*I and *Bgl*II. Owing to these restriction sites, the bulk of the transposable element can be excised from the cointegrate, leaving behind an insertion of four codons (resolution with *Sal*I) or 11 codons (resolution with *Bgl*II). Two conceptually similar systems, which depend on MoLV and Tn*4430*, respectively, (Singh *et al.*, 1997; Hallet *et al.*, 1997), give 12- and five-codon insertions, respectively. Insertions of this length have the potential to alter functional activities of a protein such as substrate specificity (Shortle and Sondek, 1995; Hayes *et al.*, 1997). As we have shown here for cOTC, a four-codon insertion can completely abrogate an effector binding site, without compromising the catalytic activity of the enzyme. Linker insertions generated by the Tn*3*- and Tn*5*-mediated systems (Hoekstra *et al.*, 1991; Manoil and Bailey, 1997) are larger (45 and 31 codons, respectively) and potentially more disruptive.

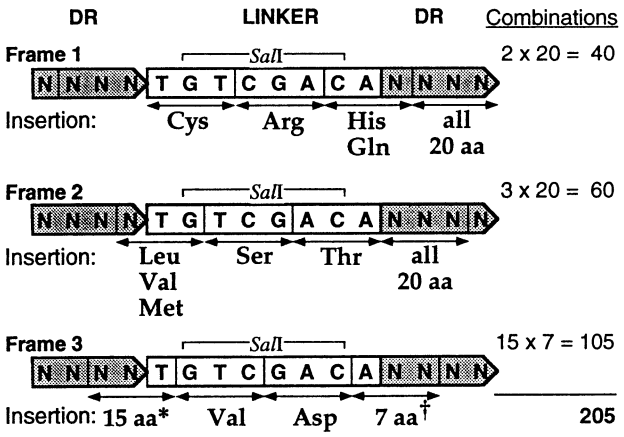

Fig. 5. Possible linker insertions after digestion with *Sal*I. DR, direct repeat, i.e. target duplication; aa, amino acid; * Phe, Ser, Tyr, Cys, Leu, Pro, His, Arg, Ile, Thr, Asn, Val, Ala, Asp or Gly; † Ile, Met, Thr, Asn, Lys, Ser or Arg.

Table II. Kinetic parameters of mutant enzymes

Enzyme	Effector	$[S]_{0.5}^{CP}$ (mM)[a]	n_H[b]	K_m^{orn} (mM)[c]
Wild-type	None	4.0 ± 0.2	4.1 ± 0.2	1.0 ± 0.05
	10 mM P_i	1.2 ± 0.1	1.0 ± 0.05	ND[d]
	10 mM AMP	1.2 ± 0.1	2.7 ± 0.1	ND[d]
2S	None	6.3 ± 0.3	4.0 ± 0.2	22 ± 2
4S	None	3.0 ± 0.2	3.0 ± 0.1	65 ± 5
6S	None	6.3 ± 0.3	5.1 ± 0.2	40 ± 2
24S	None	8.0 ± 0.5	3.1 ± 0.2	0.9 ± 0.05
	10 mM P_i	1.7 ± 0.1	1.0 ± 0.05	ND[d]
	10 mM AMP	9.0 ± 0.5	2.6 ± 0.2	ND[d]

[a]CP, carbamoylphosphate; $[S]_{0.5}$, substrate concentration required to give half-maximal velocity. The saturation curve was obtained at 10 mM L-ornithine.
[b]n_H, Hill coefficient.
[c]Apparent K_m for ornithine was determined at 10 mM carbamoylphosphate in 150 mM imidazole–HCl buffer, pH 6.8.
[d]ND, not determined.

We have previously shown that the cointegration reaction usually leads to 4 bp direct repeats, with the occasional exception of 5 bp direct repeats (Reimmann *et al.*, 1989; Schmid *et al.*, 1998). During this study, 30 insertions of pME5 and pME6 in the *arcB* gene were sequenced, including the 20 pME6 insertions analyzed in detail (Figure 4). Twenty-eight insertions were flanked by 4 bp direct repeats. One pME6 insertion had 5 bp direct repeats and one pME6 insertion 6 bp direct repeats, reflecting the fact that IS*21*, like other IS elements (Galas and Chandler, 1989), has some flexibility when interacting with a target sequence. Linker insertions flanked by atypical direct repeats change the reading frame. It is possible to avoid this complication either by filling in or by chewing back the protruding ends of the *Sal*I site. The resulting core elements of 12 or 4 bp, respectively, in combination with direct repeats of 6 or 5 bp, respectively, will give in-frame insertions.

The DNA recognition specificity of the IS*21* transposition–cointegration machinery is only partially understood. In the IS*21*–IS*21* junction donor, the terminal nucleotides 5′-CA-3′ of IS*21* are crucial for cointegrate formation (our unpublished results), whereas the mutational changes producing the *Sal*I and *Bgl*II sites (Figure 2a) do not affect the activity of

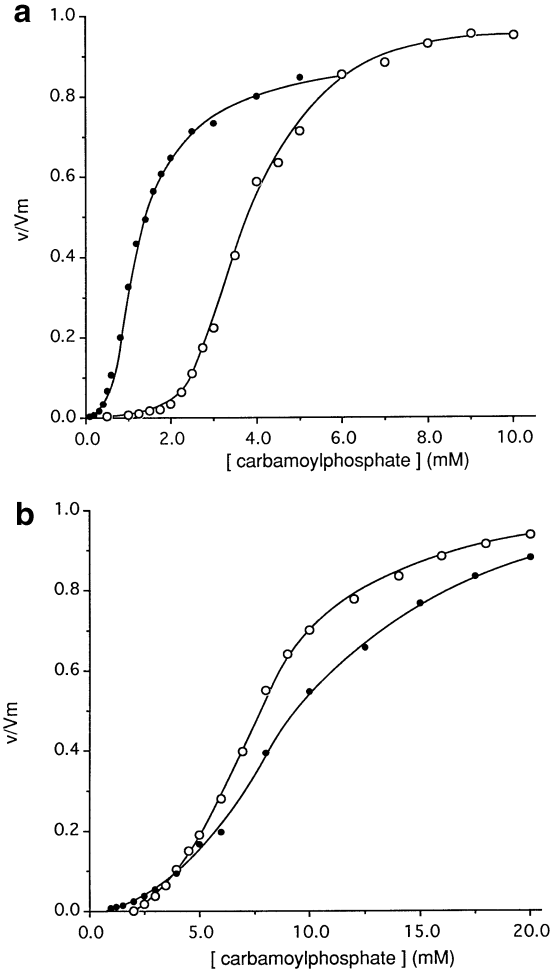

Fig. 6. Carbamoylphosphate saturation curves of (**a**) wild-type cOTC and (**b**) the AMP-insensitive mutant 24S; v/Vm, relative velocity of reaction; ○, no effector added; ●, + 10 mM AMP.

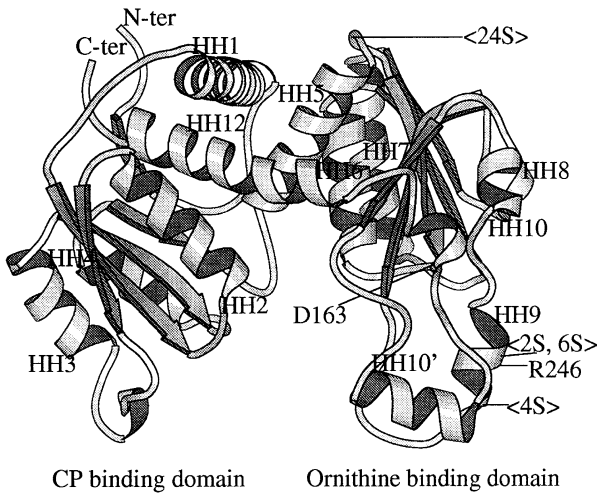

Fig. 7. Model of the cOTC monomer (Villeret *et al.*, 1995). CP, carbamoylphosphate; C-ter and N-ter, C-terminal and N-terminal residues; HH1–HH12, α-helices; <2S>, <4S>, <6S> and <24S>, sites of four-codon insertions (Figure 4b; Table II) whose properties are discussed specifically in the text; D163 and R246, amino acid residues forming an ionic interaction in the ornithine binding domain.

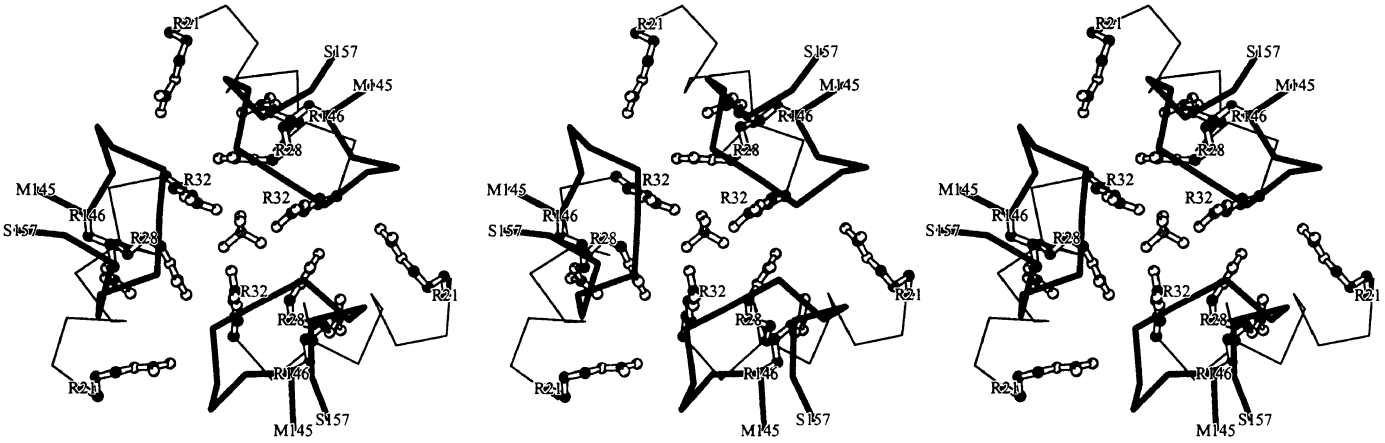

Fig. 8. Sulfate binding site at the interface between three cOTC monomers. Sulfate (tetrahedral molecule in 3-fold axis) is bound by ionic interactions with Arg32, Arg28 and Arg146. Insertion <24S> may affect these interactions, in particular with Arg146.

cointegrase. This specificity is reminiscent of that displayed by retroviral integrases. These enzymes and IS*21* cointegrase promote mechanistically similar reactions (Andrake and Skalka, 1996; Haas *et al.*, 1996). We have investigated the target specificity of the IS*21* cointegrase reaction with a number of targets, including the pBR325 *tet* gene (Reimmann *et al.*, 1989), the *P.aeruginosa arcB* gene (this study) and several other prokaryotic and eukaryotic genes (our unpublished data). No target sequence specificity has become obvious. However, some clustering of insertions in a few regions and repeated insertions into identical sites have been observed, as illustrated by Figure 4. The reasons for this behavior are not clear. Secondary structures in target DNA may play a role (Hallet *et al.*, 1994). However, most insertions appear to occur at random sites and hence are suited for linker insertion mutagenesis.

Three linker insertions in helix H9 (Figure 4b) strongly reduced the affinity of cOTC for ornithine (Table II). This effect is comparable to that seen for mutant forms of human anabolic OTC causing 'late onset' hyperammonemia. In some of these mutants, Arg277 (a conserved residue corresponding to Arg246 in cOTC) is replaced by Trp or Gln (Tuchman *et al.*, 1995; Morizono *et al.*, 1997). In both cOTC of *P.aeruginosa* and human OTC an ionic interaction between Arg246 and Asp163 (whose conserved counterpart in human OTC is Asp196) is vital to the function of the ornithine binding domain and mutational disruption of this interaction causes reduced ornithine binding, probably because Asp163 may be involved in binding the α-amino group of ornithine (Figure 7). All three tetrapeptide insertions obtained between Gly244 and Lys248 in cOTC (Figure 4b) are expected to disrupt the normal interaction between Arg246 and Asp163 and hence the elevated K_m values for ornithine can be understood.

The tetrapeptide insertion 24S giving an AMP-insensitive phenotype was mapped to the end of helix H5, near Arg146 (Figure 4b). This arginine residue, in concert with Arg32 and Arg28 of the same cOTC monomer, is located at an interface formed by three monomers belonging to different trimers (Villeret *et al.*, 1995). The nine arginine residues at this interface appear to constitute a sulfate binding site (Figure 8) and they might also interact with negatively charged activators such as inorganic phosphate or AMP (Villeret *et al.*, 1995, 1998; Tricot *et al.*, 1998). Among some 25 OTCs sequenced to date, only the cOTCs of *P.aeruginosa* and *Rhizobium etli* have these characteristic nine arginine residues; both enzymes are activated by AMP (Tricot *et al.*, 1993; D'Hooghe *et al.*, 1997). The crystal structure of the cOTC E105G mutant (Villeret *et al.*, 1995, 1998) reveals that the loops formed by amino acid residues 150–157 between helix H5 and β-sheet B6 (Figures 4b and 8) partially cover the arginine rings, potentially limiting the access of the allosteric activators. Although the structural changes taking place during the allosteric transition are unknown, the trimer–trimer interface area is likely to be involved in modulating the activity of the catalytic trimers. The insertion of the 24S linker Cys–Arg–His–Ser may indeed lower the accessibility to the arginine rings and thereby interfere with AMP binding in either the allosteric T or R state. These data can now serve as a guide to devise experiments defining more precisely the residues involved in AMP binding.

The AMP-insensitive mutant and the mutants exhibing low ornithine affinity clearly illustrate the potential of the IS*21*-based linker insertion mutagenesis. Random insertion and variable composition of linkers give rise to a multitude of protein variants. Moreover, it is pertinent to point out that the *Sal*I site in the linker insertions can be exploited to introduce new genetic information, e.g. specifying a protease cleavage site or an epitope.

Acknowledgements

We thank our colleagues who provided plasmids and strains and C.Reimmann and V.Krishnapillai for valuable discussions. This work was supported by the Swiss National Foundation for Scientific Research (project 31-45896.95), the Roche Research Foundation, the Internationale Brachet Stiftung and the Belgian Fund for Joint Basic Research.

References

Andrake,M.D. and Skalka,A.M. (1996) *J. Biol. Chem.*, **271**, 19633–19636.
Barany,F. (1985) *Gene*, **37**, 111–123.
Baur,H., Stalon,V., Falmagne,P., Luethi,E. and Haas,D. (1987) *Eur. J. Biochem.*, **166**, 111–117.
Chang,A.C.Y. and Cohen,S.N. (1978) *J. Bacteriol.*, **134**, 1141–1156.
Del Sal,G., Manfioletti,G. and Schneider,C. (1988) *Nucleic Acids Res.*, **16**, 9878.
D'Hooghe,I., Vander Wauven,C., Michiels,J., Tricot,C., de Wilde,P., Vanderleyden,J. and Stalon,V. (1997) *J. Bacteriol.*, **179**, 7403–7409.
Diederich,L., Rasmussen,L.J. and Messer,W. (1992) *Plasmid*, **28**, 14–24.
Dykxhoorn,D.M., St. Pierre,R., Van Ham,O. and Linn,T. (1997) *Nucleic Acids Res.*, **25**, 4209–4218.

Galas,D.J. and Chandler,M. (1989) In Berg, D.E. and Howe,M.M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 109–162.
Goff,S.P. and Prasad,V.R. (1991) *Methods Enzymol.*, **208**, 586–603.
Graf,R. and Schachman,H.K. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 11591–11596.
Ha,Y., McCann,M.T., Tuchman,M. and Allewell,N.M. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 9550–9555.
Haas,D., Berger,B., Schmid,S., Seitz,T. and Reimmann,C. (1996) In Nakazawa,T., Furukawa,K., Haas,D. and Silver,S. (eds), *Molecular Biology of Pseudomonads*. ASM Press, Washington, DC, pp. 238–249.
Hallet,B., Rezsöhazy,R., Mahillon,J. and Delcour,J. (1994) *Mol. Microbiol.*, **14**, 131–139.
Hallet,B., Sherratt,D.J. and Hayes,F. (1997). *Nucleic Acids Res.*, **25**, 1866–1867.
Hayes,F., Hallet,B. and Cao,Y. (1997) *J. Biol. Chem.*, **272**, 28833–28836.
Hennecke,H., Günther,I. and Binder,F. (1982) *Gene*, **19**, 231–234.
Herrero,M., de Lorenzo,V. and Timmis,K.N. (1990) *J. Bacteriol.*, **172**, 6557–6567.
Hirsch,P.R. and Beringer,J.E. (1984) *Plasmid*, **12**, 139–141.
Hoekstra,M.F., Burbee,D., Singer,J., Mull,E., Chiao,E. and Heffron,F. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 5457–5461.
Jeenes,D.J., Soldati,L., Baur,H., Watson,J.M., Mercenier,A., Reimmann,C., Leisinger,T. and Haas,D. (1986) *Mol. Gen. Genet.*, **203**, 421–429.
Kegler-Ebo,D.M., Docktor,C.M. and DiMaio,D. (1994) *Nucleic Acids Res.*, **22**, 1593–1599.
Manoil,C. and Bailey,J. (1997) *J. Mol. Biol.*, **267**, 250–263.
Marcq,S., Diaz-Ruano,A., Charlier,P., Dideberg,O., Tricot,C., Piérard,A. and Stalon,V. (1991) *J. Mol. Biol.*, **220**, 9–12.
Miller,V.L. and Mekalanos,J.J. (1988) *J. Bacteriol.*, **170**, 2575–2583.
Morizono,H., Listrom,C.D., Rajagopal,B.S., Aoyagi,M., McCann,M.T., Allewell,N.M. and Tuchman,M. (1997) *Hum. Mol. Genet.*, **6**, 963–968.
Nelson,B.D., Manoil,C. and Traxler,B. (1997) *J. Bacteriol.*, **179**, 3721–3728.
Nguyen,V.T., Tricot,C., Stalon,V., Dideberg,O., Villeret,V. and Haas,D. (1994) *FEMS Microbiol. Lett.*, **124**, 411–418.
Nguyen,V.T., Baker,D.P., Tricot,C., Baur,H., Villeret,V., Dideberg,O., Gigot,D., Stalon,V. and Haas,D. (1996) *Eur. J. Biochem.*, **236**, 283–293.
Penfold,R.J. and Pemberton,J.M. (1992) *Gene*, **118**, 145–146.
Prentki,P. and Krisch,H.M. (1984) *Gene*, **9**, 303–313.
Reimmann,C. and Haas,D. (1990) *EMBO J.*, **9**, 4055–4063.
Reimmann,C., Moore,R., Little,S., Savioz,A., Willetts,N.S. and Haas,D. (1989) *Mol. Gen. Genet.*, **215**, 416–424.
Sambrook,J. Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sainz,G., Tricot,C., Foray,M.-F., Marion,D., Dideberg,O. and Stalon,V. (1998) *Eur. J. Biochem.*, **251**, 528–533.
Schmid,S., Seitz,T. and Haas,D. (1998) *J. Mol. Biol.*, **282**, 571–583.
Singh,I R., Crowley,R.A. and Brown,P.O. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 1304–1309.
Shortle,D. and Sondek,J. (1995) *Curr. Opin. Biotechnol.*, **6**, 387–393.
Tricot,C., Nguyen,V.T. and Stalon,V. (1993) *Eur. J. Biochem.*, **215**, 833–839.
Tricot,C., Villeret,V., Sainz,G., Dideberg,O. and Stalon,V. (1998) *J. Mol. Biol.*, **283**, 695–704.
Tsang,T.C. and Bentley,D.R. (1988) *Nucleic Acids Res.*, **16**, 6258.
Tuchman,M., Morizono,H., Reish,O., Yuan,X. and Allewell,N.M. (1995) *J. Med. Genet.*, **32**, 680–688.
Ubben,D. and Schmitt,R. (1986) *Gene*, **41**, 145–152.
Valentini,G., De Gregorio,A., Di Salvo,C., Grimm,R., Bellocco,E., Cuzzocrea,G. and Iadarola,P. (1996) *Eur. J. Biochem.*, **239**, 397–402.
Vander Wauven,C., Piérard,A., Kley-Raymann,M. and Haas,D. (1984) *J. Bacteriol.*, **160**, 928–934.
Villeret,V., Tricot,C., Stalon,V. and Dideberg,O. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 10762–10766.
Villeret,V., Clantin,B., Tricot,C., Legrain,C., Roovers,M., Stalon,V., Glansdorff,N. and Van Beeumen,J. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 2801–2806.

Received August 31, 1999; revised December 20, 1999; accepted January 6, 2000