Genomic distribution and functions of uptake signal sequences in *Actinobacillus actinomycetemcomitans*

Ying Wang,1† Joshua Orvis,2 David Dyer2 and Casey Chen1

1Division of Primary Oral Health Care, University of Southern California School of Dentistry, 925 W 34th Street, Room 4107, Los Angeles, CA 90089-0641, USA
2Department of Microbiology and Immunology, University of Oklahoma, Oklahoma City, OK 73190, USA

Actinobacillus actinomycetemcomitans is naturally competent for transformation, with a transformation system similar to that of *Haemophilus influenzae* that preferentially takes up DNA bearing uptake signal sequences (USS) with the same 9-base USS core. This study examined the function of the extended 29-base USS, which comprises a highly conserved 1st region (containing the 9-base core) and 2nd and 3rd semi-conserved AT-rich regions, in transformation of *A. actinomycetemcomitans*. Transformation frequency was not affected by either location (in middle or at 5’ end) or quantity (one or two) of USS in donor DNA. Relative transformation efficiencies (in comparison to the positive control) were 28–67 % for linear DNA with single-base mutations in the USS 1st region, and 47 % and 73 %, respectively, for linear DNA with USS that contained either a non-consensus 2nd or a non-consensus 3rd region. Plasmids with a stand-alone 1st or a stand-alone 2nd–3rd region exhibited 21 % and 6 % relative transformation efficiencies, respectively. It was also noted that *A. actinomycetemcomitans* and *H. influenzae* were similar in the frequencies and distribution patterns of USS in their genomes. In conclusion, all three regions of the extended 29-base USS are required for optimum transformation in *A. actinomycetemcomitans*.

INTRODUCTION

Natural transformation is a process by which bacteria take up extracellular DNA and incorporate it into the host genome by homologous recombination. This may result in the expression of a new genotype (Dubnau, 1999; Lorenz & Wackernagel, 1994). Bacteria are the only organisms known to have this capability. Natural transformation can be found in phylogenetically distant bacteria, including both Gram-positive and Gram-negative bacteria.

The mechanisms used by competent bacterial species are also distinct. Several transformation systems have been well characterized, including those of Gram-positive *Streptococcus–Bacillus*, Gram-negative *Haemophilus–Neisseria* (Dubnau, 1999; Lorenz & Wackernagel, 1994; Smeets & Kusters, 2002) and type IV secretion-dependent *Helicobacter pylori* (Smeets & Kusters, 2002).

In competent bacterial species employing the *Haemophilus–Neisseria* system, the transformation efficiency is greatly enhanced if the donor DNA contains specific oligonucleotides designated uptake signal sequences (USS). Furthermore, USS sites are highly represented in the genomes of these bacteria. The USS of *Haemophilus influenzae* have been defined experimentally as both a 9-base core of 5’-AAGTGCGGT and its complementary sequence. Genomic sequence analysis of the USS in the Rd strain of *H. influenzae* revealed an extended 29-base consensus sequence (Smith et al., 1995, 1999). For the purpose of this study this extended USS will be described as comprising 3 regions: the 1st region (10-base, 5’-AAGTGCGGT), which contains a semi-conserved adenosine followed by the originally defined 9-base USS core, followed by the 2nd (9-base, 5’-nrwwwwwn), n: any nucleotide, r: A or G, w: A or T) and 3rd (10-base, 5’-nnnrwwwww) semi-conserved AT-rich regions. The role of these AT-rich regions in facilitating transformation remains to be determined.

The Gram-negative, capnophilic bacterium *Actinobacillus actinomycetemcomitans*, a member of the family *Pasteurellaceae*, is implicated as a major causative agent of localized aggressive periodontitis and other forms of periodontitis (Asikainen & Chen, 1999; Slots, 1999). *A. actinomycetemcomitans* is naturally competent for transformation (Fujise et al., 2004; Tonjum et al., 1990; Wang et al., 2002, 2003). The transformation system of *A. actinomycetemcomitans* is remarkably similar to that of *H. influenzae*.
A. actinomycetemcomitans also displays preferential uptake of donor DNA with the same 9-base USS core as *H. influenzae*.

The objective of this study was to further examine the requirements for the extended 29-base USS, in particular the 2nd and the 3rd semi-conserved regions, in transformation of *A. actinomycetemcomitans*. Transformation frequencies and relative transformation efficiencies were determined. The frequency and distribution pattern of USS sites in the genome of *A. actinomycetemcomitans* were also examined and compared to those in *H. influenzae*.

**METHODS**

**Bacteria and culture conditions.** Not all *A. actinomycetemcomitans* strains are competent for transformation (Fujise et al., 2004). A naturally transformable *A. actinomycetemcomitans* strain D7S-smooth was selected for this study (Wang et al., 2002). The strain was a nonfimbriated variant of strain D7S that was recovered from a patient with aggressive periodontitis. The nonfimbriation was the result of a spontaneous point mutation in the promoter region of the fimbrina/Tad locus (Wang et al., 2003). *A. actinomycetemcomitans* strain D7S was a nonfimbriated variant of strain D7S that was recovered from a patient with aggressive periodontitis. The strain was routinely grown in sTSB medium [3% tryptase soy broth (TSB), 0.3% yeast extract, 1.5% agar and 5% horse serum], or mTSB broth (3% TSB and 0.6% yeast extract) at 37°C in 5% CO₂. *Escherichia coli* plasmids were propagated in *E. coli* DH5α by standard methods (Sambrook et al., 1989). When required, the media were supplemented with spectinomycin (Spe, 50 μg ml⁻¹).

**Plasmids.** The recombinant plasmid *pDpilB870-770* was constructed previously (Wang et al., 2003). It contains a cloned recombinant DNA (*pilA-SpeR-pilC*) that was derived from the pilABC locus of *A. actinomycetemcomitans* strain D7S. The internal 846 bp of *pilB* was replaced with a 1.2 kb SpeR cassette leaving fragments of 870 bp (*pilD*) and 770 bp (*pilC*) flanking the cassette. There are two naturally occurring USS sites in *pDpilB870-770* (Figs 1 and 2). The first USS site is located upstream of *pilA*. The second USS site is located at the 5' end of *pilC* next to the SpeR cassette. The recombinant plasmid *pDpilB870-694* (Figs 2 and 3) is a derivative of *pDpilB870-770*. It was constructed by removing a 76 bp *SalI–MluI* fragment containing the USS site nearest the SpeR cassette from *pDpilB870-770*.

![Fig. 1. Transformation frequencies of linear donor DNA fragments with varying lengths of homologous flanking regions in competent *A. actinomycetemcomitans* strain D7S-smooth. The donor DNA fragments were generated by PCR amplification using *pDpilB870-770* as the template. The arrows indicate the primers and their locations relative to the template. The resultant amplicons contain a SpeR marker flanked by an invariant left arm and a right arm of varying length. All transformation assays were performed at least three times. By ANOVA with Tukey-HSD multiple range test (*P*<0.05), significantly different from: **500 bp and 766 bp, ***766 bp. No significant differences were detected among 160 bp, 250 bp or 354 bp.)

![Fig. 2. Transformation frequencies of linear donor DNA with one or two USS sites at different locations in competent *A. actinomycetemcomitans* strain D7S-smooth. The templates for generating donor amplicon DNA and primer locations (arrows) are depicted at the top. Donor DNA fragments with no USS or a USS at one end were generated by PCR amplification with the template *pDpilB870-694*. The primers used were an invertible reverse primer PilC-K and the forward primers PilA and USS-PilA respectively. Donor fragments with a USS in the middle or 2 USS were generated using the template *pDpilB870-770* with reverse primer PilC-K and forward primers PilA and USS-PilA respectively. All transformation assays were performed at least three times. By ANOVA with Tukey-HSD multiple range test: *significantly different from the other three donor DNA (P<0.05); NS, no significant differences (P>0.05).
The recombinant plasmid pD-HM was constructed by first replacing hmsFR of the hmsHFRD gene cluster with a 1-2 kb SpeI cassette in A. actinomycetemcomitans strain D7S-smooth by a previously described method (Wang et al., 2003). The genomic DNA of the resultant ΔhmsFR mutant was used as a template to amplify the SpeI marker and its flanking sequences with primers HMpre1, 5′-CGCGGTATG-TGACATCGGA (710 bases upstream of SpeI), and HM-USS, 5′-GGAAAATCCAAATGTTTTTTA (407 bases downstream of SpeI). The 2-3 kb amplicon was then cloned in pBluescript KS (Stratagene) at the EcoRI site to generate pD-HM.

Several recombinant plasmids were constructed and used directly as donor DNA in transformation assays. These plasmids were constructed by cloning into pBluescript KS the PCR products amplified from pDpilB870-694 using appropriate primers that incorporated base substitutions in USS regions (see Fig. 4 for primer sequences).

PCR. Recombinant plasmids were used as templates for PCR amplification to generate DNA for transformation assays. The primer sequences are either listed in Table 1, or provided with the results (Figs 3 and 5). Each PCR reaction of a 60 μl mixture contained 25–50 ng pre-boiled plasmid, 30 pmol each of primers, 1-6 mM MgCl₂, and 3 units Taq polymerase (Promega), and was amplified under the following conditions: 30 cycles at 94°C for 30s, 56°C for 30s, 72°C for 3-5 min, and a final extension of 5 min at 72°C. After amplification, the DNA was digested with DpnI (New England Biolabs) at 37°C for 1 h to remove template DNA. Our preliminary studies have shown that digestion of amplicons with DpnI removes methylated DNA and destroys 80–90% of the transforming activity of the template DNA, but has no apparent effect on the transforming ability of the unmethylated amplicons. If the PCR products exhibited a single strong band on an agarose gel, the DNAs were directly used in transformation assays. Occasionally, when the qualities of amplicons were not satisfactory, the PCR reactions were repeated with slightly modified conditions, or PCR products were purified through columns (Qiagen) and quantified before use in a transformation.

Transformation assays. An agar-based transformation assay was performed as previously described (Wang et al., 2002). Briefly, bacteria were grown on sTSB agar for 20 h, collected and suspended in TSB to ~5 × 10⁷ c.f.u. ml⁻¹. The bacterial suspension (20 μl) was spotted on a prewarmed sTSB plate and incubated for 2 h until cells reached optimal competency. Ten microliters of donor DNA (0.2–0.5 μg in TE buffer) was added to the recipient bacteria and mixed with an inoculation loop. The bacteria were further incubated for 5 h and then plated on the selective medium (sTSB agar containing 50 μg spectinomycin ml⁻¹) and incubated for 2–3 days to enumerate transformants. Transforming DNA was always used at more than the saturation level (~0.1 μg DNA). The results are reported as transformation frequency (transformants per c.f.u.) or as relative transformation efficiency.

Sequence analysis. Computational analyses were performed on the genome of A. actinomycetemcomitans strain HK1651. Detection and distribution of USS sites were determined using a series of Perl scripts, which have been compiled into a publicly available, web-based application called SeqSeek (unpublished), which is available at http://microgen.ouhsc.edu/software/seqseek.htm. Utilizing the 9-base USS core sequence as a query, SeqSeek displays count data for the number of USS sites found on each strand, along with the distribution of USS sites in intragenic and intergenic spaces. SeqSeek also reports large regions of the genome that are devoid of USS sites, listing the genes contained within each of these regions. Repeat regions were discovered using the European Molecular Biology Open Software Suite (EMBOSS) package of sequence analysis tools. Specifically, inversed and etandem were used to find inverted and tandem repeats, respectively.
Statistical analysis. The relative differences in the transformation efficiencies of donor DNA were evaluated by ANOVA and Tukey-HSD multiple range test.

RESULTS

Parameters of transformation with linear donor DNA

We first determined the effects on transformation frequency of (i) the lengths of homologous flanking regions (Fig. 1) and (ii) the number of USS sites of the donor DNA (Fig. 2). The transformation frequency for donor DNA with a right arm 766 bp in length was ~2.610^23 per cell; a frequency that was similar to the optimum transformation frequency in standard transformation assays (Fujise et al., 2004; Wang et al., 2002, 2003). There was a trend of a gradual decrease of transformation frequency with decreasing length of homologous regions from ~66% (500 bp) to ~5% (160 bp) of the transformation frequency with the full-length DNA.

Fig. 2 (bottom) shows that transformation frequencies were similar with donor DNA containing one or two USS sites, and with donor DNA having a USS site in the middle or at the 5' end of the DNA (no significant differences; Tukey-HSD test, \( P > 0.05 \)).

Transformation by linear DNA with base substitutions in the 1st region of the USS

Fig. 3 shows the results of the transformation assays with amplicons (~2-7 kb) that contained a full USS site or a USS site with either a single-base substitution or missing the first 5 bases. The amplicons with single-site transition mutations of the 1st region retained 28–67% of the relative transformation efficiency of the positive control (Tukey-HSD test, \( P < 0.05 \)). The deletion of the first 5 bases in the 1st region reduced the relative transformation efficiency to 1%.

Transformation with recombinant plasmids

The use of linear DNA for transformation as described above did not address the question whether optimum transformation of DNA requires the presence of DNA 5’ to the USS in the donor DNA. Therefore, we designed a series of recombinant plasmids as donor DNA for transformation assays. Fig. 4 shows that the relative transformation efficiencies of plasmids pUS54, pUS88A and pUS23 (all with a base substitution in the 1st region) were 45–58% and were significantly different from both the positive control pUS1 and the negative control pUS0 (\( P < 0.05 \); Tukey-HSD test). The relative transformation efficiencies of the 1st region of pUS51 retained 28–67% of the relative transformation efficiency of the positive control (Tukey-HSD test, \( P < 0.05 \)). The deletion of the first 5 bases in the 1st region reduced the relative transformation efficiency to 1%.

Table 1. Sequences of primers used in Figs 1 and 2

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PilA-F</td>
<td>GGAATTCCGAATTTAATGATCGTCATCG</td>
</tr>
<tr>
<td>PilB-R2</td>
<td>GCCCTGCGGATGTTGAAT</td>
</tr>
<tr>
<td>PilB-R3</td>
<td>TAATACCGTGGATTTTCCG</td>
</tr>
<tr>
<td>PilC-D</td>
<td>GAGTGACCGTGCCATCACCTCCGTTTC</td>
</tr>
<tr>
<td>PilC-500</td>
<td>GCTGTTTTCACGGCGACAG</td>
</tr>
<tr>
<td>PilC-K</td>
<td>CGGGTGACCGTGGATTTTCCG</td>
</tr>
<tr>
<td>USS-piLA</td>
<td>AAAAGTGCGGTGGAATTTTAC</td>
</tr>
<tr>
<td>PilA</td>
<td>GCGGTGGAATTTCAGCGTGTG</td>
</tr>
</tbody>
</table>
pUSS-core (containing a perfect 1st region but with non-consensus 2nd and 3rd regions), pRegions23 (with a non-consensus 1st region and intact 2nd and 3rd regions) and the negative control pUSS-0 (without USS) were severely reduced: 21%, 6%, and 6%, respectively. The results again suggested that single-base mutations in the 1st region affected transformation efficiency, even in donor DNA with an extended DNA region 5' to the USS site. The results were also consistent with our experience that the 1st region was more important than the 2nd–3rd region in transformation.

Involvement of the 2nd and 3rd regions of the USS in transformation

The results from the above experiments did not distinguish between the individual contributions from the 2nd or 3rd regions of USS in transformation. We took advantage of a native USS site occurring at the 3' end of the hmsD gene to design amplicon donor DNA with a perfect USS, a non-consensus 2nd region, a non-consensus 3rd region, or a truncated 3rd region to test their relative transformation efficiencies (Fig. 5). The first three amplicons were identical in length but contained either a perfect USS, a USS with a non-consensus 2nd region, or a USS with a non-consensus 3rd region; all three amplicons contained a tail of 3 nucleotides 3' to the USS. The remaining amplicon (HM-USS-7) contained a truncated USS missing the last 6 nucleotides. The three regions of the native USS near hmsD and in the primers are indicated with boxes. All experiments were repeated at least three times. By ANOVA with Tukey-HSD multiple range test (P<0.01), significantly different from: *HM-USS-perfect, **HM-USS-7, #HM-USS-2'.

Fig. 5. Relative transformation efficiencies of PCR amplicon donor DNA with an intact USS, a USS with non-consensus 2nd region, a USS with a non-consensus 3rd region, or a truncated USS. The recombinant plasmid pD-HM was used as the template for PCR amplification with the forward primer HMpre1, and one of the four reverse primers HM-USS-perfect (intact USS), HM-USS-2' (non-consensus 2nd region), HM-USS-3' (non-consensus 3rd region), or HM-USS-7 (truncated USS). The first 3 amplicons were identical in length but contained either a perfect USS, a USS with a non-consensus 2nd region, or a USS with a non-consensus 3rd region; all three amplicons contained a tail of 3 nucleotides 3' to the USS. The remaining amplicon (HM-USS-7) contained a truncated USS missing the last 6 nucleotides. The three regions of the native USS near hmsD and in the primers are indicated with boxes. All experiments were repeated at least three times. By ANOVA with Tukey-HSD multiple range test (P<0.01), significantly different from: *HM-USS-perfect, **HM-USS-7, #HM-USS-2'.

The 3rd region in transformation. Truncation of the 3rd region of the USS severely reduced the transformation efficiency to ~3%, suggesting that the presence of DNA beyond the 2nd region (irrespective of its sequence) was also critical for transformation.

Frequency of USS sites in the genome of strain HK1651

We searched the genome of A. actinomycetemcomitans strain HK1651 for the highly conserved 9-base core sequence of the USS and found a total of 1759 copies, identified as 881 copies on one strand and 878 copies on the complementary strand. Fewer than 14 copies of the USS would be expected by random chance. The frequency of the USS in A. actinomycetemcomitans is similar to that in H. influenzae (both occur at 0-8 copy of USS per kb genome). We also confirmed the presence of the extended 29-base full USS in A. actinomycetemcomitans, as noted previously (Wang et al., 2002).

Distribution patterns of USS sites in the genome of strain HK1651

The location of USS sites in relation to ORFs in the genome of HK1651 was further examined. There were 1402 copies of the 9-base USS core sequence located within coding regions, while the remaining 357 copies were found to be intergenic. The large-scale distribution pattern of the USS appears to be random in the genome. There are several large genomic regions (5 kb or greater) without USS in the HK1651 genome.
genome. Some USS-free regions were found to harbour rRNA operons and loci for ribosomal proteins (Smith et al., 1999) and genomic islands (Chen et al., 2005) (see also http://www.oralgen.lanl.gov/), whereas others contain putative virulence factors of A. actinomycetemcomitans [e.g. matrix binding protein EmaA (Mintz, 2004), cytolethal distending toxin (Mayer et al., 1999; Shenker et al., 2000; Sugai et al., 1998), fimbria/Tad locus (Haase et al., 1999; Inouye et al., 1990; Ishihara et al., 1997; Kachlany et al., 2000; Planet et al., 2003) and leukotoxin (Kologrubetz et al., 1989; Kraig et al., 1990; Lally et al., 1989)].

The distribution pattern of USS within a close distance to a second USS was not random. We found 162 pairs of USS sites separated by a distance of 35 bp or less. Among these sites, 160 are inverted repeats (96 +/ − pairs, and 64 −/+ pairs), and 2 are direct repeats. The 160 inverted repeats have the potential to form stem–loop structures when transcribed into RNA. Of these, 92 (53 +/ − pairs and 39 −/+ pairs) are located completely in intergenic regions. A high number of these USS sites were observed to cluster within ~50 bp of the 3′ end of ORFs. No such clustering in relationship to the 5′ end of the coding regions was found. The remaining 68 inverted repeats (43 +/− pairs and 25 −/+ pairs) were located within coding regions. The potential regulatory functions of these USS sites remain to be determined.

DISCUSSION

The presence of USS was noted in the genome of A. actinomycetemcomitans by Thompson et al. (1999). The function of A. actinomycetemcomitans USS may be inferred from studies of transformation of H. influenzae. However, while the 1st USS region in the donor DNA was required for DNA uptake in H. influenzae, the involvement of the 2nd and the 3rd USS regions in transformation has not been fully tested (Danner et al., 1980).

In this study, we examined the involvement of three regions of the USS in the transformation of A. actinomycetemcomitans. We also examined the frequency and distribution patterns of USS in the A. actinomycetemcomitans genome. The results showed that all three USS regions were required to achieve optimum transformation efficiency. The relative contributions to transformation among individual USS regions were 1st region > 2nd region > 3rd region. We further noted that the frequency and distribution patterns of USS were similar in A. actinomycetemcomitans and H. influenzae.

Several large genomic regions (>5 kb) of A. actinomycetemcomitans strain HK1651 were found to contain no USS sites. Some of these regions also exhibited low G+C content. It is worthy of note that several of these USS-free regions contain putative virulence genes of A. actinomycetemcomitans. Perhaps these USS-free regions were DNA blocks acquired by horizontal gene transfer and constitute part of the flexible gene pool that enhances the fitness or virulence of bacteria (Hacker et al., 1997; Hacker & Carniel, 2001).

The USS of A. actinomycetemcomitans is prevalent in the genomes of some Pasteurellaceae species. Albritton et al. (1984, 1986) observed that the genomic DNA from Haemophilus parainfluenzae, Haemophilus aphrophilus, Haemophilus parahaemolyticus, Pasteurella multocida and A. actinomycetemcomitans competed for homospecific transformation in H. influenzae. Hong & Dewhirst (2002) found preliminary evidence for the frequent occurrences of the USS in Haemophilus parahaemolyticus, H. parainfluenzae, Actinobacillus suis, P. multocida, Pasteurella canis and several additional Pasteurella species.

The evolutionary basis for the high prevalence of USS in the genomes of A. actinomycetemcomitans and H. influenzae is not completely understood. Since the identical USS core sequence is found at high frequencies in the genomes of diverse Pasteurellaceae species, USS may not function as a species barrier to prevent the uptake of non-self DNA as suggested previously. Bakkali et al. (2004) examined the occurrence of USS sites in P. multocida, H. influenzae and A. actinomycetemcomitans. An identical 9-base USS sequence was identified in 927, 1205 and 1760 copies, respectively in these three species. Moreover, the locations of USS sites were conserved in homologous genes in the three species. It was postulated that the USS were selected for and maintained by a biased-DNA-uptake system (Bakkali et al., 2004). It remains to be determined whether the shared USS promote genetic exchange among these bacterial species.

ACKNOWLEDGEMENTS

This research was supported by NIH/NIDCR grant R01 DE12212 and NIH/NCRR grant P20 RR016478. We thank Karen L. McKeown, PhD, for helpful editorial comments on the manuscript.

REFERENCES


Hong, B. & Dewhirst, F. E. (2002). *Haemophilus influenzae* DNA uptake signal sequences are present in *Actinobacillus actino-
mycetemcomitans* and most other species in the family Pasteurellaceae. In *The IADR/AADR/CADR 80th General Session (March 6–9, 2002)*. *J Dent Res* 81, special issue A, abstract 1448.


Mintz, K. P. (2004). Identification of an extracellular matrix protein adhesin, EmaA, which mediates the adhesion of *Actinobacillus actino-
mycetemcomitans* to collagen. *Microbiology* 150, 2677–2688.


