Site-specific Tn7 transposition into the human genome

Prasad N. Kuduvalli, Rupak Mitra and Nancy L. Craig*

The Howard Hughes Medical Institute, Department of Molecular Biology and Genetics,
Johns Hopkins School of Medicine, Baltimore, MD 21205, USA

Received October 25, 2004; Revised and Accepted January 16, 2005

ABSTRACT

The bacterial transposon, Tn7, inserts into a single site in the Escherichia coli chromosome termed attTn7 via the sequence-specific DNA binding of the target selector protein, TnsD. The target DNA sequence required for Tn7 transposition is located within the C-terminus of the glucosamine synthetase (glmS) gene, which is an essential, highly conserved gene found ubiquitously from bacteria to humans. Here, we show that Tn7 can transpose in vitro adjacent to two potential targets in the human genome: the gfpt-1 and gfpt-2 sequences, the human analogs of glmS. The frequency of transposition adjacent to the human gfpt-1 target is comparable with the E.coli glmS target; the human gfpt-2 target shows reduced transposition. The binding of TnsD to these sequences mirrors the transposition activity. In contrast to the human gfpt sequences, Tn7 does not transpose adjacent to the gfa-1 sequence, the glmS analog in Saccharomyces cerevisiae. We also report that a nucleosome core particle assembled on the human gfpt-1 sequence reduces Tn7 transposition by likely impairing the accessibility of target DNA to the Tns proteins. We discuss the implications of these findings for the potential use of Tn7 as a site-specific DNA delivery agent for gene therapy.

INTRODUCTION

The bacterial transposon Tn7 is unique among transposons, in that it is able to target a specific site in the Escherichia coli genome called attTn7 (1). The Tn7 modus operandi is ingenious: the 35 bp target recognition sequence is entirely contained within a highly conserved, essential gene, glucosamine synthetase (glmS), with the actual site of insertion located 25 bp downstream of the coding sequence (see schematic in Figure 1) such that Tn7 insertions are non-deleterious to the host (1). The reconstituted in vitro TnsABC+D transposition reaction has been fundamental to our understanding of the biochemistry of Tn7 transposition (2) and also provides a powerful method to evaluate potential target sites for Tn7 transposition (3,4). Four Tn7 proteins are required for transposition into attTn7: TnsD, the target selector, binds in a sequence-specific manner to attTn7 and recruits TnsC, an ATP-dependent, non-sequence specific DNA-binding, regulator protein, followed by TnsA and TnsB, which together constitute the transposase and carry out the chemistry of recombination (1,5). Transposition is stringently regulated such that TnsD binding to the target DNA is the first step in assembling a nucleoprotein complex containing TnsABC+D, target DNA and donor DNA; only then can the DNA breakage and joining reactions be initiated (2,3,6).

What are the particular target DNA requirements for attracting Tn7 transposition? The DNA sequence, from nucleotide positions +23 to +58 (shown in Figure 1), which constitutes the binding site for TnsD, the target selector protein, is sufficient for targeting Tn7 transposition (7). The DNA sequence elsewhere, i.e. around the actual site of insertion (nucleotide position 0) and all the way up to nucleotide position +22 (Figure 1), can be varied without significantly affecting Tn7 transposition into the target site. The DNA sequence from +23 to +58, required for targeting Tn7 transposition, is contained entirely within the C-terminus of the glmS gene (Figure 1).

Analogs of the glmS gene are found in all organisms from bacteria and yeast to humans. There are two analogs of glmS in the human genome, glutamine-fructose-6-phosphate-transaminase-1 and 2 (gfpt-1 and gfpt-2) located at chromosome 2p13 and 5q34-q35 (8–10). In addition to the analogs of glmS, a sequence search also reveals an identical copy of the E.coli attTn7 target sequence located on chromosome 11 (clone RP11-159C12) of the human genome. The glmS analog in Saccharomyces cerevisiae, located on chromosome XI, is termed gfa-1 (11). Glucosamine synthetase, the product of the glmS gene, is the first and rate-limiting enzyme in hexosamine...
metabolism (12). The C-terminal portion of the GlmS protein, which contains the isomerase domain, is very highly conserved. Indeed, the amino acid sequence is identical for all four proteins listed above; differences in nucleotide sequences are limited primarily to the ‘wobble’ position in each case (Figure 1). Thus, the highly conserved C-terminal region of glmS analogs presents potential target sites for Tn7 transposition in the human genome, as well as that of yeast and other organisms.

The high frequency and site-specificity of Tn7 transposition make Tn7 an attractive candidate for developing a reliable, site-specific DNA delivery system for eukaryotic genomes with potential use in gene therapy. Several important questions need to be addressed in this regard: Can TnsD, the target selector protein, recognize and bind the human analogs of glmS and will Tn7 transpose into this human target sequence? How would the Tns proteins interact with a chromatinized target site?

Here, we show that Tn7 can transpose adjacent to the human gfp-t-1 sequence in vitro at a frequency comparable with the E.coli sequence; the human gfp-t-2 sequence shows reduced transposition. Transposition activity closely mirrors the degree of TnsD binding to glmS and the analogous gfp-t-1 and gfp-t-2 sequences. We also report that the positioning of a nucleosome on the human gfp-t-1 sequence reduces Tn7 transposition by likely blocking the accessibility of target DNA to the Tns proteins. Interestingly, Tn7 does not transpose adjacent to the gfa-1 sequence, the glmS analog in S.cerevisiae. We discuss the implications of these observations for the potential use of Tn7 as a site-specific DNA delivery tool.

**Figure 1.** Comparison of glmS and its analogs bearing potential Tn7 target sites. The top portion is a schematic (not drawn to scale) of the Tn7 target site in E.coli termed attTn7. The central nucleotide of the 5 bp duplication upon Tn7 insertion is denoted nucleotide position 0, the TnsD-binding site (hatched rectangle) is marked from +23 to +58, and the coding region of the glmS gene is marked by the thick arrow. The orientation-specificity of the Tn7 insertion with the right end located proximal to the glmS gene and the TnsD-binding site is also shown. The DNA sequence from +23 to +58 is sufficient for Tn7 transposition in vitro and in vivo. Below the arrow are listed the conserved amino acid sequence for all four gene products. This is followed by the sequence comparison of the 3’ proximal to the glmS gene. The central nucleotide of the 5 bp duplication upon Tn7 insertion is denoted nucleotide position 0, the TnsD-binding site (hatched rectangle) is marked by the thick arrow. The sequence of the human glmS analogs—the human glmS gene, Homo sapiens gfp-t-1 and gfp-t-2 genes and the S.cerevisiae gfa-1 gene. The sequence of the human clone RP11-159C12, which is an exact copy of the E.coli glmS target sequence, is also listed. Nucleotides in capitals are conserved and nucleotides in lower case are divergent from the E.coli sequence. The sequence of the 5 bp duplication, which would result from Tn7 transposition, is also listed for each target sequence.

**MATERIALS AND METHODS**

**Plasmids and DNA fragments: glmS analogs—the human gfp-t-1, gfp-t-2 and yeast gfa-1 sequences**

50 bp sequences around the putative Tn7 insertion site (marked 0 on the sequences in Figure 1) of the human gfp-t-1 and yeast gfa-1 genes were PCR-cloned into pCRRII-TOPO (Invitrogen) to yield target plasmids pPK222 and pPK221, respectively. Similarly, the plasmid pPK230 was constructed by cloning the analogous human gfp-t-2 sequence into the plasmid pLITMUS28. Plasmid pPK13 containing the E.coli glmS sequence has been described previously (3). For the in vitro transposition reactions, the Tn7-bearing Donor plasmid pPK21 contains 174 bp from the left end of Tn7 and 199 bp from the right end of Tn7 in a pUC19 vector. The chloramphenicol resistance containing Tn7 donor pGPS2.1 and the kanamycin resistance carrying Tn7 donor pEMΔ were used to isolate and sequence Tn7 insertions into pPK222 and pPK230, respectively (see below).

170 bp EcoRI-digested fragments from plasmids pPK222 and pPK211, the 180 bp SpeI–XbaI fragment from pPK230, and the attTn7-containing 205 bp XbaI–HindIII fragment from pPK13 (3) were all 3’ end-labeled and purified using standard protocol (13). These end-labeled target DNA fragments containing glmS and its analogs were used for the in vitro experiments detailed below.

**Tn7 transposition reactions**

Tn7 transposition reactions contained the following components: ~0.02 pmol 3’ end-labeled target DNA, 0.05 pmol Tn7 donor containing plasmid pPK21, 0.5 pmol TnsD, 0.5 pmol...
TnsC, 1.2 pmol TnsA and 0.3 pmol TnsB in a 24 μl reaction. The reaction buffer included 24 mM HEPES, pH 7.6, 8 mM Tris, pH 7.6, 1.7 mM ATP, 3.7 mM DTT, 0.2 mg/ml BSA, 15.4 mM magnesium acetate, 0.8 μg/μl sonicated salmon sperm DNA, 73 mM KCl, 51 mM NaCl, 0.1 mM EDTA, 0.4 mM CHAPS and 4.7% glycerol. The reaction was staged as follows: the target DNA, TnsC and TnsD were pre-assembled in the reaction buffer by incubating at 30°C for 20 min; followed by the addition of the donor, TnsA, TnsB and Mg2+ and further incubation at 30°C for 30 min. The reaction was stopped by phenol/chloroform extraction and precipitation. The products of the transposition reaction were run out on a 6% polyacrylamide gel, dried and subject to phosphorimaging (Storm/Molecular Dynamics) to yield the phosphorimage shown in Figure 2A. The bands on the gels were quantified using ImageQuant (Molecular Dynamics) and KaleidoGraph software (for Figure 2B).

The products of transposition include: the predominant simple insertion results from the excision of the Tn7 by double-strand breaks at both ends of the transposon followed by joining of both ends to the target, and a product termed SEJ, where a single end of the transposon has undergone a double-strand break followed by joining to the target. The product marked with an asterisk in Figures 2A and 4 is not dependent on the presence of the Tns proteins or transposition. These bands were cut out of the gel, the DNA isolated and separated on denaturing PAGE; they correspond in size to the unreacted target DNA fragment (data not shown). The simplest explanation for the observed altered mobility on native polyacrylamide gels is that these are different conformers of the target DNA fragment. Transposition reactions using the gfpt-2 target sequence show a doublet for the simple insertion product. This doublet most likely results from the loss during electrophoresis of the short DNA fragment on either side of the 5 bp gap resulting from transposition. Tn7 Transposition into the larger pPK230 plasmid containing the gfpt-2 target sequence results in a single simple insertion product, which has been confirmed by sequencing (see below).

**Sequencing of Tn7 insertions into human gfpt-1 and gfpt-2 sequences**

Standard in vitro transposition reactions were carried out using pPK222 and pPK221 as targets and pGPS2.1 (New England Biolabs) as the Tn7 donor. For pPK230, the standard pEMΔ donor was used. After transformation in E.coli and selection for the appropriate antibiotic resistance, 20 independent colonies bearing Tn7-containing targets were picked, the plasmid DNA isolated and sequenced. The 5 bp duplication resulting from Tn7 transposition as well as the orientation of Tn7 insertion were mapped for each case.

**Electrophoretic mobility shift assay (EMSA)**

Various amounts of TnsD were incubated with radiolabeled target DNA in standard binding reactions and run out on 5% polyacrylamide gels as described elsewhere (3).

**Reconstitution of nucleosome core particles on the human gfpt-1 sequence**

Nucleosome core particles were assembled on the 3’ end-labeled, 170 bp DNA fragment containing the human gfpt-1 sequence by the salt-gradient exchange method (14). Reconstitution was monitored on a 5% polyacrylamide gel as described elsewhere (15). The gfpt-1 containing sequence does not yield a completely reconstituted product; a mixture comprised 85% nucleosome core particles and 15% free DNA. When the conditions of reconstitution were altered, such as decreased non-specific competitor DNA or increased histone octamers, to improve the ratio of nucleosome core particles to free DNA, we observed multimers and aggregates. Also, the lack of unique translational positioning of the nucleosome core particle on the DNA fragment and the large size of the Tn7 frame. The products of the transposition reaction were run out on a 6% polyacrylamide gel, dried and subject to phosphorimaging (Storm/Molecular Dynamics) to yield the phosphorimage shown in Figure 2A. The bands on the gels were quantified using ImageQuant (Molecular Dynamics) and KaleidoGraph software (for Figure 2B).

The products of transposition include: the predominant simple insertion results from the excision of the Tn7 by double-strand breaks at both ends of the transposon followed by joining of both ends to the target, and a product termed SEJ, where a single end of the transposon has undergone a double-strand break followed by joining to the target. The product marked with an asterisk in Figures 2A and 4 is not dependent on the presence of the Tns proteins or transposition. These bands were cut out of the gel, the DNA isolated and separated on denaturing PAGE; they correspond in size to the unreacted target DNA fragment (data not shown). The simplest explanation for the observed altered mobility on native polyacrylamide gels is that these are different conformers of the target DNA fragment. Transposition reactions using the gfpt-2 target sequence show a doublet for the simple insertion product. This doublet most likely results from the loss during electrophoresis of the short DNA fragment on either side of the 5 bp gap resulting from transposition. Tn7 Transposition into the larger pPK230 plasmid containing the gfpt-2 target sequence results in a single simple insertion product, which has been confirmed by sequencing (see below).

**Sequencing of Tn7 insertions into human gfpt-1 and gfpt-2 sequences**

Standard in vitro transposition reactions were carried out using pPK222 and pPK221 as targets and pGPS2.1 (New England Biolabs) as the Tn7 donor. For pPK230, the standard pEMΔ donor was used. After transformation in E.coli and selection for the appropriate antibiotic resistance, 20 independent colonies bearing Tn7-containing targets were picked, the plasmid DNA isolated and sequenced. The 5 bp duplication resulting from Tn7 transposition as well as the orientation of Tn7 insertion were mapped for each case.

**Electrophoretic mobility shift assay (EMSA)**

Various amounts of TnsD were incubated with radiolabeled target DNA in standard binding reactions and run out on 5% polyacrylamide gels as described elsewhere (3).

**Reconstitution of nucleosome core particles on the human gfpt-1 sequence**

Nucleosome core particles were assembled on the 3’ end-labeled, 170 bp DNA fragment containing the human gfpt-1 sequence by the salt-gradient exchange method (14). Reconstitution was monitored on a 5% polyacrylamide gel as described elsewhere (15). The gfpt-1 containing sequence does not yield a completely reconstituted product; a mixture comprised 85% nucleosome core particles and 15% free DNA. When the conditions of reconstitution were altered, such as decreased non-specific competitor DNA or increased histone octamers, to improve the ratio of nucleosome core particles to free DNA, we observed multimers and aggregates. Also, the lack of unique translational positioning of the nucleosome core particle on the DNA fragment and the large size of the Tn7 sequence by the salt-gradient exchange method (14). Reconstitution was monitored on a 5% polyacrylamide gel as described elsewhere (15). The gfpt-1 containing sequence does not yield a completely reconstituted product; a mixture comprised 85% nucleosome core particles and 15% free DNA. When the conditions of reconstitution were altered, such as decreased non-specific competitor DNA or increased histone octamers, to improve the ratio of nucleosome core particles to free DNA, we observed multimers and aggregates. Also, the lack of unique translational positioning of the nucleosome core particle on the DNA fragment and the large size of the Tn7 sequence by the salt-gradient exchange method (14). Reconstitution was monitored on a 5% polyacrylamide gel as described elsewhere (15). The gfpt-1 containing sequence does not yield a completely reconstituted product; a mixture comprised 85% nucleosome core particles and 15% free DNA. When the conditions of reconstitution were altered, such as decreased non-specific competitor DNA or increased histone octamers, to improve the ratio of nucleosome core particles to free DNA, we observed multimers and aggregates. Also, the lack of unique translational positioning of the nucleosome core particle on the DNA fragment and the large size of the Tn7
post-transposition complex prevented isolation of nucleosomal DNA after the Tn7 transposition reaction.

Transposition reactions with free and nucleosomal DNA were identical to those described above, except that they also contained the following components from the nucleosome reconstitution reactions, i.e. 2 mM Tris, pH 7.6, 0.2 mM EDTA and 21 mM NaCl in addition to the buffer conditions in the reactions described earlier.

RESULTS

Tn7 transposes into the gfpt-1 and gfpt-2 human target DNA sequences

End-labeled DNA fragments containing the human gfpt-1 and gfpt-2 sequences were used as targets for TnsABC+D-mediated transposition reaction in vitro (3). Under limiting conditions of the target selector protein, TnsD, the human gfpt-1 sequence shows levels of transposition comparable with that of the E.coli glmS sequence (compare the simple insertion product, 32% with glmS and 31% with gfpt-1, in lanes 2 for each sequence in Figure 2A). The human gfpt-2 sequence also targets Tn7 transposition, but at a reduced level; 23% simple insertion with the gfpt-2 target as compared with 32% with the glmS target, shown in Figure 2A. Interestingly, the analogous yeast gfa-1 sequence shows no Tn7 transposition (Figure 2A, right panel).

In order to better compare the ability of the E.coli and human sequences to serve as targets for Tn7 transposition, we analyzed the time course of recombination, the results of which are graphed in Figure 2B. Tn7 transposition into these targets, particularly at the earlier time points (2–20 min), follows the following trend: glmS > gfpt-1 > gfpt-2 as seen in Figure 2B. We know from longer reactions (shown in Figure 2A) that the glmS and human gfpt-1 sequence show comparable levels of transposition, and the human gfpt-2 sequence shows reduced transposition.

We also performed standard in vitro Tn7 transposition reactions using supercoiled target plasmids containing the human gfpt-1 and gfpt-2, and the yeast gfa-1 sequences, respectively, and mini-Tn7 containing donor plasmids. When the products of the in vitro transposition reactions were transformed and selected for targets containing Tn7 insertions (4), the yeast target did not yield any colonies. Twenty independent Tn7 insertions into each of the two human sequence-containing targets were selected and sequenced (4). As expected, the 5 bp duplication, bearing the 5′-AAGCA-3′ sequence for the gfpt-1 sequence and the 5′-CATCA-3′ sequence for the gfpt-2 gene, was observed for each Tn7 insertion (Figure 1). Moreover, all the insertions were in the appropriate orientation, i.e. the right end of Tn7 was located proximal to the gfpt sequence as shown in Figure 1. Thus, Tn7 insertions adjacent to the human gfpt-1 and gfpt-2 sequences in vitro are identical to those adjacent to the glmS gene in E.coli in terms of site- and orientation-specificity.

The binding of TnsD to the human gfpt-1 target is robust

Is the Tn7 transposition activity seen above due to different binding affinities of TnsD to the various target sequences? We performed EMSAs (3) to determine the extent of TnsD binding to the potential target sites. Not surprisingly, we found that TnsD binding to the human gfpt-1 target was robust and similar to the glmS target over a range of concentrations of TnsD (Figure 3). Once again, quantification of the binding data revealed reduced binding of TnsD to the human gfpt-2 sequence as compared with the corresponding E.coli sequence. In contrast, there was no detectable binding of TnsD to the analogous yeast sequence (Figure 3) in agreement with the lack of transposition activity seen in Figure 2A. Thus, the levels of transposition observed for the various target sequences closely mirror the extent of TnsD binding.

A nucleosome positioned on the human gfpt-1 sequence reduces Tn7 transposition

We have shown above that the human gfpt-1 sequence is an efficient target for Tn7 transposition in vitro. However, DNA in the eukaryotic cell is complexed with histones and other proteins in the form of chromatin and the basic repeating unit of the resulting structure is the nucleosome core particle (16). The wrapping of DNA around the histone octamer produces distinct structural perturbations in the target DNA that could potentially enhance recombination, as observed for HIV Integrase (17,18). Similarly, is a nucleosome core particle assembled on the human gfpt-1 sequence a preferred target for Tn7 transposition in vitro?

Using standard reconstitution methods (15), we assembled a single nucleosome core particle on the 170 bp DNA fragment from pPK222 containing the human gfpt-1 sequence to mimic the putative target site for Tn7 transposition in human cells, i.e. chromatinized DNA. It is important to note that unlike ‘model’ nucleosome positioning sequences, such as the 5S rDNA from Xenopus borealis (16), this gfpt-1 containing sequence does not yield a uniquely positioned, completely reconstituted product; a mixture comprising 85% nucleosome
but avoids the analogous gfa-1 sequence served as a Tn7 target in vivo, can reduce Tn7 transposition. The yeast gfa-1 sequence, which also bears considerable homology to glmS and the human target analogs, is neither able to bind TnsD nor target Tn7 transposition by our assays.

Our results show the following trend in the ability of the eukaryotic analogs to function as target sites for Tn7: glmS ≅ gftp-1 ≫ gftp-2 >> gfa-1; and most importantly, Tn7 transposition in vitro, or the lack thereof, closely mirrors TnsD binding to these eukaryotic target sequences.

It has been reported that TnsD binds the human gftp-1 and gftp-2 sequences, and the yeast gfa-1 sequence with significantly reduced affinity when compared with the glmS sequence (20). This TnsD-binding study used 20 bp DNA duplexes containing the sequence from +29 to +48 (see Figure 1), which does not completely correspond to the Tn7-binding site (20).

Sequence-specific recognition of target DNA by TnsD

It is intriguing that Tn7 recognizes and makes insertions adjacent to the human gftp-1 and gftp-2 sequences but avoids the yeast gfa-1 sequence. A comparison of these DNA sequences (Figure 1) shows that they primarily differ at the third ‘wobble’ position of multiple codons; certain changes in the glmS sequence continue to allow TnsD binding, whereas other alterations completely abolish TnsD binding. Therefore, TnsD presents an interesting case of a sequence-specific DNA-binding protein that recognizes a highly conserved sequence, but allows some flexibility within the binding site to incorporate DNA sequence changes due to base wobble. It is very likely that Tn7 target sites exist within glmS analogs in the genomes of other organisms ranging from bacteria to the higher chordates. Indeed, the sequencing of bacterial genomes has revealed Tn7 and Tn7-like transposons adjacent to the glmS analogs in multiple cases.

Tn7 is unique among transposons, in that it encodes a sequence-specific DNA-binding protein, TnsD, for the site-specific recognition of its target DNA. In E.coli, Tn7 transposition to locations other than attTn7, termed pseudo-att sites (bearing homology to the glmS sequence), is observed at a very low frequency (10 000 times less frequent than the glmS target sequence) (21). The inability of TnsD to bind the analogous yeast gfa-1 sequence and use it as a target for Tn7 transposition further attests to the stringent target site selectivity of Tn7, a characteristic that is highly desirable in a DNA delivery agent for potential gene therapy applications.

Multiple human targets for site-specific transposition by Tn7

We have found that Tn7 can transpose in vitro adjacent to the human gftp-1 and gftp-2 sequences, in a site- and orientation-specific manner; transposition activity for the gftp-1 target is at least comparable with the corresponding bacterial target DNA, and slightly reduced for the gftp-2 target. In addition to the glmS analogs in the human genome, there also exists a target site on chromosome 11, which is identical to the E.coli attTn7 target sequence. We already know from the sequence that this

**DISCUSSION**

TnsD binding and Tn7 transposition into targets containing the eukaryotic analogs of glmS

Here, we have shown that the bacterial transposon Tn7 makes insertions adjacent to the human gftp-1 and gftp-2 sequences, but avoids the analogous gfa-1 sequence from S.cerevisiae. It was previously reported that a subset of the human gftp-1 sequence served as a Tn7 target in vivo when transposition was performed in E.coli (19). However, this study did not precisely determine the relative frequency of Tn7 transposition into the human sequence as compared with the E.coli

**Figure 4.** Assembly of a nucleosome core particle on the human gftp-1 target reduces Tn7 transposition. The left panel shows an EMSA containing: free target DNA in lane 0, the nucleosome core particle when assembled on the target DNA in lane 1 and the assembled nucleosome core particle under the conditions of the Tn7 transposition reaction in lane 2. The letters F and N mark lanes containing reactions with either free or nucleosomal DNA, respectively. Bands corresponding to free and nucleosomal DNA are also marked. The right panel shows the results of the Tn7 transposition reaction on free and nucleosome-bound target DNA. Lanes 4 and 5 correspond to TnsABC+D-mediated transposition into free and nucleosomal DNA. Lanes 3 and 6 are controls to which none of the Tns proteins was added. Rest as in Figure 2A.
target site on chromosome 11 (clone RP11-159C12), which is not associated with any genes, should serve as an efficient target site for Tn7. Although glmS sequences amongst bacteria tend to show some variation, the att sequence in the human clone is identical to the E.coli attTn7 sequence, suggesting that this att sequence could be the result of contaminating E.coli sequence. (A BLAST search of the 180 bp sequence surrounding the att site in the human clone (RP11-159C12) shows that the entire sequence is identical to sequences found in several pBAC cloning vectors, which are known to contain the attTn7 target sequence.) Further studies need to be carried out to show that there is a chromosomal version of the E.coli attTn7, as seen in clone RP11-159C12.

These observations imply that Tn7 transposition could deliver DNA sequences of choice in a site- and orientation-specific manner at three locations within the human cell: adjacent to the sequence identical to attTn7 on chromosome 11, and also adjacent to the gfpt-1 (location 2p13) and the gfpt-2 (location 5q34–q35) genes.

We have also shown that a nucleosome positioned over the gfpt-1 target sequence can reduce transposition by likely occluding the target DNA from the Tns proteins. However, this observation does not preclude the ability of the human gfpt-1/2 sequences to serve as a target for Tn7 transposition for the following reasons: first, very little is known about chromatin structure around the C-terminus of the gfpt-1 and gfpt-2 genes (8) and the Tn7 target site DNA may not be assembled into nucleosomes in vivo; and second, even if it were to contain nucleosomes, chromatin structure is dynamic in vivo, when the putative target site DNA may become at least transiently accessible to the Tns proteins. Although chromatin structure may affect the relative Tn7 transposition frequency in vivo, the presence of three separate targets sites and the dynamic nature of chromatin structure mean that it does not pose an insurmountable obstacle to Tn7 transposition in human cells.

We have shown directly target sequences in the human genome for site- and orientation-specific Tn7 transposition. We have also seen that the Tn7 target selector protein, TnsD, can tolerate limited and very specific changes in target sequence. Do the abovementioned observations imply that a large number of potential target sites will be available for Tn7 transposition in the human genome? While we cannot rule out the existence of additional target sites, our experience with Tn7 transposition in bacterial genomes points to a very limited number of high affinity target sites. For example, the frequency of Tn7 transposition into attTn7 is very high and estimated to be ~1–10% in E.coli (22) and transposition into attTn7-like or pseudo-att sites, when attTn7 is unavailable, occurs at a very low frequency (10000 times less than attTn7) (21). Moreover, our results with the yeast sequence here also attest to Tn7’s ability to avoid other attTn7-like target sites in the genome. Similarly, degenerate att-like target sequences may be available in the human genome, but the transposition frequency to these sites will be very significantly reduced and we believe that Tn7 will preferentially transpose to the three high affinity human target sites on chromosomes 2, 5 and 11 described here. It is, therefore, reasonable to expect Tn7 to serve as an efficient tool for targeted DNA delivery.

Several other observations bode well for the potential use of Tn7 as a specific DNA delivery tool for gene therapy. (i) The repertoire of Tn7 also includes a mechanism of transposition immunity, i.e. the presence of a Tn7 element at the target site prevents further use of the target and multiple insertions are avoided (1). Thus, one would expect a single copy of Tn7 to be inserted adjacent to the gfpt-1 or the gfpt-2 genes in an orientation-specific manner. (ii) The actual site of Tn7 insertion is located downstream of the gfpt-1/2 genes, and so, it is very likely that Tn7 insertion will prove to be non-deleterious to the cell. (iii) There may also be particular advantages in employing a bacterial transposon for site-specific delivery in human cell lines, in that Tn7 may be able to bypass cellular mechanisms specifically evolved to restrict eukaryotic mobile elements. (iv) Finally, bacterial proteins have been successfully expressed in human cell lines, including the Cre recombinase for the very purpose of gene therapy (23). All of these factors are encouraging for the potential expression of functional Tns proteins to carry out site-specific Tn7 transposition in the human cell.

ACKNOWLEDGEMENTS

This work was supported by NIH grant P01CA16519. N.L.C. is a Howard Hughes Medical Institute (HHMI) Investigator. Funding to pay the Open Access publication charges for this article was provided by HHMI.

REFERENCES

1. Craig, N., Craigie, R., Gellert, M. and Lambowitz, A. (2002) Mobile DNA II. ASM Press, Washington, DC, pp. 423–456.
2. Bainton, R.J., Kubo, K.M., Feng, J.-N. and Craig, N.L. (1993) Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. Cell, 72, 931–943.
3. Kuduvalli, P., Rao, J.E. and Craig, N.L. (2001) Target DNA structure plays a critical role in Tn7 transposition. EMBO J., 20, 924–932.
4. Rao, J.E., Miller, P.S. and Craig, N.L. (2000) Recognition of triple-helical DNA structures by transposon Tn7. Proc. Natl Acad. Sci. USA, 97, 3936–3941.
5. Peters, J.E. and Craig, N.L. (2001) Tn7: smarter than we thought. Nature Rev. Mol. Cell Biol., 2, 806–814.
6. Skelding, Z., Samovsky, R. and Craig, N.L. (2002) Formation of a nucleoprotein complex containing Tn7 and its target DNA regulates transposition initiation. EMBO J., 21, 3494–3504.
7. Waddell, C.S. and Craig, N.L. (1989) Tn7 transposition: recognition of the attTn7 target sequence. Proc. Natl Acad. Sci. USA, 86, 3958–3962.
8. Milewski, S. (2002) Glucosamine-6-phosphate synthase—the multi-facets enzyme. Biochim. Biophys. Acta, 1597, 173–192.
9. McKnight, G.L., Mudri, S.L., Mathewes, S.L., Traxinger, R.R., Marshall, S., Sheppard, P.O. and O’Hara, P.J. (1992) Molecular cloning, cDNA sequence, and bacterial expression of human glutamine:fructose-6-phosphate amidotransferase gene (GFPT) to chromosome 2p13. Genomics, 14, 422–423.
10. Watzele, G. and Tanner, W. (1989) Cloning of the glutamine:fructose-6-phosphate amidotransferase gene from yeast. Pheromonal regulation of its transcription. J. Biol. Chem., 264, 8753–8758.
11. Dutka-Malen, S., Mazodier, P. and Badet, B. (1988) Molecular cloning and overexpression of the glucosamine synthetase gene from Escherichia coli. Biochimie, 70, 287–290.
12. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbor Press, Cold Spring Harbor, NY.
13. Lidstrom, H.R., Kuduvalli, P.N., Bengtsson, M., Cao, H., Tullius, T.D. and Kubista, M. (1999) Nucleosome structural features and intrinsic properties of the TATAAAGCC repeat sequence. J. Biol. Chem., 274, 31847–31852.
14. Hayes, J.J. and Lee, K.M. (1997) In vitro reconstitution and analysis of mononucleosomes containing defined DNAs and proteins. Methods, 12, 2–9.
16. Wolfe, A. (1998) *Chromatin Structure and Function, 3rd edn*. Academic Press, San Diego, CA, pp. 19–40.

17. Pruss, D., Bushman, F.D. and Wolfe, A.P. (1994) Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. *Proc. Natl Acad. Sci. USA*, 91, 5913–5917.

18. Pruss, D., Reeves, R., Bushman, F.D. and Wolfe, A.P. (1994) The influence of DNA and nucleosome structure on integration events directed by HIV integrase. *J. Biol. Chem.*, 269, 25031–25041.

19. Cleaver, S.H. and Wickstrom, E. (2000) Transposon Tn7 gene insertion into an evolutionarily conserved human homolog of *Escherichia coli* attTn7. *Gene*, 254, 37–44.

20. Chakrabarti, A., Desai, P. and Wickstrom, E. (2004) Transposon Tn7 protein TnsD binding to *Escherichia coli* attTn7 DNA and its eukaryotic orthologs. *Biochemistry*, 43, 2941–2946.

21. Kubo, K.M. and Craig, N.L. (1990) Bacterial transposon Tn7 utilizes two classes of target sites. *J. Bacteriol.*, 172, 2774–2778.

22. Craig, N.L. (1996) Transposon Tn7. *Curr. Top. Microbiol. Immunol.*, 204, 27–48.

23. Craig, N., Craigie, R., Gellert, M. and Lambowitz, A. (2002) *Mobile DNA II*. ASM Press, Washington, DC, pp. 38–58.