A novel approach to design of *cis*-acting DNA structural elements for regulation of gene expression *in vivo*

Partha S. Sarkar, Rajesh Bagga, P. Balagurumoorthy and Samir K. Brahmachari*†

Molecular Biophysics Unit, and *Centre for Genetic Engineering, Indian Institute of Science, Bangalore 560 012, India

Taking advantage of the degeneracy of the genetic code we have developed a novel approach to introduce, within a gene, DNA sequences capable of adopting unusual structures and to investigate the role of such sequences in regulation of gene expression *in vivo*. We used a computer program that generates alternative codon sequences for the same amino-acid sequence to convert a stretch of nucleotides into an inverted-repeat sequence with the potential to adopt cruciform structure. This approach was used to replace a 51-base-pair *Eco*RI–*Hin*dIII segment in the N-terminal region of the β-galactosidase gene in plasmid pUC19 with a 51-bp synthetic oligonucleotide sequence with the potential to adopt a cruciform structure with 18 bp in the stem region. In selecting the 51-bp sequence, care was taken to include those codons that are preferred in *E. coli*. *E. coli* DH5α cells harbouring the plasmid containing the redesigned sequence showed drastic reduction in expression of the β-galactosidase gene compared to cells harbouring the plasmid with the native sequence. This approach demonstrates the possibility of introducing DNA secondary-structure elements to alter regulation of gene expression *in vivo*.

GENE expression is a highly regulated process, much of the regulation being exerted at the level of transcription initiation[1,2]. Recent evidences show that the regulation of gene expression can also occur at the level of transcription elongation[3,4]. Although some *cis*-acting DNA sequences[5] and RNA secondary structures[6] have been implicated in regulation of transcription elongation, no definite evidence exists to show the possibility of DNA secondary structure *per se* being involved in modulation of the transcription-elongation process *in vivo*. Several *in vitro* studies have shown that unusual DNA structures, such as the left-handed Z-form[7], cruciform DNA[4] and heteronomous DNA[8], can influence transcription. Insertion of a synthetic DNA sequence capable of adopting a cruciform structure within the promoter region was found to inhibit transcription initiation *in vivo*[9]. The sensitivity to tetracycline observed in *E. coli* cells harbouring plasmid pRM36 has been attributed to the insertion of a 'structural cassette' with short stretches of left-handed Z and right-handed B helices in the tet^R promoter region[10,11]. Our earlier studies using pBR322 form V DNA as a template for *E. coli* RNA polymerase *in vitro* showed that a supercoil-stabilized cruciform structure present within the *rep* gene could block elongation by RNA polymerase, and this block could be abrogated once the cruciform structure was destabilized[4,10]. However, because of the high superhelical density in this molecule, the physiological relevance of these results is limited.

To delineate the DNA structural elements responsible for transcriptional control *in vivo* we have developed a novel approach to introduce structural cassettes within a gene and to follow the effect of such unusual DNA structures, stabilized *in vivo*, on the transcription-elongation process. We have used expression of the β-galactosidase gene as a model system for this study. Taking advantage of the degeneracy of the genetic code we have incorporated silent mutations to redesign the 5′-end of the β-galactosidase gene in plasmid pUC19 so as to introduce an inverted-repeat sequence that has the potential to adopt a cruciform structure *in vivo*. This is, to our knowledge, the first report that demonstrates '*in vivo*' that a potential cruciform structure present within the gene could act in *cis* to down-regulate the expression of a gene.

Methods

Synthesis and annealing of complementary oligodeoxyribonucleotides

The complementary 51-nucleotide-long oligonucleotides were synthesized on a Pharmacia Gene Assembler Plus using phosphoramidite chemistry. The oligomers were purified on a 20% polyacrylamide–8 M urea gel and by passing through an NAP (Pharmacia) column[12]. To anneal the complementary strands, the oligonucleotides were mixed in equimolar ratio, heated at 95°C for 3 min, and slowly brought to room temperature. The annealed sample was stored at −20°C. Circular-dichroism (CD) spectra of the annealed product were recorded on a Jasco J-500 spectropolarimeter equipped with a data processor[12].

*For correspondence

Cloning of 51-mer duplex into pUC19

Plasmids used in this study were prepared by the modified method of Maniatis[13] and purified on a CsCl density gradient. Vector pUC19 was double-digested with *Eco*RI and *Hin*dIII (all restriction enzymes from New England Biolabs, USA, and used according to vendor's specifications) and the 2635-bp fragment was eluted from an agarose gel. The sticky ends of this linear fragment were dephosphorylated using calf intestinal phosphatase. The synthetic 51-mer duplex was phosphorylated using polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP. A molar ratio of 2:1 of vector to insert was used for cloning. Dephosphorylated vector was mixed with phosphorylated insert, the mixture vortexed, and ligation carried out for 16 h at 4°C using T4 DNA ligase (New England Biolabs). One hundred ng of ligation product was used for transformation of competent *E. coli* DH5α cells. The recombinant plasmid is called pSBC1.

β-Galactosidase assay

β-Galactosidase was assayed as described by Miller[14]. In brief, *E. coli* cells containing either pSBC1 or pUC19 were grown in presence of 1–2 mM isopropyl β-D-thiogalactoside (IPTG; Sigma). Aliquots were taken at different times until cells reached stationary phase. Enzyme was released from the cells by treatment with toluene. β-Galactosidase activity was measured as release of *o*-nitrophenol from *o*-nitrophenyl β-D-galactopyranoside (ONPG; Sigma) by reading absorbance at 420 nm. Protein was estimated by the method of Bradford[15]. Specific activity of the enzyme was calculated as described by Miller[14].

Strategy for designing a potential cruciform structure

We selected a 51-nucleotide-long *Eco*RI–*Hin*dIII segment of pUC19 which encompasses the polylinker region and also codes for the N-terminal region of β-galactosidase. This segment of the β-galactosidase gene codes for amino acids 5 to 22. To design a sequence with the potential to adopt a cruciform structure a computer program was developed (details will be published elsewhere). This program takes a gene sequence as input and generates all possible DNA sequences that code for the same amino-acid sequence taking advantage of codon degeneracy. All these DNA sequences were searched for the presence of an inverted repeat with stem length $n \geq 8$ bp and loop length $m = 3$ or 4 bases. Of several possible combinations of degenerate sequences only a few sequences were found to have the desired inverted repeat (with $n \geq 8$ and $m = 4$). In some of these sequences we also altered some of the codons to extend the stem length of the inverted-repeat sequence and to include A/T nucleotides within the loop region. Extended stem length and presence of A/T nucleotides within the loop region have been shown to provide thermodynamic stability to the extruded cruciform[16,17]. Figure 1,*a* outlines our strategy. A sequence was chosen with $n = 18$ and $m = 4$ (Figure 1,*b*). To achieve maximum stem length so as to stabilize the potential cruciform structure *in vivo*, the codons for amino acids at positions 10, 11 and 15 were changed from those for Arg, Ser and Asp to those for Ser, Thr and Asn respectively (Figure 1,*a*). It has been shown earlier that the first 27 amino acids of the N-terminus of β-galactosidase can be removed and substituted by a wide variety and different number of amino acids without affecting enzyme activity[18]. For instance, even complete inversion of the *Eco*RI–*Hin*dIII segment in pUC18 *vis-a-vis* pUC19 does not change the activity of β-galactosidase. In the modified sequence care was taken to select those codons that showed positive codon score, i.e. those codons that are favoured in *E. coli* (Table 1)[19]. Suitable restriction sites were also included in the modified sequence to facilitate the characterization of the clone and detection of cruciform structure under physiological conditions (Figure 1,*a*).

Cloning the redesigned sequence

The complementary oligonucleotides synthesized were purified on urea–polyacrylamide gels prior to annealing. The inverted-repeat sequence necessitated purification at higher temperature. The annealed product was checked to determine whether the two strands had hybridized to give a duplex or each strand had self-hybridized to give a hairpin loop. Figure 2 shows the CD spectrum of the annealed oligomers, with a positive peak at 274 nm and a negative peak at 242 nm, suggesting a double-stranded B-DNA-like structure. The annealed product was phosphorylated at the 5'-OH using [γ-^{32}P]ATP. The phosphorylated product migrated as one band corresponding to 51 bp on an 8% polyacrylamide gel (Figure 3).

These oligomers had been designed such that, upon annealing, cohesive termini for *Eco*RI and *Hin*dIII are generated. This enabled us to replace the small *Eco*RI–*Hin*dIII segment of pUC19 corresponding to the N-terminal region of β-galactosidase with the annealed synthetic oligomer without changing the reading frame. All the transformants were screened for the presence of the insert by examining sensitivity to *Xho*I and insensitivity to *Bam*HI and *Pst*I (see Figure 1,*a*). Figure 4 shows the restriction pattern of one of the transformants, indicating complete cleavage by *Xho*I and no cleavage by *Pst*I and *Bam*HI. The restriction pattern of the insert region of pSBC1, one of the clones which showed the presence of the insert, is shown in Figure 5. Plasmid pSBC1 was digested with *Hin*dIII and partially

a

N-TERMINAL SEQUENCE OF β-GALACTOSIDASE GENE

```
   HindIII          PstI                          BamHI                          EcoRI
5'-AGC TTG CAT GCC TGC AGG TCG ACT CTA GAG GAT CCC CGG GTA CCG AGC TCG - 3'
    5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22
     3'-AC GTA CGG ACG TCC AGC TGA GAT CTC CTA GGG GCC CAT GGC TCG AGC TTA A- 5'
```

CORRESPONDING mRNA SEQUENCE

```
5'-AGC UUG CAU GCC UGC AGG UCG ACU CUA GAG GAU CCC CGG GUA CCG AGC  UCG AAU-3'
    5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20   21  22
```

CORRESPONDING AMINO ACIDS SEQUENCE

```
   Ser-Leu-His-Ala-Cys-Arg-Ser-Thr-Leu-Glu-Asp-Pro-Arg-Val-Pro-Ser-Ser-Asn
    5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22
```

MODIFIED mRNA SEQUENCE USING DEGENERATE CODONS

```
5'-AGC UUG CAU GCC UGC UCG ACG ACA CUC GAG AAU CCU CGA GUG CCG UCG AGC AAU-3'
    5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22
```

CORRESPONDING POLYPEPTIDE SEQUENCE

```
   Ser-Leu-His-Ala-Cys-Ser-Thr-Thr-Leu-Glu-Asn-Pro-Arg-Val-Pro-Ser-Ser-Asn
    5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22
```

CORRESPONDING MODIFIED SEQUENCE OF DNA

```
   HindIII                          XhoI  HinfI  XhoI
5'-AGC TTG CAT GCC TGC TCG ACG ACA CTC GAG AAT CCT CGA GTG CCG TCG AGC -3'
    5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22
     3'-AC GTA CGG ACG AGC TGC TGT GAG CTC TTA GGA GCT CAC GGC AGC TCG TTA A-5'

                  <----------------------      ---------------------->
```

b

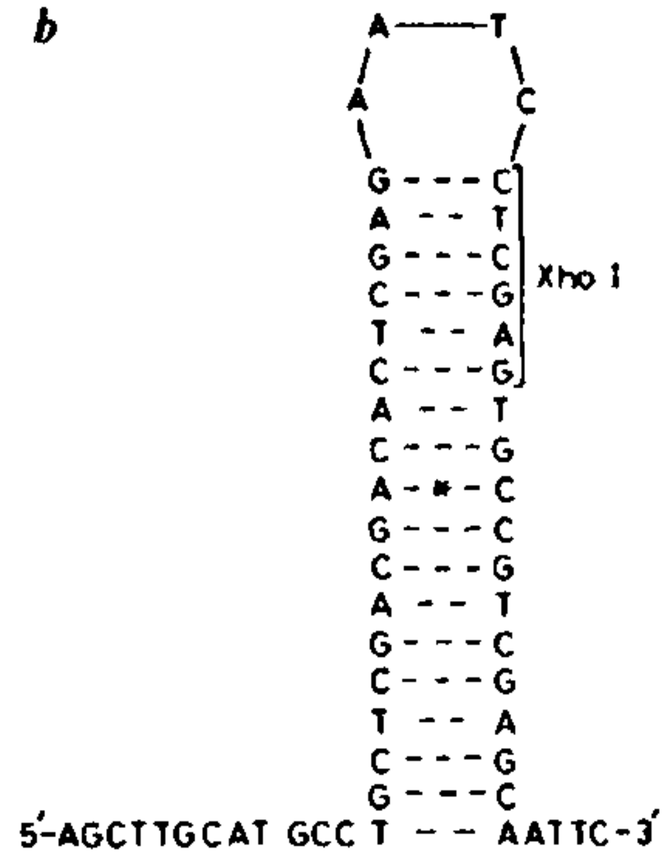

Figure 1. *a*, Strategy for designing inverted-repeat sequence within the β-galactosidase gene. Both strands of the modified DNA sequence were synthesized and purified before cloning. *b*, Potential cruciform structure, * indicates mismatch

digested with *Xho*I, and the 3′-recessive ends were gap-filled using [α-^{32}P]dATP and DNA polymerase Klenow fragment. As expected two bands corresponding to 26 and 36 bp (see Figure 1,*a*) were observed.

In vivo expression of β-galactosidase

E. coli cells harbouring either pUC19 or pSBC1 were plated on LB plates containing 5-bromo-4-chloro-3-

Table 1. Comparison of codon frequencies for pUC19 and pSBC1 in the *Eco*RI–*Hin*dIII segment of β-galactosidase gene.

Codon position	Original codon	New codon	Original aa coded	New aa coded	Codon frequency*		Average codon frequency†	
					Original	New	Original	New
9	UGC	UGC	Cys	Cys	0 639	0 639	0 569	0 569
10	AGG	UCG	Arg	Ser	0 003	0 047	0 027	0.133
11	UCG	ACG	Ser	Thr	0 047	0 099	0 133	0.187
12	ACU	ACA	Thr	Thr	0 364	0 053	0 271	0 084
13	CUA	CUC	Leu	Leu	0 007	0 076	0 024	0 099
14	GAG	GAG	Glu	Glu	0 242	0 242	0 273	0 273
15	GAU	AAU	Asp	Asn	0 363	0 076	0 491	0 307
16	CCC	CCU	Pro	Pro	0 009	0 099	0 111	0 116
17	CGG	CGA	Arg	Arg	0 003	0 009	0 064	0 044
18	GUA	GUG	Val	Val	0 274	0 191	0 214	0 264
19	CCG	CCG	Pro	Pro	0 768	0 768	0 607	0 607
20	AGC	UCG	Ser	Ser	0 191	0 047	0 197	0 133
21	UCG	AGC	Ser	Ser	0 047	0 191	0 133	0 197
22	AAU	AAU	Asn	Asn	0 076	0 076	0 307	0 307

*Codon frequencies[19] are calculated from the codon used in strongly expressed genes.
†Average codon frequencies[19] are calculated considering both strongly and weakly expressed genes
'Original' implies pUC19
'New' implies pSBC1

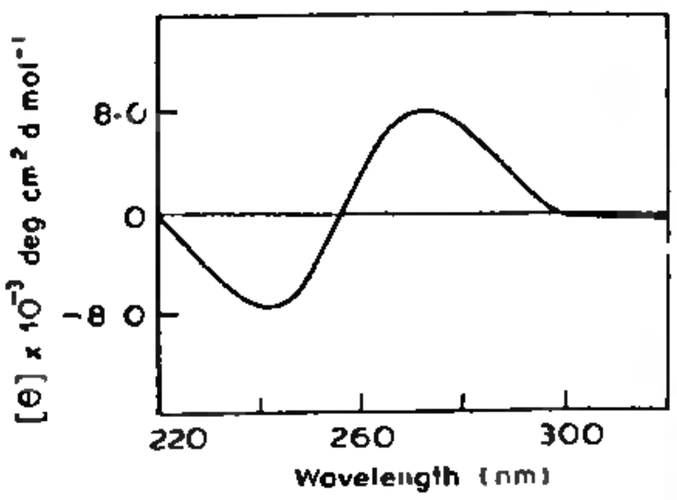

Figure 2. CD spectrum of 51-mer duplex in 10 mM Tris, 1 mM EDTA, pH 7.6.

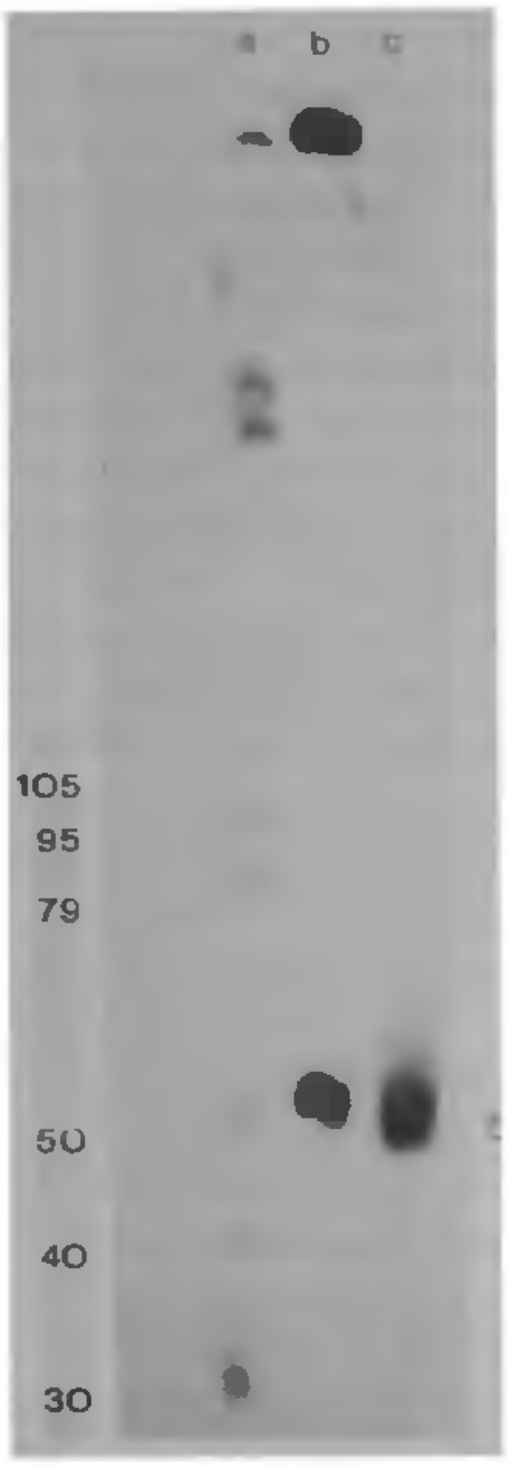

Figure 3. Confirmation of the double-stranded nature of the annealed synthetic oligonucleotides Annealed product was ^{32}P-labelled at 5'-OH and electrophoresed on an 8% polyacrylamide gel. Lane a, marker pBR322 cut with *Sau*3AI; b, pUC19 cut with *Eco*RI/*Hin*dIII; c, annealed 51-mer duplex.

indolyl β-D-galactopyranoside (X-Gal; Sigma) and 2 mM IPTG. Colonies of pUC19-containing cells showed deep blue colour after overnight incubation at 37°C. However, the colonies of pSBC1-containing cells

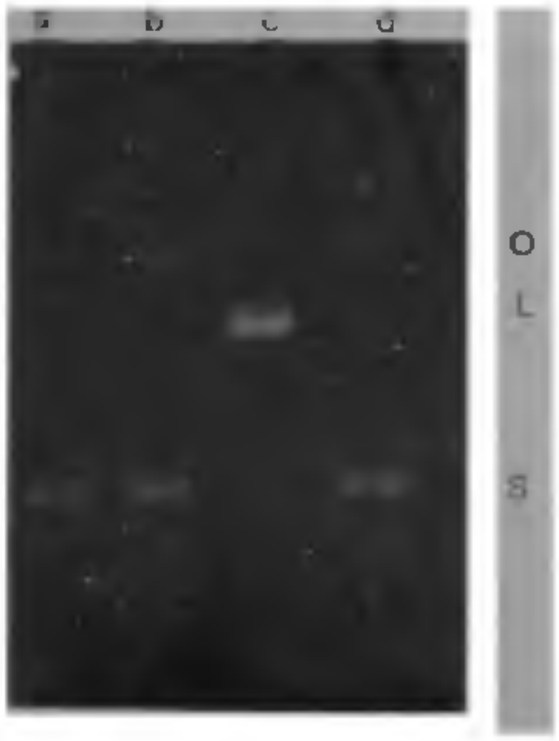

Figure 4. Characterization of pSBC1 clone Plasmid pSBC1 was digested with various restriction enzymes and run on 1% agarose gel in 1XTAE buffer, pH 8.3. O, Open circle, L, linear; S, supercoiled Lane a, Control, b, *Pst*I, c, *Xho*I; d, *Bam*HI.

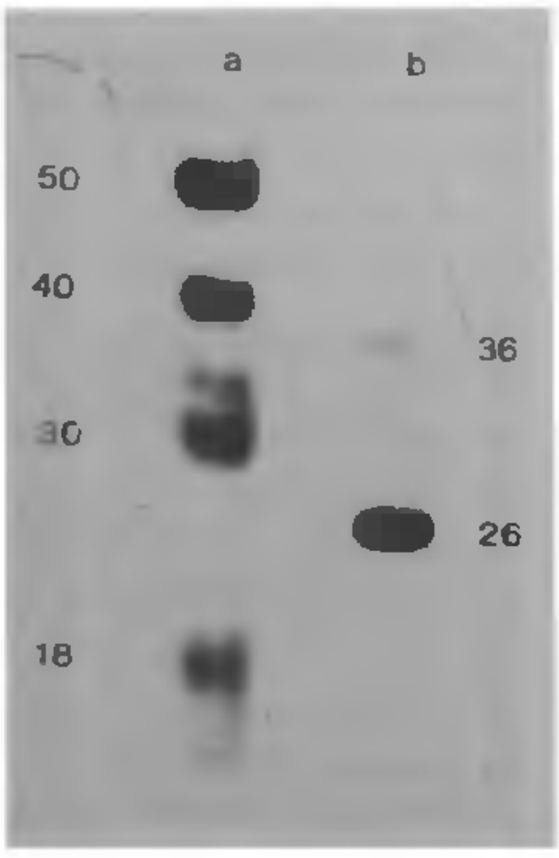

Figure 5. Restriction mapping of the insert region of pSBC1. Plasmid pSBC1 was fully digested with *Hin*dIII and partially digested with *Xho*I. After labelling the recessive termini products were electrophoresed on a 12% polyacrylamide gel. Lane a, marker pBR322 cut with *Sau*3AI, b, pSBC1 cut with *Hin*dIII and *Xho*I (partial).

Figure 6. *In vivo* expression of β-galactosidase on X-Gal-containing plates (*a*) after 12 h incubation, (*b*) after 72 h incubation

were white (Figure 6). Further incubation for three days resulted in light blue colour in the case of colonies of pSBC1-containing cells.

β-Galactosidase activity in cell extracts

Activity of β-galactosidase was measured (Figure 7) in whole-cell extracts at different times following induction with IPTG. Specific activity of β-galactosidase was also measured in whole-cell extracts after 6 h of induction and the data are shown in Table 2. β-Galactosidase activity was 28-fold lower in cells harbouring pSBC1 compared to those harbouring pUC19.

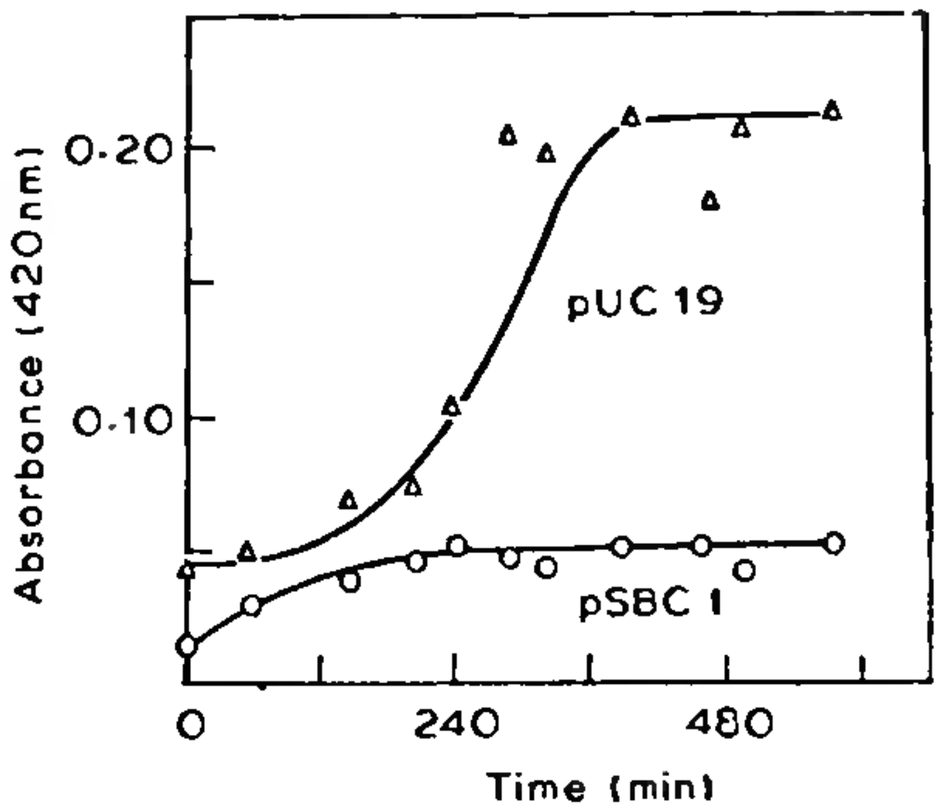

Figure 7. Comparison of β-galactosidase activity of *E. coli* cells harbouring pUC19 (△) or pSBC1 (○).

Table 2. Specific activity of β-galactosidase in whole-cell extracts

Clone	Specific activity* (units hr mg protein)
pUC19	350,450 (100)
pSBC1	12,520 (3.6)

*Numbers in parenthesis are per cent activity

Discussion

We have used the β-galactosidase gene as a model system to test our hypothesis[4,10] that alteration of the secondary structure of the DNA template *in vivo* could influence gene expression in *cis*. We have demonstrated by a novel approach that, by introducing a potential cruciform structure within the β-galactosidase gene carried on a recombinant plasmid (pSBC1), the expression of the gene can be down-regulated, as assessed by decreased enzyme activity in whole-cell extracts. Our earlier studies on pBR322 form V DNA, a topologically unlinked, highly supercoiled molecule with unusual structures around or within the coding region[20,21], have shown that cruciform structure could block elongation by *E. coli* RNA polymerase *in vitro*[4]. Formation of attenuated transcripts from the *rep* gene, with RNA polymerase pausing before the cruciform structure within pBR322 form V DNA template, was convincingly demonstrated[4]. Our present studies suggest that a cruciform structure in the 5′ coding region of the β-galactosidase gene could be blocking *E. coli* RNA polymerase, resulting in low level of β-galactosidase expression *in vivo*. Presently we do not know whether, in the case of pSBC1, RNA polymerase is stopping before the cruciform region and generating attenuated transcripts, or RNA polymerase movement is hindered by the cruciform structure, resulting in the low level of transcript formation. Dot-blot analysis using 5′ and 3′ probes with respect to the inverted repeat and S1 nuclease hybridization studies are in progress to substantiate these observations.

It is known that, in prokaryotes, transcription elongation of sequences with a GC-rich hairpin loop followed by stretches of U residues could lead to *rho*-independent termination[22]. To look for the effects of a DNA cruciform structure in transcription regulation, care was taken to eliminate, within the redesigned sequence, the characteristic features of *rho*-independent terminators. To rule out the possibility of RNA hairpin loop structure-mediated translational attenuation[23] leading to low-level β-galactosidase expression in pSBC1-harbouring cells, we introduced a left-handed Z-potential sequence between codons 7 and 13 of the β-galactosidase gene. Interestingly, this has also led to decreased β-galactosidase expression *in vivo* (unpublished results). Since, in this case, nascent RNA cannot adopt a hairpin structure, it is probable that natural DNA sequences capable of adopting unusual DNA structures within the transcribing region of the gene could down-regulate the expression of the gene.

Our results provide a possible mechanism by which unusual DNA structures within a gene could act in *cis* to regulate the expression of the gene *in vivo*. Transcription-elongation control has been observed in the case of several genes[24–29], e.g. *hsp70* gene of *Drosophila*, c-*myc*, c-*myb*, c-*fos* genes of mouse and man, and late genes of SV40 and adenovirus. In fact deletion analysis has shown that in the case of c-*myc* a 95-bp region containing an inverted repeat may be responsible for attenuation[25]. But due to the instability of the truncated RNA it could not be concluded whether the attenuation involved recognition sequences or structures within the RNA transcript or in the DNA template. Furthermore, abundance of potential cruciform structures within coding regions could reflect a biological significance[30]. The existence of a cruciform structure *in vivo* has been demonstrated recently using T7 endonuclease as a probe[31]. Observations of cruciform-DNA-binding protein also substantiate the possibility of occurrence of stabilized cruciform structures *in vivo*[32]. Our observations suggest that mutations in the stem region could destabilize the cruciform structure and lead to enhanced expression. This provides a mechanism by which loss of regulation could be envisaged through point mutations or deletions.

Stability of cruciform structure in a plasmid *in vitro* and *in vivo* has been shown to be dependent on the level of supercoiling of the plasmid[33]. Supercoiling can be modulated by several factors, including the activity of DNA gyrase versus that of DNA topoisomerase, level of transcription, and interaction of the template with proteins[34–36]. Thus the transition between cruciform and transcribable B-form could be easily facilitated *in vivo*. Such a mechanism could provide rapid control of gene expression. The usefulness of such a mechanism can be envisaged in the case of certain cellular responses that require rapid modulation in the expression of particular genes, e.g. heat-shock genes.

The implications of our result are far-reaching. For instance, poor expression of certain eukaryotic genes in prokaryotes could be due to the occurrence of such potential secondary-structure elements within the transcribable regions. In such cases expression of these genes could possibly be enhanced by a reverse of our approach, i.e. redesigning the genes in such a way that secondary-structure elements within the coding region are eliminated. Thus the approach described in this paper opens up the possibility of studying the role of various DNA secondary-structure elements in *cis* in the regulation of transcription *in vivo*.

Additional note

Since the submission of the manuscript we have reconstructed plasmid pSBC1 by interchanging the codons for Ser at 10 and 21, for Thr at 11 and 12, and for Pro at 16 and 19 (shown below). This mutant plasmid, pSBmC1, codes for the same amino acids as pSBC1, but does not have the inverted-repeat sequence which can adopt cruciform structure owing to the codon reshuffling. *E. coli* DH5α cells harbouring this plasmid show complete reversal of the suppression of β-galactosidase activity *in vivo* seen in cells harbouring pSBC1. This observation conclusively supports our proposition and shows that codon-usage bias has no role in decreased expression of β-galactosidase gene in pSBC1.

pSBC1

	Ser	Leu	His	Ala	Cys	Ser	Thr	Thr	Leu	Glu	Asn	Pro	Arg	Val	Pro	Ser	Ser	
5	AGC	UUG	CAU	GCC	UGC	UCG	ACG	ACA	CUC	GAG	AAU	CCU	CGA	GUG	CCG	UCG	AGC	3
		5		7		9		11		13		15		17		19		21

pSBmC1

5 AGC UUG CAU GCC UGC AGC ACA ACG CUC GAG AAU CCG CGA GUG CCU UCG UCG 3
5 7 9 11 13 15 17 19 21

The sequence synthesized and used to construct pSBmC1 by replacing the *EcoR*I *Hind*III fragment of pUC19

5 AGC TTG CAT GCC TGC AGC ACA ACG CTC GAG AAT CCG CGA GTG CCT TCG TCG-3
7 9 11 13 15 17 19
3 AC GTA CGG ACG TCG TGT TGC GAG CTC TTA GGC GCT CAC GGA AGC AGC TTA A 5

1. Darnell, J. E., *Nature*, 1987, **297**, 365.
2. Struhl, K., *Annu. Rev. Biochem.*, 1989, **58**, 1051.
3. Proudfoot, N. J, *Trends Biochem. Sci*, 1989, **14**, 105.
4. Bagga, R., Ramesh, N. and Brahmachari, S. K., *Nucleic Acids Res.*, 1990, **18**, 3363.
5. Bentley, D L. and Groudine, M, *Cell*, 1988, **52**, 245.
6. Mok, M., Maderious, A. and Chen-Kiang, S., *Mol. Cell. Biol.*, 1984, **4**, 2031.
7. Peck, L. J. and Wang, J. C., *Cell*, 1985, **40**, 129
8. Kinnıburgh, A. J., *Nucleic Acids Res.*, 1989, **17**, 7771.
9 Horowitz, M S. and Loeb, L. A, *Science*, 1988, **241**, 703.
10. Brahmachari, S. K. *et al.*, in *Structures and Methods* (eds Sarma, R H. and Sarma, M. H), Adenine Press, New York, 1990, vol. 2, p. 33.
11 Brahmachari, S. K., Mishra, R. K., Bagga, R. and Ramesh, N., *Nucleic Acids Res.*, 1989, **17**, 7273
12. Mishra, R. K., Latha, P. K. and Brahmachari, S. K., *Nucleic Acids Res.*, 1988, **16**, 4651.
13. Sambrook, J., Fritsch. E. F. and Maniatis, T., eds., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989
14. Miller, J. H., ed, *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1972.
15. Bradford, M. M, *Anal. Biochem*, 1976, **72**, 248.
16 Singelton, C. K, *J. Biol. Chem.*, 1983, **258**, 7661.
17. Courey, A. J. and Wang, J. C., *J. Mol. Biol.*, 1988, **202**, 35.
18 Thomas, D Y., Dubuc, G. and Narang, S., *Gene*, 1982, **19**, 211.
19 Grosjean, H. and Feirs, W., *Gene*, 1982, **18**, 199.
20. Brahmachari, S K., Shouche, Y. S., Cantor, C. R. and McClelland, M., *J. Mol. Biol.*, 1987, **193**, 201.
21. Shouche, Y. S., Ramesh, N. and Brahmachari, S. K., *Nucleic Acids Res.*, 1990, **18**, 267.
22. Platt, T., *Annu. Rev Biochem.*, 1986, **55**, 339
23. Landick, R., Carey, J. and Yanofsky, C, *Proc Natl. Acad. Sci. USA*, 1987, **84**, 1507.
24. Rougive, A. E. and Lis, J. T., *Cell*, 1988, **54**, 795
25. Wright, S. and Bishop, J. M., *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 505.
26. Bender, T., Thompson, C. and Kuehl, W., *Science*, 1987, **237**, 1473.
27. Piechaczyk, M, Blanchard, J. M. and Jeanteur, P., *Trends Genet.*, 1987, **3**, 47.
28 Hay, N., Skolnik-David, H and Aloni, Y., *Cell*, 1982, **29**, 183.
29 Hawley, D. K. and Roeder, R. G, *J Biol. Chem.*, 1985, **260**, 8163.
30. Muller, U. R. and Fitch, W. M., *Nature*, 1982, **298**, 582
31. Panayotatos, N. and Fontaine, A, *J Biol. Chem.*, 1987, **262**, 11364.
32. Bianchi, M. E., *EMBO J.*, 1988, **7**, 843
33. Panayotatos, N. and Wells, R. D., *Nature*, 1981, **289**, 466.
34 Pruss, G. J and Drlica, K., *Cell*, 1989, **56**, 521.
35. Liu, L. F. and Wang, J. C., *Proc. Natl. Acad Sci. USA*, 1987, **84**, 7024.
36. Boroweic, J. A, Dean, F B., Bultock, P. B. and Hurwitz, J., *Cell*, 1990, **60**, 181.

ACKNOWLEDGEMENTS. P. S S. and R. B are recipients of CSIR research fellowships. Financial support from Director, IISc, and the Department of Biotechnology towards 'Central Facility for Oligonucleotides Synthesis' is gratefully acknowledged We thank Prof. M. R. N Murthy for helping in the development of the computer program and 'Distributive Information Centre', IISc, for providing computational facilities; Ms K. Usha and Mr P. Ramasamy for technical help and for efficient typing of the manuscript; and all the other members of S K. B.'s lab for helpful discussions

Received 31 October 1990; revised accepted 2 January 1991

RESEARCH COMMUNICATIONS

Cloning of the catalytic determinant of soluble phosphofructokinase of yeast

Z. Lobo, R. Mehta and P. K. Maitra

Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India

The soluble phosphofructokinase of *Saccharomyces cerevisiae* is composed of two polypeptides α and β encoded by two unlinked genes *PFK1* and *PFK2* respectively. We report here cloning of the gene *PFK1* by complementation of a glucose-negative double mutant *pfk1 pfk2* using a genomic library of yeast in a multicopy vector. The anaerobic growth of the transformants on glucose is also rescued when the plasmid carries *PFK1* but not when it carries *PFK2*. The gene *PFK1* alone restores soluble-enzyme activity even if the resident *pfk2* allele is mutant. This suggests that the α subunit is the catalytic determinant of the enzyme.

GENETIC and biochemical studies on the soluble phosphofructokinase of yeast have defined two genes *PFK1* and *PFK2* involved in its synthesis. They determine, respectively, the two subunits α and β that constitute this enzyme[1-3]. Both genes have been cloned and sequenced[4,5]. Studies with a large number of mutants isolated by us have shown that mutations in *PFK1* extinguish soluble-enzyme activity without affecting aerobic growth on glucose. A second mutation in *PFK2* renders *pfk1* mutants glucose-negative[2,6,7] owing to the loss of a particulate phosphofructokinase. *pfk2* Mutants themselves retain the soluble-enzyme activity. This asymmetry in the function of the two gene products has been resolved by the observation that mutations in *PFK2* affect only the regulatory properties of the enzyme, leaving its catalytic capability intact[8]. (In the genetic nomenclature used here, *PFK1* refers to the gene for the catalytic α subunit and *PFK2* to the gene