Volume 11 Number 10 1983 Nucleic Acids Research

Targeted random mutagenesis: the use of ambiguously synthesized oligonucleotides to mutagenize sequences immediately 5' of an ATG initiation codon*

M.D.Matteucci and H.L.Heyneker

Department of Molecular Biology, Genentech, Inc., 460 Pt. San Bruno Boulevard, South San Francisco, CA 94080, USA

Received 25 February 1983; Revised and Accepted 25 April 1983

ABSTRACT

The nine base pairs immediately 5' of the initiation codon for the bovine growth hormone (BGH) structural gene have been mutagenized. The mutagenesis method employs the ligation of an ambiguously synthesized oligonucleotide duplex into a previously engineered gap in an expression plasmid for BGH. The mutation method, coupled with hybridization screening, is efficient at isolating 1 and 2 base pair changes within the targeted region. The E. coli cultures harboring the mutant plasmids were assayed for relative levels of BGH expression. The most notable result is the varied effect of substitution of G into the mRNA at various positions in this region.

INTRODUCTION

Translation initiation of mRNA in bacteria requires recognition of a specific region of the RNA by the ribosome and various protein factors (1). The Shine and Dalgarno (S.D.) sequence, for instance, interacts with the 3' end of the 16S ribosomal RNA (1,2). The distance between the S.D. sequence and the initiation ATG (spacer region) plays an important role in efficient translation (3,4). Recent attention has focused on the nonrandom nature of sequences surrounding the initiation ATG (5,6,7). Work presented by deBoer (8) and in this paper suggest that the sequence of the spacer region is also very important. The scope of this paper is to introduce a precisely targeted random mutagenesis method and apply it toward mutagenizing the spacer region of a plasmid expressing bovine growth hormone (BGH) (9).

The desire was to isolate transversions and transition mutations in a defined region 5' of the initiation ATG for BGH. Current methods of site directed random mutagenesis, while very powerful, appeared to be somewhat imprecise in the targeting of the mutagenesis (10,11,12). A recent method utilizing heteroduplex loop formation and bisulfite treatment suffers from only being capable of C to T and G to A changes (13). The method presented

here is precise in its targeting of random point mutations and is not limited in the possible base pair changes.

MATERIALS AND METHODS

Restriction digests were performed using commercial enzymes. The enzymes were used as directed by the suppliers. The S1 nuclease digests were done on 10 µg of XbaI cut DNA in 50 µl of 30mM sodium acetate (pH 4.6), 50mM NaCl, 1mM $ZnSO_4$ using S1 at a concentration of 100 U/ml (Miles Labs). Incubations were done for 15 minutes at room temperature (4). The Klenow exonuclease reaction was carried out on 5 µg of plasmid in 50 µl using 8 units DNA polymerase I Klenow fragment (NEN) and dGTP at 25 µM. The incubation was for 15 minutes at 37°C in 20 mM Tris HCl (pH 7.5), 7 mM $MgCl_2$, 60 mM NaCl and 6 mM 2-mercaptoethanol.

Ligations were performed by combining tailored plasmid vector (1 µg) and kinased oligomers in 1 to 10 molar ratio in 20 µl of ligase buffer (50 mM Tris HCl (pH 7.5), 6 mM $MgCl_2$ 5 mM DTT), heating to 65° for 5 min., quick cooling to 4°C and adding ATP to 1 mM and 1 unit ligase (BRL). After an overnight 4°C incubation, the reaction was used to transform E. coli K12 Strain 294 (14).

Synthetic oligomers bearing ambiguous bases were manually synthesized on a cellulose support by monomer addition triester chemistry by previously described methods (15). The hybridization probe was synthesized on an automated synthesizer (SAM-Biosearch) using previously described chemistry (16,17). Hybridizations were carried out as described (18). Sequencing was performed using the Sanger dideoxy-method directly on plasmid DNA (19,20).

Bovine growth hormone was assayed in E. coli extracts made from cultures grown in minimal media (9) from a 1/100 dilution of a stationary phase Luria Broth cultures. All media contained 5µg/ml tetracycline. Cells were harvested in log phase at approximately $A_{550} = 1$.

Whole cell extracts made by the boiling SDS procedure (14) were run on SDS-polyacrylamide 12.5 percent gels (21). Radioimmune assays (RIA) specific for BGH were performed on cell extracts prepared by sonication in a buffered 5 percent SDS solution (9).

RESULTS

The gene coding for BGH had been cloned and expressed in E. coli under the control of the trp promoter and the trp leader S.D. sequences. A plasmid had been constructed which introduces a unique XbaI restriction site

between the S.D. sequence and the initiation ATG (9). This allows for enzymatic manipulation and the insertion of a synthetic oligonucleotide mutant duplex pool into this region.

Figure 1 shows the scheme used to generate mutations all within the 9 basepair synthetic duplex. After digestion with XbaI and enzymatic removal of the sticky ends with S_1 nuclease, new (asymmetric) sticky ends were generated by limited digestion using the 3'-exonuclease activity of the Klenow fragment of DNA polymerase I in the presence of dGTP. This allowed for the efficient insertion of the ambiguously synthesized duplexes in a defined orientation.

The choice of the "wild type" sequence (starting point for mutagenesis) of the oligomer duplex was based on previous work in which this spacer region was shown to result in easily measured levels of BGH expression. The synthesis of the oligomers for the mutant duplexes was designed such that

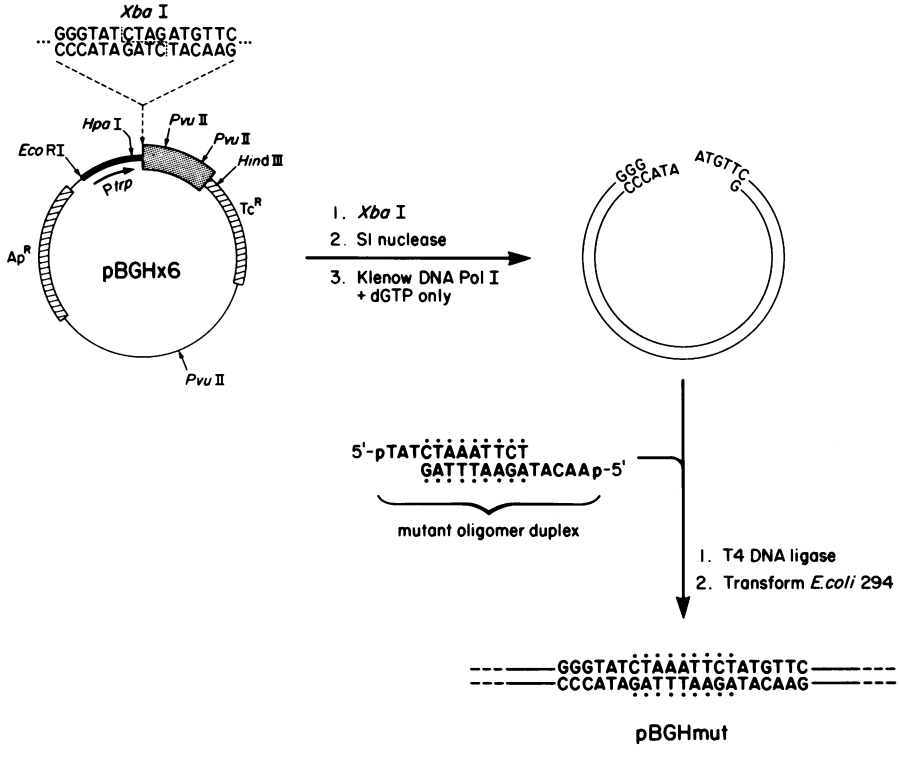

Figure 1. Scheme for targeted random mutagenesis.

for each condensation cycle corresponding to positions marked with an asterisk (Fig. 1) 75 percent of the indicated "wild type" nucleoside and 8 percent of each of the other three were added, whereas positions not marked were composed of 100 percent of the indicated wild type base. Assuming equal coupling kinetics and purity of the monomer blocks added, the resulting oligomer would be $.75^9$ or 7.5 percent "wild type," the rest would be mutant at one or more positions within the nine base pair spacer region. It was further assumed that mismatches of up to two basepairs between complementary strands would still allow duplex formation under annealing conditions yielding molecules which could serve as a substrate for ligation.

After ligation and transformation of the ligation product into competent E. coli 294, the cells were incubated for 2 hrs. in Luria Broth before plating to allow mismatches to be replicated and segregated. Plasmid DNA isolated from randomly selected ampicillin resistant colonies was analyzed by restriction digest (HpaI-PvuII) and polyacrylamide gel electrophoresis. Approximately 50 percent had synthetic DNA insertions.

A hybridization screening procedure was employed which allowed the convenient discrimination between wild type inserts (strong hybridization), 1 and 2 base pair change mutants (weak hybridization) and colonies bearing no insert (no hybridization). A 15mer probe, 5'-TATCTAAATTCTATG was synthesized. This probe is completely homologous to the wild type insertion sequence. Colony hybridizations were done on 400 randomly selected colonies (18). 48 negative hybridizers from the lysed colony hybridizations were rescreened using plasmid DNA purified by the alkaline SDS method (Fig. 2) (22). Purified DNA was necessary to reduce the background in the 4°C wash experiment. Twenty-one intermediate hybridization clones were sequenced through the region of interest (20). They all contained 1 or 2 base pair mutations (Fig. 3). To check whether some such mutations were missed by the above screening procedure, all non-hybridizing clones from the experiment shown in Fig. 2 were subjected to restriction analysis (HpaI-PvuII). Only one contained an insert which upon sequencing contained 3 base pair changes (Mutant 4) from wild type. This suggests that the hybridization screen on purified plasmid DNA is highly efficient at detecting 1 or 2 base pair changes from the wild type sequence. However, it is possible that some 1 base pair changes could have been among the strong hybridizers (Fig. 2) and consequently would have been missed.

E. coli 294 bearing mutant plasmids were assayed for BGH production.

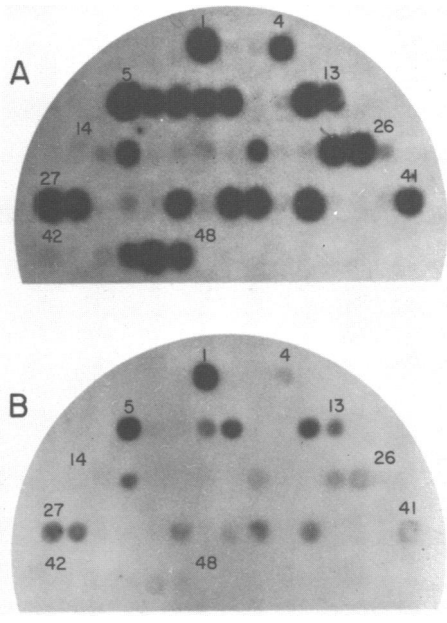

Figure 2. Hybridization Comparison. Film A results from binding of denatured purified plasmid from the 48 colonies described in Methods. The filter was hybridized with probe in 6X NET, 5X Denhardts, 250 µg/ml tRNA, 0.5 percent NP40 for 18 hrs at 4°C. Washings were performed 4 times at 4°C with 6X SSC and the autoradiogram exposed. Film B was obtained in a similar manner, with the only difference being that the washings were performed at 37°C. Plasmids No. 1 and No. 5 were wild type sequence controls.

Assays were done qualitatively by analysis of whole cell protein extracts on SDS PAGE gels (21) and by a specific radioimmune assay for BGH (9). The results are summarized in Figure 3. There is good qualitative agreement between the SDS gel data (not shown) and the RIA's. One distinct up mutation was discerned. This is a C → T change at position -9. Down mutations, however, were more frequent. Most notably T to G changes at positions -1 or -3 were found to lead to a large decrease in expression. Most mutations had little to no effect on BGH expression. Interestingly, a change to G at -2, -5, -6, -7, or -8 had little effect.

DISCUSSION

The aim of the mutagenesis method described here was to randomly mutate all positions in the 9 base pair region immediately 5' to the ATG initiation

Mutant Number	Position 5' of ATG										Relative RIA Activity	Dot Number in Fig. 2
	-9	-8	-7	-6	-5	-4	-3	-2	-1			
1	T	–	–	–	–	–	–	–	–		1.5	28
2	–	–	–	–	–	–	A	–	–		1.3	8
3	–	–	–	–	–	–	–	T	–		1.1	13
4	–	C	G	–	–	–	A	–	–		1.0	48
wildtype GGGTAT	C	T	A	A	A	T	T	C	T	ATG	1.0	1, 5
5	–	–	–	G	–	–	–	A	–		1.0	7
6	–	G	–	–	–	–	–	–	–		1.0	35
7	–	–	–	–	C	–	–	–	–		.8	24
8	–	–	–	–	T	–	–	–	–		.7	21
9	–	–	–	–	–	–	C	–	–		.7	9, 12
10	–	–	–	T	–	–	–	–	–		.7	16
11	–	–	–	G	–	–	–	–	–		.7	32
12	–	–	–	–	–	–	–	A	–		.6	27
13	–	–	–	–	–	–	–	G	–		.6	34, 45
14	–	–	–	–	G	–	–	–	–		.6	37
15	–	–	G	–	–	–	–	–	–		.6	41
16	–	–	C	–	–	–	–	–	–		.5	25
17	–	–	–	–	–	–	–	–	G		<.2	46
18	–	–	G	–	–	–	G	–	–		<.2	6
19	–	–	–	–	–	–	G	–	–		<.3	4, 47

Figure 3. Correlation of sequence and relative BGH expression for isolated mutants.

codon. This goal was achieved in an efficient manner. All mutations identified were localized precisely within the desired region.

Figure 4 shows the distribution of mutations both in position and base pair changes. The bias toward mutagenesis at position -3 and against it at position -4 is difficult to explain. Positions -1 and -9 appear to be mutagenized only rarely. Possibly a terminal mismatch in the oligomer duplex selects against that duplex from being ligated into the vector or is efficiently repaired upon introduction to the cell.

The base pair change distribution has a clear bias toward A → G changes. This change could arise from either an A to G change in the upper strand base pairing to a T or a T → C change in the lower strand base

Position distribution:

T	A	T	C	T	A	A	A	T	T	C	T	A	T	G
0	0	0	1	2	5	4	3	0	6	4	1	0	0	0

Base change distribution:

T to A	2	C to A	2	A to T	3
C	3	T	2	C	2
G	4	G	1	G	7

Figure 4. Distribution of mutations within targeted region.

pairing with A. The former mismatch, a G-T base pair, is more likely to occur, because of the relative stability of such a mismatch (23). This would lead to a more stable duplex and consequently a more likely ligation event.

This mutagenesis method allows the saturation of a precise area with point mutations. The specificity of oligonucleotide synthesis allows for "random" mutagenesis in a very defined region. Any defined area could be studied using similar techniques. The hybridization screening detects genotypic mutations and doesn't depend on phenotypic changes for detection. This is important because it is useful to identify changes which do not affect phenotype.

The most curious feature about differences in expression levels between the spacer mutants is the effect of the substitution of G into the mRNA. Substitution of G at positions -1 or -3 resulted in significant down mutations while at positions -2, -5, -6, -7, -8, little change occurred. mRNA secondary structure predictions made according to the rules of Tinoco et al. (23) did not yield different spacer region structures for any of the described mutants or wild type sequences (results not shown). Arguments that such substitutions of G lead to spurious SD sequences and consequent inhibitory competition for the 3'-end of the 16S RNA of the ribosome are not supported by this data since there is no recognizable homology between the spacer region and 3'-end of the 16S RNA.

It is interesting to note that substitution of G at -3 creates a new out of frame ATG (Mutant 19, Fig. 3). As a test of whether this new ATG is responsible for the significant down mutation observed with mutant 19 the following mutation was constructed. A plasmid (mutant 20) bearing the sequence GGGTATCTAAAAGCTATG was constructed in an analogous fashion to the mutagenesis scheme outlined in Figure 1. E. coli (Strain 294) harboring this plasmid showed an intermediate level of BGH expression relative to the wild type plasmid and mutant 19 by qualitative assay on SDS PAGE gel (Fig. 5). This observation is consistent with the possibility that out of frame ATG inhibits correct translational initiation.

It is clear that there is a poison effect on expression of BGH when there is a G in the message at -1 and to a lesser extent -3. There are several cases of G in the -1 position relative to the ATG in E. coli genes (6). A plausible explanation might be that this is a means of limiting the translation of mRNA coding for a gene product needed in only limited amounts. Inefficient translation has been shown to be partially responsible

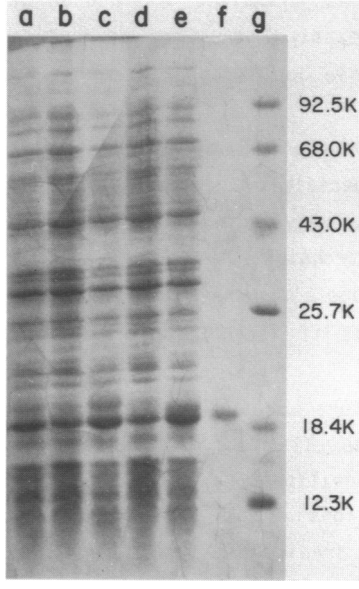

Figure 5. Laemmli gel comparing BGH expression in key mutants. Cultures of E. coli harboring plasmids were grown and prepared for electrophoretic analysis as described in Methods. The sequences of the plasmids in the region of interest are:

a) mutant 20 ----------AG-----
b) mutant 19, Fig. 3 ----------G-----
c) mutant 1, Fig. 3 ------T----------
d) mutant 17, Fig. 3 ----------G---
e) wild type GGGTATCTAAATTCTATG

Lane f is an authentic BGH standard.

for the low production of trp aporepressor in E. coli (25).

The up mutation (No. 1), a C → T change at position -9, creates a TTT region. Suggestions have been made that this can serve as the ribosomal protein S_1 binding site (24). It may be important that this site be adjacent to the S.D. sequence and not preceding the ATG as in mutant 3. Two other mutations appear to be up mutations: C to T (-2) and T to A (-3). The wild type sequence in this work is already efficient at the expression of bovine growth hormone. It is possible that translational initiation is no longer tne rate limiting step in the translational process and that some other parameter has become rate limiting. The weak up mutations seen in this system might actually become more dramatic in a system in which transcription is dependent on a weaker promoter or in which the transcriptional unit is located in the E. coli chromosome as a single copy.

ACKNOWLEDGMENTS

We thank Michelle Sanda for performing the BGH radioimmune assays and the Organic Chemistry Department for the synthesis and purification of the oligonucleotides.

*Genentech Contribution No. 158

REFERENCES

1. Steitz, J.A. (1978) In Biological Regulation and Control (Plenum Publishing Co., R. Goldberger, Ed.), Vol. 1.
2. Shine, J. and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342-1346.
3. Roberts, T.M., Kacich, R., and Ptashne, M. (1979) Proc. Natl. Acad. Sci. USA 79, 760-764.
4. Shepard, H.M., Yelverton, E., and Goeddel, D.V. (1982) DNA 1, 125-131.
5. Scherer, G.F.E., Walkinshaw, M.D., Arnott, S., and Morre, D.J. (1980) Nucleic Acids Research 8, 3895-3907.
6. Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B.S., and Stormo, G. (1981) Annual Review of Microbiology.
7. Stormo, G.D., Schneider, T.D., and Gold, L.M. (1982) Nucleic Acids Research 10, 2971-2995.
8. de Boer, H.A., Comstock, L.J., Hui, A., Wong, E., and Vasser, M. (submitted) Expression of Cloned Genes in Procaryotic and Eucaryotic Vectors (Elsevier/North-Holland, Inc., T.S. Papas, Ed.)
9. Seeburg, P.H., Sias, S., Adelman, J., de Boer, H.A., Hayflick, J., Jhurani, P., Goeddel, D.V., and Heyneker, H.L. (in press) DNA.
10. Shortle, D. and Nathans, D. (1978) Proc. Natl. Acad. Sci. USA 75, 2170-2174.
11. Weiher, H., Schaller, H. (1982) Proc. Natl. Acad. Sci. USA 79, 1408-1412.
12. Shortle, D., Grisafi, P., Benkovic, S.J. and Botstein, D. (1982) Proc. Natl. Acad. Sci. USA 79, 1588-1592.
13. Kalderon, D., Oostra, B.A., Ely, B.K., and Smith, A.E. (1982) Nucleic Acids Research 10, 5161-5171.
14. Goeddel, D.V., Heyneker, H.L., Hozumi, T., Arentzen, R., Itakura, K., Yansura, D.G., Ross, M.J., Miozzari, G., Crea, R., and Seeburg, P.H. (1979) Nature 281, 544-548.
15. Crea, R. and Horn, T. (1980) Nucleic Acids Res. 8, 2331-2348.
16. Matteuci, M.D. and Caruthers, M.H. (1981) Journal Am. Chem. Soc. 103, 3185-3189.
17. Beaucage, S.L. and Caruthers, M.H. (1981) Tet Letters 22, 1859-1862.
18. Grunstein, M. and Hogness, D.S. (1975) J. Cell Biol. 67, 835-862.
19. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
20. Wallace, R.B., Johnson, M.J., Suggs, S.Y., Miyoshi, K.I., Ghatt, R., and Itakura, K. (1981) Gene 16, 1-3.
21. Laemmli, U.K. (1970) Nature 227, 680-685.
22. Birnboim, H.C. and Doly, J. (1979) Nucleic Acid Research 7, 1513-1523.
23. Tinoco, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M., and Gralla, J. (1973) Nature New Biol. 246, 40-41.
24. Szer, W., Hermoso, J.M., and Leffler, S. (1975) Proc. Natl. Acad. Sci. USA 72, 2325-2329.
25. Kelley, R.L. and Yanofsky, C. (1982) Proc. Natl. Acad. Sci. USA 79, 3120-3124.