Volume 11 Number 16 1983 Nucleic Acids Research

Formation of genes coding for hybrid proteins by recombination between related, cloned genes in *E. coli*

Hans Weber and Charles Weissmann

Institut für Molekularbiologie I, Universität Zürich, 8093 Zürich, Switzerland

Received 4 July 1983; Accepted 19 July 1983

ABSTRACT

We describe a method for the formation of hybrid genes by in vivo recombination between two genes with partial sequence homology. DNA structures consisting of plasmid vector sequences, flanked by the α2 interferon gene on the one side and a portion of the α1 interferon gene (homology about 80%) on the other, were transfected into E.coli SK1592. Appropriate resistance markers allowed the isolation of colonies containing circular plasmids which arose by in vivo recombination between the partly homologous interferon gene sequences. Eleven different recombinant genes were identified, six of which encoded new hybrid interferons not easily accessible by recombinant DNA techniques.

INTRODUCTION

Twelve or more α-interferons are encoded in the human genome (for a review, see ref. 1), most of which are expressed to some degree (2,3,4). Some of these exhibit widely different antiviral activities on cultured cells of different animal origin (5,6,7). For example, the specific activity of interferon α1 is comparatively low on human cells but high on mouse cells, whereas the opposite is the case for interferon α2 (5,6,7). Genetic engineering techniques were used in vitro to construct hybrid interferon genes consisting of α1 and α2 specific sequences joined at either of two restriction sites present in both genes at homologous positions (5,6,7). The specific antiviral activities of hybrid interferons containing the C-terminal half of interferon α1 were high on mouse cells; on the other hand, high activity on human cells was dependent on the presence of the N-terminal half of interferon α2.

The types of recombinants that can be created by this approach are limited by the number of appropriate restriction

sites. As the nucleotide sequence homology between the genes for interferon α1 and α2 is about 80% it seemed likely that the recombination machinery of E.coli cells could be used *in vivo* to obtain recombinant genes with crossovers at any site showing sufficient homology.

In this paper we show that by an appropriate choice of constructions it is possible to generate with minimal effort a variety of hybrids not easily obtained by ordinary genetic engineering techniques. The positions of the desired crossovers can be directed to predetermined regions. The method should be generally applicable to the formation of recombinants between not too distantly related genes.

MATERIALS AND METHODS

1) Plasmids and bacteria.

Plasmids pM11 and pM21 (see description in Results and Discussion section) were constructed by Dr. M. Mishina (unpublished results). For the construction of plasmid pM11kan a derivative of plasmid pBR322 conferring resistance to tetracycline, ampicillin and kanamycin was first prepared. pBR322 was partially digested with HaeII in presence of 50 μg/ml ethidium bromide, to yield mostly full-size linear molecules. These were ligated with an equivalent amount of a complete HaeII digest of a pCRI plasmid (containing, for incidental reasons, a bacteriophage Qβ cDNA insert; kindly supplied by Dr. M. Billeter). Transfectants were selected for resistance against kanamycin and ampicillin, and subsequently tested for tetracycline resistance. For unknown reasons simultaneous selection for resistance against all three antibiotics yielded no clones. Restriction mapping showed that the 1430 bp HaeII fragment conferring kanamycin resistance had been inserted in pBR322 at the HaeII site in position 2352. Plasmid pM11kan was then constructed by ligating the following 3 components: (1) The 4400 bp PstI-SalI fragment of the kanamycin-resistant pBR322 derivative, (2) the 1000 bp ClaI-PstI fragment of pM11, (3) the ClaI-SalI fragment (about 800 bp) of a pBR322 derivative carrying a small insertion (a small lambda DNA fragment linked by poly(dG:dC) tracts) in the BamHl site. The 3 fragments were isolated by electrophoresis on 1% low-melting agarose gels

(SeaPlaque agarose, FMC Corporation). Aliquots of the gel slices were melted at 65° and used directly for the ligation reaction (8). Transfectant colonies were selected for kanamycin resistance and tested for tetracycline and ampicillin sensitivity. The structure of the desired plasmid pM11kan was confirmed by restriction mapping.

E.coli SK1592, a phage T_1-resistant derivative of strain SK1590 (9), was obtained from Dr. S. Kushner (University of Georgia) and E.coli 803 from Dr. K. Murray (University of Edinburgh). Enzymes were purchased from New England Biolabs Inc. and labeled nucleotides from Amersham.

2) In vivo recombination.

The SalI-PstI fragment (1580 bp) of pM21 and the SalI-BglII fragment (4970 bp) of pM11kan were isolated by electrophoresis on low melting agarose (1.2% and 0.8%, respectively) in 40 mM Tris-acetate, 1 mM EDTA (pH 7.8). Ligation was in reaction mixtures (20 µl) containing 20 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP, 10-20 fmol of each DNA fragment (2-5 µl aliquots of agarose slices melted at 65°, final concentration of agarose 0.35%) and DNA ligase (400 units, New England Biolabs) at 16° for 18 h. Aliquots (2.5 µl) were used in transfection mixtures (100 µl) containing 10 mM $CaCl_2$, 10 mM $MgCl_2$, 10 mM Tris-HCl (pH 7.5), and 65 µl of a suspension of $CaCl_2$-treated E.coli (prepared according to ref. 10 and stored in 50 mM $CaCl_2$, 20% glycerol at -80°). The suspensions were kept at 0° for 15 min, at 42° for 2 min, then diluted with L-broth (1 ml) and incubated with shaking at 37° for 90 min. Aliquots were spread on agar plates containing L-broth plus kanamycin (50 µg/ml) and tetracycline (20 µg/ml). The number of colonies varied between 1'600 and 90'000 per pmol DNA fragments; control plates with tetracycline alone gave 5-12x10^7 colonies/pmol pBR322. Plasmids were prepared by modifications of published methods (11,12). DNA sequences were determined as described by Guo and Wu (13), using the EcoRI, BglII or PvuII sites as first or second cleavage sites.

RESULTS AND DISCUSSION

The principle of the experiment is outlined in Fig. 1. The two parental interferon genes (or overlapping portions thereof)

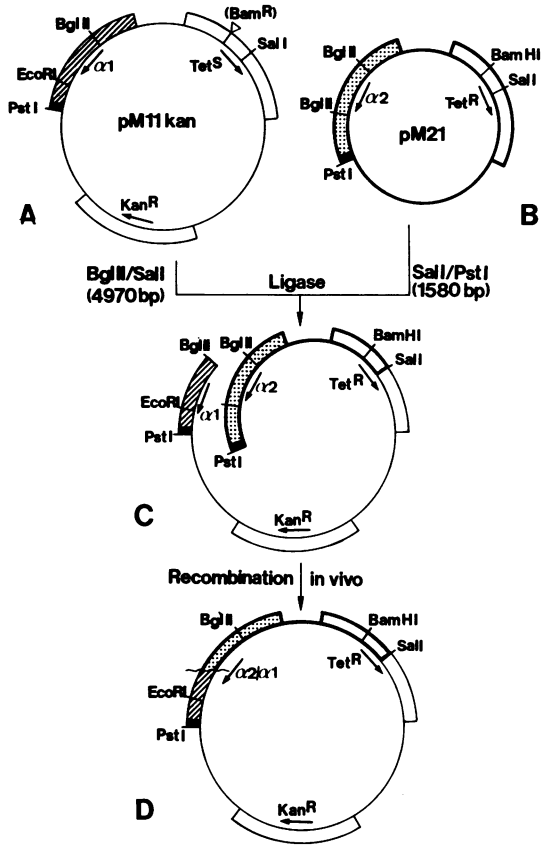
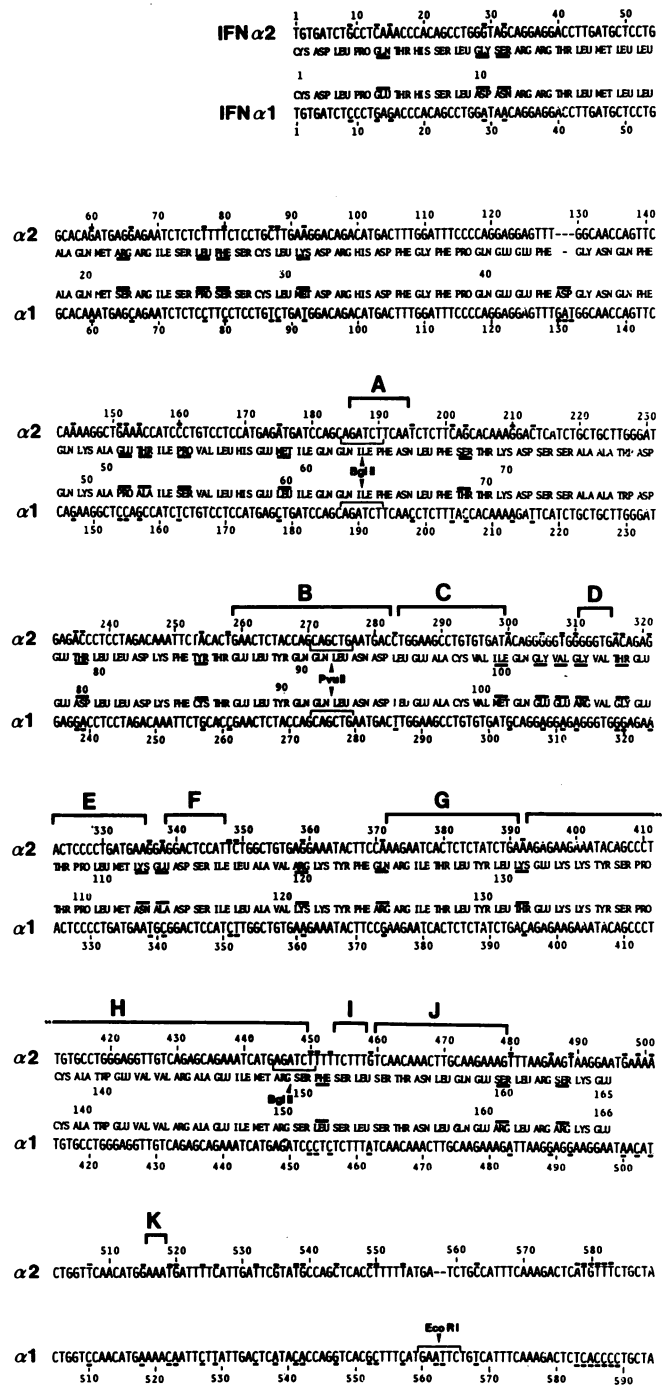


Figure 1. Construction of DNA amenable to recombination in vivo to yield α2-α1-interferon hybrid genes. A. Plasmid pM11kan is a pBR322-derived expression plasmid for interferon α1 production in E.coli. It contains a DNA segment conferring kanamycin resistance in position 2352 and has a small insertion in the BamH1 site, which abolishes tetracycline resistance. The mature interferon α1 coding sequence is joined to the β-lactamase promoter and ribosome binding region and replaces the N-terminal portion of the β-lactamase gene. B. In plasmid pM21 the interferon α2 gene is inserted in the same way into pBR322 as α1 in pM11kan; the remaining pBR322 sequences are unmodified. C. Appropriate fragments from both parent plasmids are ligated to yield linear (probably concatenated) DNA structures linked via the SalI site, thus reconstituting the tetracycline resistance gene. Details are given in the text. D. General structure of the interferon α2-α1 recombinant plasmids. The dotted regions indicate α2-specific, the hatched regions α1-specific sequences. The black regions adjacent to the PstI sites indicate short poly(dG:dC) tracts (15-30 bp). Bam^R, destroyed BamH1 site; Tet^R, Kan^R, tetracycline and kanamycin resistance genes; Tet^S, inactivated tetracycline resistance gene.

with the vector sequences between them are supplied to the host cell as parts of a linear DNA structure. Circularization of such structures by recombination within the interferon genes leads to replicating plasmids. An appropriate arrangement of two antibiotic resistance genes allows the easy selection of recombinants.

The linear DNA structures (presumably concatenates of the molecules shown in Fig. 1C) were constructed by ligation of two restriction fragments, each derived from a plasmid containing one of the parental interferon genes and an antibiotic resistance gene. Plasmid pM21, constructed for high expression of interferon α2 in E.coli (M. Mishina, unpublished results), contained the DNA sequence coding for mature interferon α2, fused to an AUG triplet 6 nucleotides downstream from the Shine-Dalgarno sequence of the β-lactamase gene of plasmid pBR322. The interferon segment replaced the β-lactamase sequences up to the PstI site, to which it was joined by a poly(dG:dC) tract. The remainder of the plasmid was identical to pBR322, including the intact tetracycline resistance region. In the other parental plasmid, pM11kan, the interferon α1 sequence was inserted into pBR322 in exactly the same way as α2 in pM21. In addition, the kanamycin resistance gene of plasmid pCRI (1430 bp HaeII fragment; ref. 14) had been inserted into the HaeII site at position 2352 and the tetracycline resistance region inactivated by a small insertion into the BamHI site.

To obtain linear DNA structures suitable for *in vivo* recombination between the α1 and α2 genes, the SalI-PstI fragment (1580 bp) of plasmid pM21 was ligated to the SalI-BglII fragment (4970 bp) of plasmid pM11kan and the resulting DNA transfected into $CaCl_2$-treated E.coli SK1592. Selection on agar plates containing both tetracycline and kanamycin yielded only bacteria transformed by ligation products containing elements of both parental moieties linked correctly at the SalI site, as the proximal portion of the tetracycline resistance region has to be provided by the α2 parent plasmid whereas the distal portion of the tetracycline resistance region as well as the kanamycin resistance gene originate from the α1 parent. As the IFN gene segments were the only homologous regions on the linear ligation products, it was expected that circularization by recombination


IFN α2
IFN α1
α2
α1
A
B
C
D
E
F
G
H
I
J
K
PvuII
Bgl II
Eco RI

would occur predominantly if not exclusively between the BglII site (around position 190) and the end of the interferon DNA, i.e. the PstI site.

Plasmids from tetracycline and kanamycin resistant clones were subjected to restriction and sequence analysis. Of 63 clones analyzed, 62 appeared to have arisen by correct homologous recombination, i.e., without any gaps or insertions, as judged by restriction mapping and by the finding that 20 out of 20 cell-free extracts from the bacterial clones had levels of interferon activity similar to the α2 parent strain when tested on human HEp2 (CCL23) cells. Thirteen recombinants lacked the unique EcoRI site located downstream of the IFN-α1 coding sequence and probably contained entirely α2-specific coding sequences; they were not analyzed further. Forty-four of the remaining plasmids were sequenced between the upstream BglII site and the EcoRI site. As shown in Fig. 2, 11 different crossover regions were identified in this interval. The 11 recombinant sequences encoded 9 different interferons, 8 of which were hybrids. Only two of these had been obtained previously using conventional genetic engineering techniques (5,6,7).

Five of the clones yielded plasmid DNA preparations which were heterogeneous with regard to the downstream BglII site, one also with regard to the EcoRI site. Heterogeneity could be due to segregation from a heteroduplex recombination intermediate or to transfection by several copies of (possibly concatenated) recombination-competent DNA. Homogeneous plasmid preparations were obtained after retransfection and cloning, except for one subclone, where the BglII restriction pattern again indicated heterogeneity, with two components being present in about equal quantity. Since this DNA preparation consisted mainly of dimeric plasmids, it seemed likely that two different recombinants were linked in a tandem dimeric circle.

The number of recombinant plasmids recovered for each cross-

Figure 2. Location of 11 crossover regions (A through K) between the interferon α2 and α1 genes observed in the in vivo recombination experiments. Recombinants with crossovers in regions D, E, F, G, H and I/J code for new hybrid interferons. Crossovers in B and C as well as I and J give sequences differing only by silent nucleotide changes.

over region varied greatly, however, the experimental conditions used for selection did not ensure that all hybrids were independent isolates. The dependence of recombination frequency on the degree of homology, on the length of the homologous region and on specific sequence features remains to be determined. In any event, crossovers were found in regions with as few as five or even three bp of uninterrupted homology (Fig. 2, regions D and K). It should be noted however that the actual crossover point need not correspond to the region in which recombination is initiated, and that recombination might require regions of higher homology.

In additional experiments the actual recombination step between the plasmid components was carried out in vitro leaving only heteroduplex repair to take place in the host cell. Plasmid pM11kan was linearized at the BglII site (position 189); plasmid pM21 was partially digested with BglII to yield mostly linear full-size molecules, 60-70% of which were cut at the downstream BglII site (position 446) as shown by restriction analysis. Both linearized plasmid preparations were digested with T_4-DNA polymerase to convert the terminal 300-400 nucleotides into a single-stranded form; they were mixed in about equal proportion and annealed. Remaining gaps were filled in with DNA polymerase, the products were cleaved with restriction endonuclease SalI and circularized by ligase. Transfections were carried out into E.coli strains HB101 ($recA^-$), 803 (rec^+) and SK1592 ($recA^+$sbcB). Colonies resistant to both tetracycline and kanamycin were obtained from all three strains, but among 21 plasmids analyzed all but one, a hybrid generated in E.coli HB101 with a crossover in region C, were found by restriction mapping to contain deletions or otherwise rearranged sequences. It would seem that the in vivo process is both more efficient and simpler to carry out, but that in principle similar results could be produced in vitro.

The method we have described can be further refined to yield crossovers in predetermined regions. For example, if the linear concatemers are formed between the pM11kan SalI-BglII fragment and a pM21 fragment extending from the SalI to the PvuII site in position 273, recombination would be confined to

the 89 bp region between the BglII and the PvuII sites.

It will be of interest to determine how much homology between two genes is required to allow this type of recombination.

ACKNOWLEDGEMENTS

We thank Mr. Peter Meyer for competent technical assistance, Dr. David Valenzuela for analyzing several of the hybrid clones and Mr. Michel Streuli for $CaCl_2$-treated E.coli SK1592. This work was supported by the Schweizerische Nationalfonds and the Kanton Zürich.

REFERENCES

1. Weissmann, C., Nagata, S., Boll, W., Fountoulakis, M., Fujisawa, A., Fujisawa, J.-I., Haynes, J., Henco, K., Mantei, N., Ragg, H., Schein, C., Schmid, J., Shaw, G., Streuli, M., Taira, H., Todokoro, K. and Weidle, U. (1982) Phil. Trans. R. Soc. Lond. B299, 7-28.
2. Rubinstein, M., Levy, W.P., Moschera, J.A., Lai, C.-Y., Hershberg, R.D., Bartlett, R.T. and Pestka, S. (1981) Archs. Biochem. Biophys. 210, 307-318.
3. Allen, G. and Fantes, K.H. (1980) Nature 287, 408-411.
4. Goeddel, D.V., Leung, D.W., Dull, T.J., Gross, M., Lawn, R.M., McCandliss, R., Seeburg, P.H., Ullrich, A., Yelverton, E. and Gray, P.W., (1981) Nature 290, 20-26.
5. Streuli, M., Hall, A., Boll, W., Stewart II, W.E., Nagata, S. and Weissmann, C. (1981) Proc. Natl. Acad. Sci. USA 78, 2848-2852.
6. Weck, P.K., Apperson, S., Stebbing, N., Gray, P.W., Leung, D., Shepard, H.M. and Goeddel, D.V. (1981) Nucl. Acids Res. 9, 6153-6166.
7. Rehberg, E., Kelder, B., Hoal, E.G. and Pestka, S. (1982) J. Biol. Chem. 257, 11497-11502.
8. Frischauf, A.M., Garoff, H. and Lehrach, H. (1980) Nucl. Acids Res. 8, 5541-5549.
9. Kushner, S.R. (1978) In "Genetic Engineering" (H.W. Boyer and S. Nicosia, eds.) Elsevier North Holland, Amsterdam, pp. 17-23.
10. Cohen, S.N., Chang, A.C.Y. and Hsu, L. (1972) Proc. Natl. Acad. Sci. USA 69, 2110-2114.
11. Currier, T.C. and Nester, E.W. (1976) Anal. Biochem. 76, 431-441.
12. Holmes, D.S. and Quigley, M. (1981) Anal. Biochem. 114, 193-197.
13. Guo, L.-H. and Wu, R. (1982) Nucl. Acids Res. 10, 2065-2084.
14. Oka, A., Sugisaki, H. and Takanami, M. (1981) J. Mol. Biol. 147, 217-226.