An approach to construction of hybrid polypeptide molecules—homologue recombination method

Alexey A.Schulga*, Ilya V.Levichkin¹, Feruz T.Kurbanov¹, Andrei L.Okorokov,
Galina E.Pozmogova¹ and Mikhail P.Kirpichnikov

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov Street 32,
Moscow 117984 and ¹Center 'Bioengineering', Russian Academy of Sciences, Vavilov Street 34/5,
Moscow 117334, Russia

Received July 7, 1994; Accepted July 18, 1994

Mapping of functionally important sites (FIS) is a problem of great importance in the study of protein molecules. Success in gene cloning and enlargement of the number of cloned genes has allowed new insights into this problem. Proteins belonging to a single family may exhibit consistent functional peculiarities. These make it feasible to map the FIS by recombining fragments of homologous genes. This process entails substitution of desired segments in gene under study for homologous ones from another representative of the family and subsequently comparing the wild-type and hybrid protein (1,2). Because structurally important residues are usually conserved it is likely that hybrid molecules formed will have maintained a correct and stable conformation. In choosing the segment to be substituted one should consider both the currently available data on the protein and/or any theoretical predictions (2,3). Eventually one may choose to substitute a great many segments that may overlap in aggregate the whole gene. In general two techniques are used for this purpose—site-directed mutagenesis and PCR (1,4). These methods are time-consuming, elaborative and expensive, requiring the extensive chemical synthesis of oligonucleotides. In this article we present a new approach, the homologue recombination (HR) method, which allows a set of hybrid molecules to be produced in one experiment. The method is based upon the observation that the thermostable Vent-polymerase (New England Biolabs) which has 3'–5' proofreading exonuclease activity shortens the single-stranded DNA fragments synthesized in unidirectional PCR from their 3'-termini. An assortment of such fragments transcribed from a donor gene may be used as primers in the site-directed mutagenesis of a related acceptor gene. This will lead to the production of a number of hybrid genes. Segments of various lengths in the acceptor gene are substituted for homologous ones in the donor gene.

As an example this method has been successfully applied in the construction of hybrid barnase/binase molecules in our laboratory. These bacterial RNases, barnase and binase, are close natural homologues (84% identical residues) (5,6) and consist of a single polypeptide chain of 110 and 109 residues, respectively, without any disulfide bonds. Their 3D-structures were resolved at a high resolution (7–9). The proteins have very similar secondary and tertiary structures. However there are distinct dissimilarities in their kinetics and thermostability (10/11).

The study of hybrid barnase/binase molecules might allow causes of the dissimilarities to be elucidated. The scheme of the experiment is depicted in Fig. 1. The barnase gene was cloned into the MCS of M13 mp18 vector and uracil-containing single stranded template was purified (12,13). Single-stranded (ss) DNA copies of the binase gene were obtained by PCR in two steps. First the double stranded copies were prepared by amplification using two primers, Bin5: 5' CAAAAGCCGCCGTCATTAAT-ACGTTTG 3' and Bin3: 5' TCGTTATCGAATACGTGTGA-AAG 3', that flank the mature binase coding region from the 5' and 3' termini respectively. The reaction mixture (50 µl included: 65 pg of plasmid DNA with chromosomal binase gene (14) 50 pmoles of each of the two primers, 10 nmoles of each of the four dNTPs and 1 unit of Vent-polymerase in the 1× buffer supplied with the enzyme. The amplification was performed in a Bio-Med Thermal Cycler using a program that denatured at 96°C for 1 min, annealed at 60°C for 1 min and extended at 74°C for 30 sec, over a total 30 cycles. Amplified DNA was purified from agarose gel slices using Glassmilk (Geneclean, Bio 101). Second, the single-stranded DNA was prepared by unidirectional PCR using the isolated DNA (0.5 pmoles) as a template and phosphorylated oligonucleotide Bin3 (10 pmoles) as a primer. The experimental conditions were as described above. The reaction mixture was extracted with phenol/chloroform (1:1), saturated in 100 mM Tris–HCl, pH 8.0. DNA was precipitated by the addition of 1/10 volume 0.3 M sodium acetate and 3 volumes of ethanol at room temperature and sedimented at 12,000 g. The product of the unidirectional PCR was used as a primer in the subsequent step of site-directed mutagenesis according to Kunkel (13). One third part of the DNA was annealed to single-stranded uracil template containing the barnase gene by incubating the reaction mixture at 90°C for 2 min and subsequently cooling slowly to 30°C over the course of approximately 1 hour. The mutagenesis was conducted as described in the Bio-Rad MutaGene In Vitro Mutagenesis Kit manual. The *XL-1 Blue* strain was used for transformation. The DNA clones were isolated and sequenced using the Sanger method. A set of hybrid genes was accordingly obtained (Fig. 2). The following segments in the barnase (which is 110 residues in length) were found to be substituted for the binase ones: 1–110, 7–110, 12–110, 21–110, 26–110, 30–110, 37–110,

* To whom correspondence should be addressed

Figure 1. The scheme of HR-method as applied to the construction of hybrid barnase/binase molecules. The experiment comprised the following essential steps: (I) uracil template production with the acceptor-gene (barnase), amplification of (II) the double- and (III) single-stranded DNA copies of the donor-gene (binase) by PCR and finally (IV) site-directed mutagenesis according to Kunkel (13).

Figure 2. Diagrammatic representations of the hybrid barnase/binase molecules constructed by HR-method. Filled boxes are the binase portions of the molecules; open boxes are the barnase portions. On the left there are names of the hybrids. The names indicate the first and the last residue in the binase portion of the hybrid molecules. On the right there are the frequencies of the hybrids calculated from the sequence analysis of 25 random clones.

67–110, 73–110, 90–110, 93–110, 101–110 a.a. The efficiency of HR-method can be estimated from the results of sequence analysis of 25 random clones (Fig. 2). Non-mutant sequences were represented in 12 clones; in one case there was the substitution of a complete barnase gene for the binase one; hybrid genes were found in 12 clones.

Similar results have been obtained in constructing hybrid human/porcine growth hormones (GH) (Fig. 3). The sequences of these proteins are about 67% identical and their tertiary structures are similar (15–17). To our surprise, besides the hybrid molecules containing various N-terminal lengths of human GH linked to the remaining C-terminal portion of porcine GH

Figure 3. Structures of the hybrid human/porcine growth hormone molecules. Black bars are pGH domains; white bars are hGH domains. On the left there are names of the hybrids. They indicate the spanning of pGH domains. On the right there are the frequencies of hybrids calculated from the sequence analysis of 32 random clones. A) hybrid porcine/human GHs, type I; B, C) hybrid porcine/human GHs, type II.

and analogous to that obtained in the barnase/binase experiment (Fig. 3A, the hybrid proteins, type I) another type of hybrid molecule was found (Fig. 3B, C, hybrid proteins type II) in which in contrast the porcine GH region is spread from N-terminus. Type II hybrid molecules are presumed to be the result of 5'–3' exonuclease editing from the exposed 5'-end of the mutagenic primer *in vivo* (18). To all appearance in some cases the whole polynucleotide primer was eliminated giving yield to non-hybrid phenotype. Also the number of hybrid clones might be decreased because of the methyl-directed mismatch repair *in vivo*, correcting the unmethylated mutant strand in favor of the plus strand template (19). Thus the transfection of the heteroduplexes into repair deficient strains of *E. coli* should enhance the yield of hybrids. Elaboration of a more efficient protocol for the construction of hybrid molecules is the subject of ongoing investigation in this laboratory.

In conclusion, the method we present here could be applied in preliminary searches for functionally important amino acid residues. In practice the choice of the PCR primers is arbitrary and need not be confined to the gene termini. For example these sites may be located close to the regions under study. In succeeding stages more precise methods can be used, such as oligonucleotide-directed mutagenesis. The only requirement for application of the approach is the availability of at least two homologous genes encoding proteins with detectable difference in properties. Finally the method can also be used to express a number of related genes in a vector constructed for the expression of a single one. Potentially it allows manipulation of the coding regions alone, without disturbance of the flanking nucleotide sequences. For instance we were able to express the binase in *E. coli* by substituting the barnase coding region in the available expression vector (pMT702) for the binase one (20). To simplify selection of the clones in which the whole target region was replaced, PCR primers with incorporated restriction enzyme sites could be used.

ACKNOWLEDGEMENTS

The authors thank Dr R.W.Hartley (National Institute of Health, USA) for providing the plasmids pMT702, pMT416 and pMT316 and Prof M.Ya.Karpelsky (Engelhardt Institute of Molecular Biology, Russia) for stimulating discussions and Dr J.M.Gulbis (Rockefeller University, NY, USA) for help in manuscript preparation. This work was supported by Grants No. 102 and 103 from the Protein Engineering Council (Russia) and No. 93-04-21295 from Russian Foundation of Fundamental Research. The research described in this publication was made possible in part by Grant No. MF6000 from the International Science Foundation.

REFERENCES

1. Cunningham,B.C., *et al.* (1989) *Science* **243**, 1330–1336.
2. Schulga,A.A., *et al.* (1992) *Biotechnologiya (Russia)* **5**, 22–26.
3. Gabrielian,A.E., Schulga,A.A. and Kirpichnikov,M.P. (1992) *Mol. Biol. (Russia)* **26**, 617–623.
4. Jones,D.H. and Howard,B.H. (1990) *BioTechniques* **8**, 178.
5. Paddon,C.J. and Hartley,R.W. (1986) *Gene* **40**, 231–239.
6. Schulga,A.A., *et al.* (1992) *Nucleic Acids Res.* **20**, 2375.
7. Mauguen,Y., *et al.* (1982) *Nature* **297**, 162–164.
8. Pavlovsky,A.G., *et al.* (1989) *Krystallografia (USSR)* **34**, 137–142.
9. Bycroft,M., *et al.* (1991) *Biochemistry* **30**, 8697–8701.
10. Serrano,L., Day,A.G. and Fersht,A.R. (1993) *J. Mol. Biol.* **233**, 305–312.
11. Makarov,A.A., *et al.* (1993) *J. Biomol. Struct. Dyn.* **10**, 1047–1065.
12. Hartley,R.W. (1988) *J. Mol. Biol.* **202**, 913–915.
13. Kunkel,T.A., Roberts,J.D. and Zakour,R.A. (1987) *Methods Enzymol.* **154**, 367–382.
14. Nurkiyanova,K.M., *et al.* (1989) *Dok. Akad. Nauk SSSR (USSR)* **309**, 1476–1479.
15. Abdel-Meguid,S.S., *et al.* (1987) *Proc. Natl. Acad. Sci. USA* **18**, 6434–6437.
16. Pavlovsky,A.G., *et al.* (1989) *Dok. Akad. Nauk SSSR (USSR)* **305**, 861–864.
17. Skryabin,K.G., *et al.* (1989) In Jardetzky,O. (ed.), *Protein Structure and Engineering*. Plenum Press, New York, pp. 309–324.
18. Carter,P., Bedouelle,H. and Winter,G. (1985) *Nucleic Acids Res.* **13**, 4431–4443.
19. Kramer,W., Schughart,K. and Fritz,H.-J. (1982) *Nucleic Acids Res.* **10**, 6475–6485.
20. Schulga,A.A., *et al.* (1994) *Molek. Biol. (Russia)* **28**, 453–463.